

October 2020

# A Framework for Augmenting Building Performance Models Using Machine Learning and Immersive Virtual Environment

Chanachok Chokwitthaya  
*Louisiana State University at Baton Rouge*

Follow this and additional works at: [https://digitalcommons.lsu.edu/gradschool\\_dissertations](https://digitalcommons.lsu.edu/gradschool_dissertations)



Part of the [Construction Engineering and Management Commons](#)

---

## Recommended Citation

Chokwitthaya, Chanachok, "A Framework for Augmenting Building Performance Models Using Machine Learning and Immersive Virtual Environment" (2020). *LSU Doctoral Dissertations*. 5385.  
[https://digitalcommons.lsu.edu/gradschool\\_dissertations/5385](https://digitalcommons.lsu.edu/gradschool_dissertations/5385)

This Dissertation is brought to you for free and open access by the Graduate School at LSU Digital Commons. It has been accepted for inclusion in LSU Doctoral Dissertations by an authorized graduate school editor of LSU Digital Commons. For more information, please contact [gradetd@lsu.edu](mailto:gradetd@lsu.edu).

# A FRAMEWORK FOR AUGMENTING BUILDING PERFORMANCE MODELS USING MACHINE LEARNING AND IMMERSIVE VIRTUAL ENVIRONMENT

A Dissertation

Submitted to the Graduate Faculty of the  
Louisiana State University and  
Agricultural and Mechanical College  
in partial fulfillment of the  
requirements for the degree of  
Doctor of Philosophy

in

The Department of Engineering Science

by  
Chanachok Chokwitthaya  
B.S., Kasetsart University, 2007  
M.S., Lamar University, 2013  
December 2020

*Dedicated to my parents, Mr. Jedthabut and Mrs. Panit Chokwitthaya, and to my siblings, Mr. Warat Chokwitthaya, Mr. Nutt Choikwittaya, and Miss Piyaporn Chokevittaya for their unconditional love, support, and encouragement through all these years in the United States of America.*

## ACKNOWLEDGEMENTS

I would like to express my special appreciation and thanks to my advisor Dr. Yimin Zhu for his warmth, support, and encouragement through all the difficulties of graduate school. I am grateful to him for giving me the opportunity to be a part of this emerging, challenging, and exciting research perspective. He has been a tremendous mentor in the truest sense, and his guidance has been instrumental in shaping both my dissertation research and my overall development as a professional. Without his immense patience and encouragement throughout my PhD, this journey would have been much more grueling.

I am thankful to my co-advisor, Dr. Supratik Mukhopadhyay, for his great insights, encouragement, willingness to help, and constant support.

I also would like to thank Dr. Carol Friedland, Dr. Yongcheol Lee, and Dr. Aly Mousaad Aly for serving as my committee members. Their timely and valuable inputs and willingness to evaluate this work are greatly appreciated.

I would also like to thank everyone in the Construction Management Department, especially Dr. Charles Berryman, Dr. Marwa Hassan, and Dr. Chao Wang, Dr. Kimberley Williams, and Ms. Stephanie Heumann for their valuable input, support, and smiles, as well as the administrative staff, Ms. Jo Ann Landry, Ms. Becky Labatut, Ms. Elizabeth Green, and Ms. Lauren Zabaski for their kind assistance in so many ways.


I would like to thank all my friends in Baton Rouge, in particular Julia Gabriela Hernandez Aguilar, Arash Taghi Nezhad Bilandi, Sanaz Saeidi, Asal Emamjomeh, Robert Dibiano, Edward Collier, and all of my Thai friends, and I would like to give special thanks to Thanida Chuacharoen. All of my friends have always been great friends, patient listeners, and great sources of inspiration.

This work would not be possible without the love, affection, support, and faith of my family. I am thankful to my beloved parents and my siblings for providing me with the best education, support, and encouragement necessary to make me the person I am today. Without them, I would never have been able to achieve my goals.

Finally, I gratefully acknowledge the funding provided by the Louisiana Board of Regents Award No. LEQSF (2016-17)-RD-C-07 and the U.S. National Science Foundation Award #1640818. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Louisiana Board of Regents and the National Science Foundation.

# TABLE OF CONTENTS

ACKNOWLEDGEMENTS .....	iii
LIST OF TABLES .....	vii
LIST OF FIGURES .....	viii
ABSTRACT .....	x
CHAPTER 1. INTRODUCTION .....	1
1.1. Background .....	1
1.2. Problem Statement .....	12
1.3. Goals and Objectives .....	14
1.4. Research Hypotheses .....	14
1.5. Research Scope .....	17
CHAPTER 2. COMBINING CONTEXT-AWARE DESIGN-SPECIFIC DATA AND BUILDING PERFORMANCE MODEL TO IMPROVE BUILDING PERFORMANCE PREDICTIONS DURING DESIGN .....	19
2.1. Introduction .....	19
2.2. Research Objective .....	22
2.3. Overview of the Computational Framework .....	24
2.4. Application of the Computational Framework .....	28
2.5. Results .....	45
2.6. Limitations of the Study .....	49
2.7. Conclusion and Future Work .....	50
CHAPTER 3. AUGMENTING BUILDING PERFORMANCE PREDICTIONS DURING DESIGN USING GENERATIVE ADVERSARIAL NETWORK AND IMMERSIVE VIRTUAL ENVIRONMENTS .....	54
3.1. Introduction .....	54
3.2. Comparison of the GAN-Based Framework and the ANN-Based Greedy Algorithm .....	57
3.3. Research Objective .....	59
3.4. Overview of the Computational Framework .....	60
3.5. Application of the GAN-Based Framework .....	64
3.6. Results .....	80
3.7. Discussion .....	88
3.8. Limitations of the Study .....	92
3.9. Conclusions and Future Work .....	93
CHAPTER 4. ROBUSTNESS ANALYSIS OF A GAN-BASED FRAMEWORK FOR AUGMENTING BUILDING PERFORMANCE MODELS .....	94
4.1. Introduction .....	94
4.2. The GAN-Based Framework with Robustness Analysis .....	98
4.3. Robustness Analysis of the GAN .....	105
4.4. Results and Discussion .....	114

4.5. Limitations of the Study .....	125
4.6. Conclusions and Future Work.....	126
CHAPTER 5. CONTRIBUTIONS, CONCLUSIONS, AND FUTURE WORKS .....	128
5.1. Contributions.....	128
5.2. Conclusions .....	129
5.3. Future Works.....	130
APPENDIX. COPYRIGHT INFORMATION  .....	132
Chapter 2 Copyright.....	132
Chapter 3 Copyright.....	132
REFERENCES .....	133
VITA .....	152

## LIST OF TABLES

2.1. Contextual factors and the independent variable in the application. ....	36
2.2. Examples of virtual and auditory cues used in the IVE experiment. ....	36
2.3. The sequence of the IVE experiment. ....	37
2.4. Details of IVE experimental days. ....	39
2.5. The summary of t-test ( $\alpha = 0.05$ ) analysis ....	48
3.1. The definition of the errors to prove the hypothesis. ....	59
3.2. Variables and their values considered in the IVE experiment. ....	69
3.3. Audio cues of office tasks and switch locations. ....	71
3.4. Choices of the likelihood of switching on and their interpretation. ....	71
3.5. Groups of the <i>sub-synthetic IVE dataset</i> . ....	74
3.6. The mixture ratio ( $\alpha$ ) of computational cases. ....	80
3.7. The absolute errors to prove the hypothesis. ....	84
3.8. The summary of tests of significant difference between $E_1$ and $E_2$ . ....	85
3.9. The summary of tests of significant difference between $E_1$ and $E_2$ . ....	86
3.10. The summary of the hypothesis testing. ....	87
4.1. The summary of the parameters and their corresponding components in the application...	109
4.2. The ratios of adding AWGNs to the data in the perturbed training datasets. ....	110
4.3. The ratios of changing data of the probability of switching on to random data in the perturbed training datasets. ....	112
4.4. <i>P-values</i> corresponding to adding data noises. ....	119
4.5. <i>P-values</i> corresponding to replacing the probability of switching on with random data. ...	122
4.6. <i>P-values</i> corresponding to altering work area illuminance. ....	125



## LIST OF FIGURES

2.1. Computational framework. ....	24
2.2. Greedy heuristic algorithm. ....	26
2.3. Probability of switching on under work area illuminance of the Hunt model. ....	30
2.4. The physical environment. ....	31
2.5. The IVE configuration. ....	31
2.6. The layout of the physical environment and the locations of the sensors. ....	32
2.7. Probability of switching on under work area illuminance ....	33
2.8. Diagram of factors included in the IVE experiment. ....	37
2.9. Probability of switching on under work area illuminance (HMM). ....	41
2.10. Scheme of ANNs of the computational framework. ....	42
2.11. Observations of updated BPMs obtained from the computational framework using various mixture ration ( $\alpha$ ) ....	45
2.12. The hypothesis testing. ....	47
2.13. Plots of probability of switching on obtained from the synthetic IVE testing dataset and prediction of the ANN. ....	49
3.1. The GAN-based framework. ....	60
3.2. The Hunt model. ....	65
3.3. The model of the office. ....	67
3.4. The IVE configuration. ....	67
3.5. A participant exploring the virtual office and selecting a need of turning the light on. ....	72
3.6. The Da Silva model. ....	74
3.7. Scheme of the generative adversarial network (GAN) of this application. ....	75
3.8. Scheme of the ANN-based greedy algorithm framework. ....	77

3.9. The scheme of the ANN of this application.....	78
3.10. Efficient greedy algorithm. ....	79
3.11. The context-aware design-specific data.....	80
3.12. Plots of the mean of the <i>augmented BPMs</i> , the mean of the <i>updated BPMs</i> , and the <i>performance target</i> .....	82
4.1. The GAN-based framework with the robustness analysis. ....	98
4.2. <i>Augmented BPMs</i> corresponding to non-perturbed training dataset.....	115
4.3. <i>Augmented BPMs</i> corresponding to adding data noises. ....	117
4.4. <i>K-S statistics</i> corresponding to adding data noises. ....	119
4.5. <i>Augmented BPMs</i> corresponding to replacing the probability of switching on with random data.....	121
4.6. <i>K-S statistic</i> corresponding to replacing the probability of switching on with random data. ....	122
4.7. <i>Augmented BPMs</i> corresponding to altering work area illuminance.....	124
4.8. <i>K-S statistic</i> corresponding to altering work area illuminance. ....	125

## ABSTRACT

Building performance models (BPMs), such as building energy simulation models, have been widely used in building design. *Existing BPMs* are mainly derived using data from existing buildings. They may not be able to effectively address human-building interactions and lack the capability to address specific contextual factors in buildings under design. The lack of such capability often contributes to the existence of building performance discrepancies, i.e., differences between predicted performance during design and the actual performance.

To improve the prediction accuracy of *existing BPMs*, a computational framework is developed in this dissertation. It combines an *existing BPM* with context-aware design-specific data involving human-building interactions in new designs by using a machine learning approach. Immersive virtual environments (IVEs) are used to acquire data describing design-specific human-building interactions, a machine learning technique is used to combine data obtained from an *existing BPM*, and IVEs are used to generate an *augmented BPM*.

The potential of the framework is investigated and evaluated. An artificial neural network (ANN)-based greedy algorithm combines context-aware design-specific data obtained from IVEs with an *existing BPM* to enhance the simulations of human-building interactions in new designs. The results of the application show the potential of the framework to improve the prediction accuracy of an *existing BPM* evaluated against data obtained from the physical environment. However, it lacks the ability to determine the appropriate combination between context-aware design-specific data and data of the *existing BPM*. Consequently, the framework is improved to have ability to determine an appropriate combination based on a specified *performance target*. A generative adversarial network (GAN) is used to combine context-aware design-specific data and data of an *existing BPM* using the *performance target* as guide to generate an *augmented BPM*.

The results confirm the effectiveness of this new framework. The performance of the *augmented BPMs* generated using the GAN-based framework is significantly better than the *updated BPMs* generated using the ANN-based greedy algorithm.

The framework is completed by incorporating a robustness analysis to assist investigations of robustness of the GAN regarding the uncertainty involved in the input parameters (i.e., an *existing BPM* and context-aware design-specific data).

Overall, this dissertation shows the promising potential of the framework in enhancing performance of BPMs and reducing performance discrepancies between estimations made during design and in performance in actual buildings.

# CHAPTER 1. INTRODUCTION

## 1.1 Background

Building performance models (BPMs) are a tool assisting designers to analyze, understand, and optimize the performance of non-existing or future buildings. Various BPMs have been created and improved over recent decades for various purposes, such as predicting human-building interactions with building components [1][2], supporting building climate control systems [3][4], as well as supporting automated electric lighting and blind systems [5]. Most BPMs are created by collecting data on human-building interactions with building systems, such as heating, ventilation, and air conditioning (HVAC) along with artificial lights, blinds, windows, and appliances (e.g., televisions, audio systems, and refrigerators). Such data are collected by using conventional human-building interactions research methods (e.g., questionnaires, field monitoring, and laboratory experiments), which mainly rely on human-building interactions in existing buildings. However, human-building interactions are influenced by several contextual factors, such as the occupant's sense of control, building characteristics, building service systems and operations, as well as climates, all of which make developing BPMs challenging [6]. Contextual factors are hidden factors related to the human-building interactions, and multiple contextual factors may drive the interactions simultaneously [7]. Additionally, many contextual factors vary dynamically, such as weather [8]. For example, human-building interactions with light switches may depend on lighting preferences, light control positions, and weather. Therefore, data of human-building interactions obtained from the conventional human-building interactions research methods may totally differ from human-building interactions in a new building under design. Using BPMs constructed by using such

data to estimate building performance during design may contribute to significant discrepancies between estimated and actual performance [9].

Immersive virtual environments (IVEs) are alternative research methods that can be used to observe human-building interactions in design specifics and to collect context-aware design-specific data, such as human-building interactions in specific building design contexts (e.g., building components as well as indoor and outdoor environments). This approach is different from conventional methods that use existing buildings and their contexts to estimate or simply assume human-building interactions. However, IVEs have many limitations, such as short experiment sessions, small data samples, and negative impacts on participants (e.g., cybersickness), all of which make IVE-based experiments limited. The limitations of IVEs cause difficulties in continuously collecting human-building interaction data in virtual environments for extended periods of time. IVE-based experiments for data collection are often highly focused and event/purpose-driven. Therefore, data collected using IVEs are not as comprehensive as data collected in reality using conventional occupancy data collection approaches (e.g., sensing, field studies, and surveys). Consequently, it is difficult to create comprehensive BPMs as general models if only using data from IVE experiments. It is more feasible to bias a general model using observational data to fit a particular design than to produce a general BPM from only observational data obtained from IVE experiments. Therefore, the author proposes a framework to combine *existing BPMs* that are constructed by using the conventional human-building interactions research methods with observational data acquired from IVE experiments that simulate a building during design (i.e., context-aware design-specific data). The framework incorporates the advantages of both *existing BPMs* and context-aware design-specific data. Specifically, the framework preserves the general predictive power of an *existing BPM* while

addressing specific human-building interactions in the context of a new design identified by designers or engineers. The framework produces a more representative BPM specific to a building under design than an *existing BPM*, which improves prediction accuracy. The framework uses machine learning techniques [i.e., artificial neural networks (ANNs) and generative adversarial networks (GANs)] as computations to combine an *existing BPM* with context-aware design-specific data acquired from IVE experiments that simulate the building during design. Lastly, robustness analysis of computations is included to complete the framework. The ability to understand the robustness allows users of the framework to reduce risk and gain confidence during using the framework.

In sections 1.1.1 to 1.1.6, the author provides a comprehensive review of relevant literature including BPMs, the impact of human-building interactions on building performance, IVEs for observing human-building interactions during design, machine learning for augmenting an *existing BPM* with knowledge of context-aware design-specific data to generate an *augmented BPM*, and robustness of computations for estimating building performance.

### **1.1.1 Building Performance Models (BPMs)**

BPMs offer several advantages to assist designers during building design. BPMs are mainly constructed based on deterministic correlations between physical quantities, such as temperature, illuminance, occupancy, and status of building components, such as light switches, blinds, and windows. Several examples of developments and uses of BPMs are summarized in this section.

Hunt [1] developed a BPM for predicting manual lighting control based on a switch-on probability and minimum working area illuminance. The BPM was developed by using data collected in a field study in which sensors were installed in experimental offices to capture

occupant interactions with switches for artificial lighting. He developed the BPM by using probit analysis. Nicol [10] developed BPMs to predict human-building interactions with the usage of windows, lighting, blinds, heaters, and fans based on outdoor temperature in naturally-ventilated buildings from survey data. Probit analysis was used to find the relationship between human-building interactions and outdoor temperature. Newsham [2] developed and improved a computer-based thermal model “FENESTRA” by including an algorithm to describe manual blind operation and light switching drawn from Hunt’s model. From results of his model, he suggested that incorporating algorithms of occupant behavior into building thermal models could significantly affect predictions of building energy consumption. Reinhart [5] proposed an algorithm to determine manual and automatic electric lighting energy demand called “Lightswitch-2002”. It was integrated into many simulation programs, such as design support tool (Lightswitch Wizard), lighting simulation (DAYSIM), and whole building energy simulation (ESP-R). The algorithm includes an occupancy model that considers the profiles of the occupants and the minimum working area illuminance similar to Hunt’s approach along with a dynamic daylight simulation to predict electric lighting demand. The algorithm considers within-day switch-on probability in addition to the probability of switching the light on upon arrival. Similarly, Gunay, et al. [11] formulated BPMs for an adaptive lighting and blind control algorithm. Their BPMs included concurrent solar irradiance as an additional predictor for occupant lighting preferences along with minimum working area illuminance and intermediate occupancy in other works.

The major occupant behavior research methods applied to construct BPMs are questionnaires, field monitoring, and laboratory experiments. Most of them use human-building interactions from specific occupants and contexts. However, the diversity of occupants and



contexts may cause human-building interactions to vary substantially from one building to another. As a result, predicted results from BPMs during design often show discrepancies when compared with the building performance of the actual buildings after they are built.

### **1.1.2 Human-Building Interactions and Their Impact on Building Performance**

Recent studies have suggested that human-building interactions have a significant impact on building energy consumption, and they are one of the major contributing factors to large uncertainties in building performance [12][13]. A large body of research has been dedicated to study the impact of human-building interactions on building performance. For instance, Clevenger and Haymaker [14] applied building energy simulation to study the impact of human-building interactions. They showed that the predicted energy consumption changed by at least 150% if the maximum and minimum values of occupant-related inputs were applied to the simulations. Santin, et al. [15] applied historical survey data from 15,000 houses and 3 years of energy usage to conduct a statistical analysis related to human-building interactions affecting energy consumption. They reported that human-building interactions contributed to 7.2% of energy consumption variation. Similar results were found by Kavousian, et al. [16] and D'Oca, et al. [17]. Human-building interactions are one of the most significant factors that cause variations in building energy consumption even if building envelopes, functions, and environments are identical. Hence, human-building interactions are a key factor that should be considered during building design to satisfy the functional purposes of the building.

According to the Merriam-Webster dictionary, context is defined as “the interrelated conditions in which something exists or occurs”. Factors, often referred to as contextual factors, are used to describe or model the conditions. There are evidences that human-building interactions are driven by contextual factors [18], including building conditions [19]. For

instance, the illuminance of a work area is one of the contextual factors that significantly determines the lighting usage of the occupants. Multiple contextual factors may influence human-building interactions simultaneously [20]. Understanding the impact of contextual factors on human-building interactions is one of the keys to potentially enhance the accuracy of BPMs.

Human-building interactions are driven by various contextual factors, which may include type of building, time, occupants, systems, and environments [18]. For instance, indoor air quality influences window usage, work area illuminance influences blind and light usage, and indoor temperature influences thermostat usage [21]. Multiple contextual factors may drive human-building interactions simultaneously. Nicol, et al. [20] stated that in addition to indoor temperatures, contextual factor that influenced occupants to seek thermal comfort included clothing, the metabolic rates of occupants, skin moisturizer, and air movement. To satisfy their thermal comfort, occupants might also interact with multiple building systems, such as blinds, thermostats, windows, and lighting. Hong and Lin [12] studied human-building interactions in multiple single offices. They found that an occupant's attitude towards consumption (e.g., austerity, standard, and wasteful) was one of the major contextual factors that drove the behaviors of occupants. Human-building interactions have a significant impact on building performance, and they are influenced by several contextual factors. Thus, considering contextual factors is necessary in human-building interactions studies, especially those of human-building interactions for building during design.

### **1.1.3 Conventional Human-Building Interactions Research Methods**

Many research studies have been conducted to improve building designs by using BPMs. Three common approaches applied to study human-building interactions for developing BPM are questionnaires, field monitoring, and laboratory experiments.

Questionnaires are a common method to study human-building interactions.

Questionnaires can be administered to subjects that researchers want to investigate and are feasible in large-scale experiments. Attia, et al. [22] used questionnaires to collect human-building interaction data related to household device usage in residential apartments in various areas in Egypt. They applied the results of the questionnaires to construct benchmarks for building energy simulations. Feng, et al. [23] used questionnaires to observe human-building interactions related to air conditioning (AC) patterns. The questionnaires were used to categorize human-building interactions regarding behavior about switching AC on or off. Questionnaires have been used to research multiple aspects of interest in several places simultaneously. Nicol [10] studied human-building interactions regarding usage of windows, lighting, blinds, heating, and fans by using questionnaires in the United Kingdom, Sweden, France, Portugal, and Greece. Although questionnaires provide various advantages, an important disadvantage is that they are not able to quantitatively capture the relationship between contexts and human-building interactions.

The field monitoring method has been used in many studies of usage, such as light switching [24], window opening [25], energy usage for space and water heating [15], the heating set-points of occupants [26], human-building interactions regarding shading and lighting [27], and human-building interactions regarding plug-in equipment usage [28]. One of the advantages of this method is that the collected data are a longitudinal time series, and acquiring large samples is possible because multiple sensors are deployed. Another advantage is that the method is capable of providing quantitative relationships between the occupant behaviors and the contexts. However, this method has many limitations. Firstly, data are often collected at time intervals, e.g., every 30 minutes. Thus, some critical events may be unobserved if they occur

between the intervals. Also, other equipment may interfere with sensors and distort information of human-building interactions and contexts. Finally, many assumptions with respect to human-building interactions and design contexts, such as occupant schedules, variables that drive interactions, and purposes of occupant response to building systems have to be made to derive the BPMs.

Using laboratory experiment methods [29][30], human-building interactions have been studied in controlled environments, e.g., the Zero Energy Lab (ZØE) at the University of North Texas [31]. These methods have various advantages. Firstly, a wide range of scenarios can be simulated under controlled conditions. Furthermore, a variety of monitored data can be obtained, namely physical, physiological, and psychological data. An important disadvantage is that observations are often limited to laboratory conditions and contexts. Conclusions drawn from observations based on such studies may be difficult to extrapolate to different conditions and contexts.

Clearly, the three methods described above typically rely on observations of human-building interactions in existing buildings. Because human-building interactions are context sensitive, findings from such observations can certainly contain biases and uncertainties. Thus, applying such findings to new designs may lead to significant variation in predictions.

#### **1.1.4 IVEs as an Alternative for Collecting Human-Building Interactions During Design**

Due to the limitations of conventional occupant research methods, the author suggests an alternative method to study and observe human-building interactions during building design by employing IVEs. IVEs are multisensory computer-generated environments that provide the users with a sense of being mentally immersed or present in the simulations. Generally, IVEs may be classified into (1) fully immersive [e.g., head-mounted display-based (HMD)], (2) semi-

immersive (e.g., Cave Automatic Virtual Environment), and (3) non-immersive (e.g., computer desktop) integrated with other immersion capabilities. In building design, IVE applications provide many significant advantages, e.g., 3D visualizations that can be manipulated in real-time and used to virtually explore building components and construction processes [32], revision tools that investigate and address issues [33], and support for decision-making [34]. Moreover, IVEs have been applied to study many human-building interactions and energy usages. For instance, Heydarian, et al. [35] studied occupant lighting preferences in a single office by using IVEs. The participants were exposed to various design contexts and asked to adjust lights and shades in IVEs. The ability of IVEs to replicate field experiences in occupant lighting behavior has been validated. Saeidi, et al. [36] validated results obtained from IVE experiments with data collected in actual buildings and concluded that lighting stimuli in the IVE were able to produce similar behavioral responses. Niu, et al. [37] developed a framework to integrate building design with IVE to help building designers capture human-building interactions and identify contextual patterns. They concluded that integrating building design with IVE using their framework facilitated understanding of human-building interactions and identification of design contexts that guided the occupants to act according to the intentions of the designers. Recently, Saeidi, et al. [38] used IVEs to capture occupant lighting preferences in a single office based on several exposed design contexts and cues. They validated the results of human-building interactions from IVE experiments with the actual data from the physical environment. The results showed good agreement between the IVE experiments and the actual data from the physical environment.

These works show the capabilities of IVEs in studying human-building interactions, including abilities (1) to retain the control of an experimental environment, (2) to design

experimental environments, and (3) to acceptably maintain the realness of physical environment. Moreover, IVEs can be integrated with programming techniques to allow participants to interact with components in IVEs as well as to allow researchers to capture human-building interactions responding to design contexts and to elicit contextual factors. IVEs have been proven to provide several exceptional advantages as an alternative human-building interactions research method to support the development of BPMs and may play a role in reducing the performance gaps between estimations and actual buildings.

### **1.1.5 Machine Learning**

The framework in the present study uses machine learning algorithms as computations to augment an *existing BPM* using context-aware design-specific data obtained from an IVE experiment simulating a new design. Two machine learning algorithms are involved in the framework, namely ANNs and GANs. These machine learning algorithms are reviewed in this section.

#### ***Artificial Neural Networks (ANNs)***

ANNs [39] are one of the most widely used machine learning techniques in various disciplines, such as building performance [40][41], building energy consumption [42][43], and occupancy comfort [44][45]. Jing, et al. [46] used an ANN with a backpropagation algorithm to predict building energy consumption based on various contextual factors, such as building age, internal floor areas, carbon dioxide emissions, and building component energy consumption. They also compared the performance of the ANN with multiple regression approaches and concluded that the ANN generally provided more accurate forecasts than multiple regression. Ahmad, et al. [47] compared the performance of an ANN with a backpropagation algorithm to a random forest for building energy prediction models. They claimed that the ANN performed

marginally better than the random forest method. In human behavior studies, Fang, et al. [47] and Mehr, et al. [48] used ANNs with backpropagation algorithms to predict human activities in buildings based on date and time. According to the literature, ANNs have strong abilities to learn prior knowledge (or distributions) and to make predictions corresponding to the given prior knowledge. Therefore, the first stage of the present study used an ANN to combine prior knowledge obtained from an *existing BPM* with context-aware design-specific data obtained from an IVE simulating a building under design to generate an *augmented BPM* with improved accuracy.

### ***Generative Adversarial Networks (GANs)***

Deep learning has grown in popularity in recent years [49][50][51][52][53]. GANs were proposed by Goodfellow, et al. [50]. GANs have been successfully used in various domains [54], especially image synthesis. Ledig, et al. [55] used GANs to learn and recover photorealistic textures from downsampled images. They proposed super-resolution GANs (SRGANs) that can estimate photorealistic super-resolution images with high upscaling factors. Radford, et al. [56] introduced deep convolutional generative adversarial networks (DCGANs) for generating realistic and high-resolution images. They showed that DCGANs outperformed other unsupervised algorithms (k-means, random forest, and transductive support vector machines). Wang and Gupta [57] introduced style and structure GANs (S2-GANs). These addressed structure and style in the image generation process. S2-GANs have the ability to generate photorealistic high-resolution images in addition to having a more robust and stable training method compared to standard GANs. Wu, et al. [57] introduced 3D-GANs that are capable of generating 3D objects by combining volumetric convolutional networks with GANs. These previous studies have demonstrated the abilities of GANs to produce synthetic images that are close to real

images from arbitrary image clues (noises). Thus, the author used a GAN to produce an *augmented BPM* corresponding to a *performance target* by combining *existing BPMs* with the knowledge on human-building interactions responding to contextual factors in new building design (i.e., context-aware design-specific data).

### **1.1.6 Robustness Analysis of the Computation in the Framework**

The majority of computations for estimating building performance, such as building performance simulations (e.g., EnergyPlus, Autodesk Revit, ESP-r, and DesignBuilder) and machine learning techniques (e.g., ANNs, random forests, and support vector machines) are traditionally treated as black-box computations. They take given input parameters to analyze and estimate building performance. Input parameters are often subject to uncertainty since their data are usually obtained by using measurements and experiments (e.g., field observations and surveys) in which the existence of uncertainty may not be avoidable. The uncertainty causes reductions of robustness and increases the risks of using the computations. Therefore, incorporating robustness analysis in the framework assists investigations of the ability of the computation to handle such uncertainty and remain robust during execution.

## **1.2 Problem Statement**

### **1.2.1 Performance Discrepancy Between Estimations and Actual Buildings**

BPMs are tools that assist designers to estimate performance of future buildings during design. However, the performance discrepancy, i.e., the difference between predicted performance during design and actual building performance always exists. The lack of ability to accurately model human-building interactions in *existing BPMs* is among the factors contributing to the discrepancy. In the first stage of this study, the author proposes a framework that potentially reduces performance discrepancy.



### 1.2.2 Appropriate Mixture of an *Existing BPM* and Context-Aware Design-Specific Data

In the first stage, the author proposes an ANN-based greedy algorithm framework to potentially reduce the performance discrepancy between estimations during building design and actual building performance. The framework combines an *existing BPM* and context-aware design-specific data acquired from IVE experiments simulating the building during design to generate an *augmented BPM*. However, the framework needs several assumptions and trials to obtain an appropriate *augmented BPM*. It may often not be possible to obtain an appropriate *augmented BPM*. Inappropriate combination may cause poor estimations when the *augmented BPM* is used to predict the performance of buildings. Therefore, the second stage of the present study improves the capability of the framework to be able to determine the appropriate combination without trial and error. The improved framework is called the GAN-based framework.

### 1.2.3 Robustness of a Computational Component of the Framework

No physical quantity can be measured and experimented on without involving uncertainty [58]. Uncertainty may reduce the robustness of the computational component of the GAN-based framework. If the computation is not robust, the framework may generate an inappropriate *augmented BPM*. Using such *augmented BPM* to assist decision making during design increases likelihood of errors especially discrepancy between estimated and actual building performance. To understand and evaluate uncertainty impacting robustness of the computation of the framework, the third stage of the present study provides a robustness analysis of the GAN.

### 1.3 Goals and Objectives

This research contributes to the development of a novel approach for reducing the discrepancy in building performance between estimations made during building design and actual building performance, thereby improving future building designs.

The research objectives are described below:

1. To establish a framework to augment an *existing BPM* using context-aware design-specific data acquired from an IVE simulating a building under design to improve the prediction accuracy of BPMs.
2. To enhance the ability of the framework for allowing users to appropriately generate an *augmented BPM* by using a *performance target* as a guide.
3. To analyze the robustness of the computation of the framework to allow users to gain confidence in making decisions during using the framework.

To achieve these objectives, the author first identifies the components that contribute to proving the research hypothesis, and these are organized into chapters that show the connection and contribution of each component.

### 1.4 Research Hypotheses

Several research hypotheses are listed below.

#### 1.4.1 Hypothesis 1

The author hypothesized that the ANN-based greedy algorithm framework could significantly improve the prediction accuracy of BPMs during building design. To test the hypothesis, absolute errors were calculated including 1) the absolute error that measures the discrepancy between the predicted outcomes of an *existing BPM* and actual building

performance data ( $E_1$ ) and 2) the absolute error that measures the discrepancy between the predicted outcomes of an *augmented BPM* and the actual building performance data ( $E_2$ ).

The formulas for calculating  $E_1$  and  $E_2$  are shown in Equation (1-1) and (1-2), respectively:

$$E_1 = | \text{The predicted outcomes of an } \textit{existing BPM} - \text{actual data} | \quad (1-1)$$

$$E_2 = | \text{The predicted outcomes of an } \textit{augmented BPM} - \text{actual data} | \quad (1-2)$$

Both errors were used to test the hypothesis as shown in the description below:

$$H_0: \text{mean of } E_1 - \text{mean of } E_2 = 0$$

$$H_1: \text{mean of } E_1 - \text{mean of } E_2 > 0$$

A one-tailed *t-test* ( $\alpha = 0.05$ ) was applied to investigate the statistically significant difference between the means of  $E_1$  and  $E_2$ .

#### 1.4.2 Hypothesis 2

In the first stage, the author randomly generated the combination of context-aware design-specific data and an *existing BPM* without knowing the appropriate mixture between them. In the improved framework (the GAN-based framework), the author reliably determined the appropriate combination by introducing a *performance target* as a guide for the combination. The author hypothesized that the GAN-based framework would generate significantly better *augmented BPMs* compared to the ANN-based greedy algorithm proposed in the first stage.

To define a hypothesis, the GAN-based framework generated an *augmented BPM*, and the ANN-based greedy algorithm generated an *updated BPM* [59]. The comparison was based on the hypothesis that the *augmented BPM* would be more accurate than the *updated BPM*. The absolute error measured discrepancy between an *updated BPM* and the *performance target* ( $E_1$ ) and the absolute error measured discrepancy between the *augmented BPM* and the *performance*

*target* ( $E_2$ ) were calculated using Equation (1-3) and (1-4), respectively. They were used to develop the hypothesis.

$$E_1 = \quad | \text{The predicted outcomes of an } \textit{updated BPM} - \text{a performance target} | \quad (1-3)$$

$$E_2 = \quad | \text{The predicted outcomes of an } \textit{augmented BPM} - \text{a performance target} | \quad (1-4)$$

To test the performance of the *augmented BPM*, the hypothesis was defined as follows:

$$H_0: \text{mean of } E_1 - \text{mean of } E_2 = 0$$

$$H_1: \text{the null hypothesis is not true}$$

A *t-test* ( $\alpha = 0.05$ ) was applied to investigate the statistically significant difference of the performance between the *augmented BPM* and the *updated BPM*.

### 1.4.3 Hypothesis 3

In the last stage, a robustness analysis was performed to determine whether the GAN produced a resilient *augmented BPM*. If the GAN for particular assumptions about variability in inputs (e.g., uncertainty of the involved parameters) produced a similar *augmented BPM*, the GAN was considered robust for those assumptions. The robustness analysis identified whether the GAN remained robust when input datasets were uncertain. An *augmented BPM* generated by the GAN trained on a non-perturbed training dataset ( $A_{non-perturbation}$ ) was considered as the baseline. The robustness analysis determined differences between  $A_{non-perturbation}$  and an *augmented BPM* generated by the GAN trained on a perturbed training dataset ( $A_{perturbation}$ ). If  $A_{perturbation}$  was not significantly different from  $A_{non-perturbation}$ , the GAN was considered robust. Accordingly, the hypothesis was defined as follows:

$$H_0: A_{perturbation} - A_{non-perturbation} = 0$$

$$H_1: A_{perturbation} - A_{non-perturbation} \neq 0$$

## 1.5 Research Scope

Although the computational framework can be broadly applied to most BPMs, the validation of the framework and tests of hypotheses were limited to the prediction of human-building interactions regarding lighting usage. Lighting usage was selected to validate the framework and test hypotheses for several reasons. Lighting is one of the largest energy consumers in artificially lit buildings and consumes 5-15% of total electric energy in buildings [60]. Through various design solutions, the amount of lighting-related energy used may be significantly reduced [61]. BPMs for predicting human-building interactions regarding lighting usage are one of the alternatives to support design solutions that contribute to reduce lighting-related energy consumption. Furthermore, lighting is the most mature IVE capability, and IVEs have shown their ability to capture human-building interactions regarding lighting usage in many studies [62][35][63]. However, more complex IVE systems with additional sensory modalities are necessary for an IVE to fully capture other human-building interactions.

Contextual factors are the factors that are not included as independent variables in *existing BPMs*. They indirectly influence human-building interactions and are determined by situational factors that are associated with building contexts. The contextual factors considered in the present study were factors related to physical environments (e.g., illuminance) and buildings (e.g., office configurations, locations of light switches, and office tasks). Other factors such as psychological, physiological, and social ones were not considered as contextual factors in the present study.

The robustness of the computation of the framework in the present study was quantified in terms of robustness regarding the uncertainty involved in an *existing BPM* and context-aware

design-specific data. Details of elements that cause uncertainty such as participants, environments, and tools as well as computational systems were not considered in the study.

## **CHAPTER 2. COMBINING CONTEXT-AWARE DESIGN-SPECIFIC DATA AND BUILDING PERFORMANCE MODEL TO IMPROVE BUILDING PERFORMANCE PREDICTIONS DURING DESIGN**

### **2.1 Introduction**

According to the International Energy Agency, buildings in developed countries consume up to 40% of their total energy [64]. The significant consumption of fossil fuel-based energy has caused negative environmental impacts, such as ozone layer depletion, global warming, and climate change [65]. In addition, studies have confirmed that decisions made during the design phase significantly influence energy efficiency during building operations (e.g., [66][67]). Thus, improvements in decision support for building energy efficiency during design can contribute to the reduction of building energy consumption and enhancement in building energy performance [68].

BPMs are decision-support tools assisting designers and engineers to understand, analyze, and optimize building performance during design. There are different types of BPMs, including simulation models of whole-building energy consumption [69], predictive models for the performance of building systems, such as space heating [42] and air quality [70], as well as models of occupant interactions with building components, such as light switches, blinds, windows, and thermostats [71][72]. A number of research studies (e.g., [1][2] [10][11]) have successfully included human-building interactions in building performance modeling and prediction. Generally, such BPMs are constructed by collecting data of human-building

---

This chapter was previously published as C. Chokwitthaya, Y. Zhu, R. Dibiano, and S. Mukhopadhyay, “Combining context-aware design-specific data and building performance models to improve building performance predictions during design,” *Automation in construction*, vol. 107, p. 102917, 2019, doi: 10.1016/j.autcon.2019.102917. Reprinted by permission of Elsevier.

interactions and finding correlations between independent variables (e.g., temperature, illuminance, solar irradiance, and occupancy status) and dependent variables (e.g., human interactions with building components, such as light switches, blinds, and windows). For instance, BPMs for predicting artificial lighting usage (e.g., [1][5]) considered work area illuminance as an independent variable to predict whether occupants turned on artificial lighting on arrival. In addition to work area illuminance, the location of light switches [73], interiors layouts [74], and occupancy statuses [75] may influence occupant interactions with light switches. Human-building interactions (e.g., occupant responses to contextual factors and occupant habitual behaviors) are highly context-dependent. The contexts of existing buildings from which data for developing BPMs are obtained often differ from the context of a new building under design. Thus, the application of BPMs in a different context may introduce significantly large variances and contribute to the discrepancies between predictions during design and actual performance during operations [9]. An alternative is to customize existing BPMs to address the context of buildings under design.

According to the Merriam-Webster dictionary, context is defined as “the interrelated conditions in which something exists or occurs.” Contextual factors are often used to describe or model such interrelated conditions. In the present study, factors that may influence human-building interactions but are ignored in *existing BPMs* were considered as “*contextual factors*” in relation to the *existing BPMs*. There are evidences that human-building interactions are driven by contextual factors, such as building conditions [18][19]. Multiple contextual factors may influence human-building interactions simultaneously [20]. Therefore, having the capability to consider human-building interactions in a specific context, such as the context embodied in a new design, may be one of the keys to significantly enhancing the accuracy of BPMs.



IVEs are multisensory computer-generated environments and have been effectively applied to various researches in building design and engineering, such as emergency evacuation [76][77], building designs [78][79][80], and occupant behavior predictions [81][82]. IVEs have also been applied to studies related to human-building interactions and energy usage. For instance, Heydarian, et al. [35] studied occupant lighting preferences in a single office using IVEs. Saeidi, et al. [36] validated occupant light usage behavior in IVEs and showed that IVEs were capable of replicating field experiences. Niu, et al. [37] developed a framework to integrate building designs with IVEs to help building designers capture occupant preferences and identify context patterns. Studies have shown that human-building interactions are context-dependent. Since buildings under design do not physically exist, human-building interactions with buildings under design cannot be directly observed. To capture such human interactions, IVEs are proxies of reality that allow designers or researchers to observe such interactions. Overall, the main advantages of applying IVEs during design include replicating the context of buildings under design [36], allowing designers or researchers to control experimental conditions and to apply desired experimental contextual factors [38]. Therefore, IVEs have the potential to support designers or researchers to observe human-building interactions in simulated building context during design. However, IVEs have many limitations, such as short experiment sessions, small data samples, and negative impacts on participants (e.g., cybersickness) [83], all of which make IVE-based experiments limited. The limitations of IVEs cause difficulties in continuously collecting human-building interaction data in virtual environment for extended periods of time. IVE-based experiments for data collection are often highly focused and event/purpose-driven, so data collected using IVEs are not as comprehensive as data collected in real buildings using conventional occupancy data collection approaches (e.g., sensing, field studies, and surveys).

Consequently, it is difficult to create comprehensive BPMs as general models if only using data obtained from IVE experiments. Thus, it is more feasible to bias a general model using the data to fit a particular design than to produce a general BPM only from observational data obtained from IVE experiments.

To enhance the prediction accuracy of *existing BPMs*, the author created a computational framework that combines an existing model with observational data obtained from IVE experiments. Specifically, the framework preserves the general predictive power of an existing BPM while addressing specific human-building interactions in the context of a new design identified by designers or researchers. As a result, the framework produces a more representative BPM specific to a building under design to improve prediction accuracy. The framework produces a more representative BPM specific to a building under design than an *existing BPM*, which improves prediction accuracy.

In the following, the author states the research objective, provides an overview of the computational framework, and then presents the application of the framework to a single occupancy office. Results, conclusions, and future work are then discussed based on the application.

## **2.2 Research Objective**

The objective of this chapter is to determine if the computational framework can potentially improve the prediction accuracy of BPMs during design. To achieve the objective, the author applied the computational framework to the study of a single occupancy office. The framework produced an optimal BPM, which was called the *augmented BPM*. An application was designed to verify the effectiveness of the framework using the *augmented BPM*, which was achieved by testing a hypothesis.

The author hypothesized that the computation framework could significantly improve the prediction accuracy of BPMs during design. To test this hypothesis, absolute errors were analyzed, including 1) the absolute error that measured the discrepancy between the predicted output of an *existing BPM* and the actual building performance data ( $E_1$ ) and 2) the absolute error that measured the discrepancy between the predicted output of the *augmented BPM* and the actual building performance data ( $E_2$ ).

The formulas for calculating  $E_1$  and  $E_2$  are shown in Equation (2-1) and (2-2), respectively:

$$E_1 = | \text{The predicted outcome of an } \textit{existing BPM} - \text{actual data} | \quad (2-1)$$

$$E_2 = | \text{The predicted outcome of the } \textit{augmented BPM} - \text{actual data} | \quad (2-2)$$

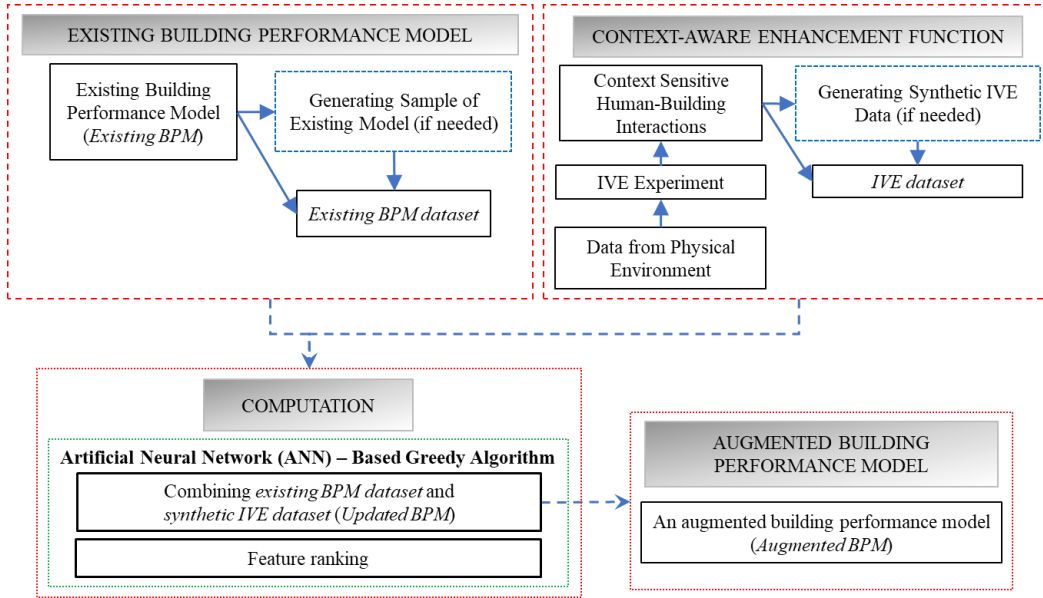
Both errors were used to test the hypothesis as shown below:

$$H_0: \text{mean of } E_1 - \text{mean of } E_2 = 0$$

$$H_1: \text{mean of } E_1 - \text{mean of } E_2 > 0$$

A one-tailed *t-test* ( $\alpha = 0.05$ ) was applied to investigate the statistically significant difference between the mean of  $E_1$  and the mean of  $E_2$ .

## 2.3 Overview of the Computational Framework



**Figure 2.1.** Computational framework.

The computational framework comprised of four main elements (Figure 2.1): (1) an *existing BPM*, (2) context-aware design-specific data, (3) computation, and (4) an *augmented BPM*. In theory, the framework is parametric and does not have any restrictions on the input datasets because it only combines an *existing BPM* with context-aware design-specific data. Datasets associated with an *existing BPM* and context-aware design-specific data were inputs to the framework. Since the framework could be applied to any existing BPM and context-aware design-specific data, datasets applied in the framework do not need to be specified. In practice, there is a need to consider the cost associated with acquiring data using IVE experiments. In the following, details of components are discussed.

### 2.3.1 Existing Building Performance Model

An *existing BPM* represents a model that already exists, but it does not necessarily capture the important contextual factors of a new building design. For example, the Hunt model uses illuminance to predict the status of light switches. While it may be effective in general, the model cannot accurately predict artificial lighting usage if a new design has a very different occupancy pattern from the pattern that the Hunt model was implicitly based on [1].

To generate a dataset using an *existing BPM*, the computational framework in the present study provided a tool to generate such a dataset using statistical approaches, e.g., Monte Carlo (MC) simulations. The dataset was called the *existing BPM dataset*.

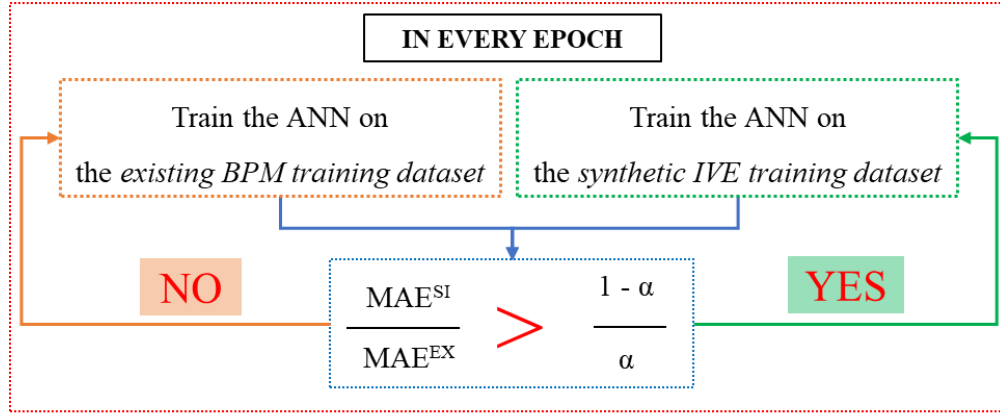
### 2.3.2 Context-Aware Design-Specific Data

Context-aware design-specific data describe contextual factors in a set of specific events of a new design. For example, a designer may believe that how occupants interact with light switches on arrival in summer mornings is critical to a design purpose. This contextual factor needs to be explicitly included in a BPM to observe occupant behaviors. In the present study, context-aware design-specific data were used to modify an *existing BPM* so that the BPM better reflected the context of a new building under design. To generate context-aware design-specific data, an IVE was used as a tool to collect context-specific data of a design, e.g., occupant's use of artificial lighting on a clear summer day. IVE-based experiments are often conducted with small samples in short period of time, leading to small IVE datasets [38][84][85]. To overcome this limitation, the framework provided an alternative solution to the small sample size issue by applying a Hidden Markov Model (HMM) statistical data learning approach to generate a large dataset, which was called a *synthetic IVE dataset*.

### 2.3.3 Computation

Two major parts are included in the computation component, i.e., combining the *existing BPM dataset* and the *synthetic IVE dataset* as well as feature ranking.

#### *Combining the Existing BPM Dataset and the Synthetic IVE Dataset*



**Figure 2.2.** Greedy heuristic algorithm.

The purpose of combining the *existing BPM dataset* and the *synthetic IVE dataset* is to generate the *augmented BPM*. The author applied an ANN [39] to this process. Compared to other methods such as Bayesian networks, regression models, Kalman filter, and other graphical models, ANNs provide several advantages for the computational framework. In many applications, ANNs have been proven to be more accurate, flexible, and consistent in predictions than Bayesian networks [86], regression models [87] [88] [89], Kalman filter [90], and K-means [91]. ANNs have the capability to combine multiple datasets during training [92], e.g., the *existing BPM dataset* and the *synthetic IVE dataset*, while other graphical models may not offer or need complex algorithms to support such a function. Among graphical models, Bayesian networks offer the capability to combine multiple datasets, but they do not allow fine-grained control [93] (mixture ratio) over the combination of datasets. Unlike Bayesian networks, a

greedy algorithm (Figure 2.2) can be used for fine-grained control to train an ANN with an appropriate mixture of the two datasets.

To train the ANN for combining the *existing BPM dataset* and the *synthetic IVE dataset* based on a mixture ratio  $\alpha$  (0 to 1), the author uses an efficient greedy heuristic algorithm (Figure 2.2). The algorithm uses the mean absolute error (MAE) to measure the effectiveness of the ANN trained on both datasets. Before training the ANN, the *existing BPM dataset* is split into the *existing BPM training dataset* and the *existing BPM testing dataset*. Similarly, the *synthetic IVE dataset* is split into the *synthetic IVE training dataset* and the *synthetic IVE testing dataset*. During training, two MAEs are calculated in every epoch. The first MAE measures the difference in the predictions of the ANN and the *synthetic IVE testing dataset* ( $MAE^{SI}$ ). The second MAE measures the difference in the predictions of the ANN and the *existing BPM testing dataset* ( $MAE^{EX}$ ). The algorithm (Figure 2.2) uses  $\alpha$  to maintain the proportion of  $MAE^{SI}$  and  $MAE^{EX}$  based on the equations below:

$$\frac{MAE^{SI}}{MAE^{SI}+MAE^{EX}} \approx 1 - \alpha \quad (2-3)$$

$$\frac{MAE^{EX}}{MAE^{SI}+MAE^{EX}} \approx \alpha \quad (2-4)$$

The Equation (2-3) is simplified by substituting value of  $MAE^{SI} + MAE^{EX}$  in Equation (2-4) to become  $\frac{MAE^{SI}}{MAE^{EX}} \approx \frac{1-\alpha}{\alpha}$ , which is used during training. During training, at every epoch, if  $\frac{MAE^{SI}}{MAE^{EX}} > \frac{1-\alpha}{\alpha}$ , the algorithm greedily attempts to reduce  $\frac{MAE^{SI}}{MAE^{EX}}$  in the epoch by training the ANN on the *synthetic IVE training dataset* to reduce  $MAE^{SI}$ . Otherwise, in that epoch, the ANN is trained on the *existing BPM training dataset* to reduce  $MAE^{EX}$ . The training continues for a pre-specified number of epochs.

Several mixture ratios ( $\alpha$ ) may be used to combine the *existing BPM dataset* and the *synthetic IVE dataset* in the training. In the computational framework, the obtained results are called *updated BPMs*. The most accurate *updated BPM* when validated using reference data from the physical building is considered as the *augmented BPM*.

### ***Feature Ranking***

Feature ranking is generally used to discern and discard weakly influential, irrelevant, or redundant features from a given set of features before performing further critical analysis [94]. Techniques that are often used to perform feature ranking can essentially be divided into three main categories including filter, wrapper, and embedded methods. The filter method directly uses properties of data to estimate the goodness of features and ignores the effects of the selected feature subsets on the performance of a classifier. The wrapper method estimates the goodness of features by learning and evaluating the performance a classifier such as an ANN using only the features of interest [95]. The embedded method is a combination of the filter and the wrapper methods. The embedded method uses the internal information of a classifier to analyze feature ranking [96]. However, there is no best method among the three [97]. In the present study, the author applies the feature ranking technique to rank the influence of factors impacting human-building interactions. The wrapper method is selected since an ANN has been used as a classifier, and the input data are classified into features of interest (e.g., occupancy, intermediate leaving, and illuminance).

## **2.4 Application of the Computational Framework**

The application of the computational framework was focused on understanding the potential of the framework and validating the hypothesis. The prediction of artificial lighting usage in a single occupancy office was used for data collection and validation. The author

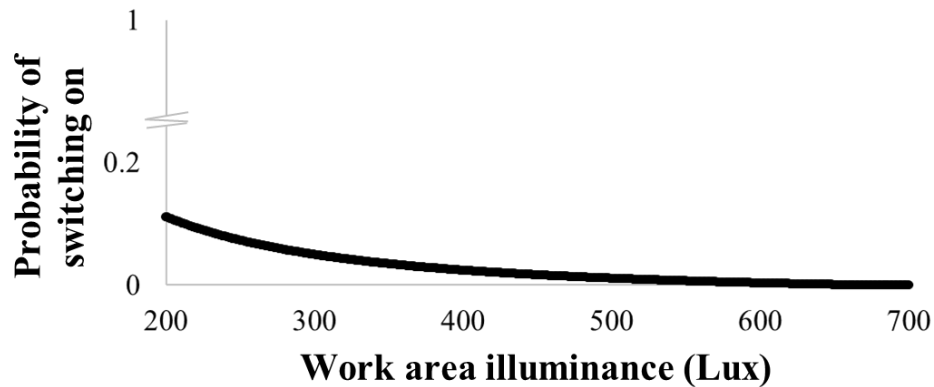


monitored a physical office for one month to collect artificial lighting usage data. The data obtained from the physical office were used for two purposes: (1) creating an IVE simulating different contextual conditions for collecting human-building interaction data and (2) validating the *augmented BPM*. An IVE was created by referring to the physical office as well as modeling conditions relevant to the variables of a selected BPM and contextual factors to be studied. The occupant who occupied the physical office also participated in the IVE experiment. The Hunt model for predicting lighting usage was selected as the *existing BPM* [1]. After computation, the most accurate *updated BPM* was selected as the *augmented BPM*. Predicted results of the *augmented BPM* were compared with predicted results of the *existing BPM* to evaluate the effectiveness of the proposed framework. In the following, the author explains the application in detail.

#### **2.4.1 Existing BPM and Existing BPM Dataset**

The light switch BPM proposed by Hunt [1] was selected as the *existing BPM*. The Hunt model was selected based on the following reasons: (1) it has been used as a baseline model for many extended models predicting artificial light use [5][98][99]; (2) it was cited as one of the major models by a recent paper in the field [100]; and (3) the framework is generic; that is, it can use the Hunt model or its expanded models as input. Moreover, the Hunt model has one independent variable (work area illuminance), which allowed the author to demonstrate the inclusion of other variables as contextual factors. Collecting data in IVE experiments for contextual factors is expensive because all virtual scenes and stimuli about the contextual factors need to be designed and modeled. Including more variables increases the expense and time consumed in an experiment. Therefore, to achieve the objective of the present study, any well-accepted model ideally with a small set of input variables was acceptable.

Hunt applied a field study approach to collect data on human-building interactions with light switches. He observed occupant light switching behaviors in six different rooms including multi-person offices, school classrooms, and open-plan teaching areas for six months. He deployed time-lapse photography to capture the lighting status in the rooms every 8 minutes throughout the day and night. Using probit regression analysis, the Hunt model predicted artificial lighting status based on work area illuminance (lux) (Figure 2.3) [1].



**Figure 2.3.** Probability of switching on under work area illuminance of the Hunt model.

A MC simulation was used to generate independent and identically distributed (IID) samples from the Hunt model. The input was work area illuminance, which was randomly generated following a uniform distribution. The output of the MC simulation was the probability of switching on under various work area illuminance levels. The input and output were arranged into pairs of work area illuminance and the corresponding probability of switch on. This data set was referred to as the *existing BPM dataset*.

## 2.4.2 Context-Aware Design-Specific Data

### *Physical Environment*

A single occupancy office located on the campus of a major state university in the south-central region of the United States of America was selected as the actual environment for the application (Figure 2.4). The office occupant was a male faculty member aged between 30 to 40

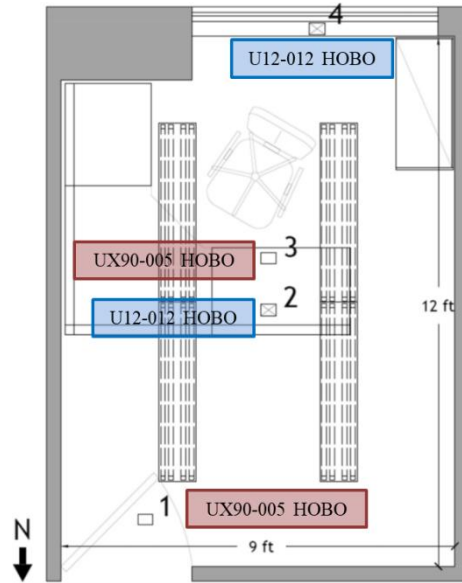
years. The dimensions of the office were 9 feet in width by 12 feet in length, and 10 feet in height (Figure 2.6). Various sensors were placed in the office to measure the lighting illuminance (lux), the artificial lighting status (on or off), and the occupancy pattern (occupied or non-occupied) from September 23<sup>rd</sup> to October 27<sup>th</sup>, 2016. One Onset UX90-005 HOBO occupancy/light runtime data loggers was placed above the door (sensor #1 in Figure 2.6) and one above the work area (desk) (sensor #2 in Figure 2.6) to identify the occupancy pattern and the lighting status (on or off), respectively. One Onset U12-012 HOBO temperature/relative humidity/light/ data loggers was placed at the work area (sensor # 3 in Figure 2.6) and one at the windows (sensor # 4 in Figure 2.6) to specifically measure the work area and outdoor light intensity, respectively. The sensors were set to collect data every 5 seconds.



**Figure 2.4.** The physical environment.



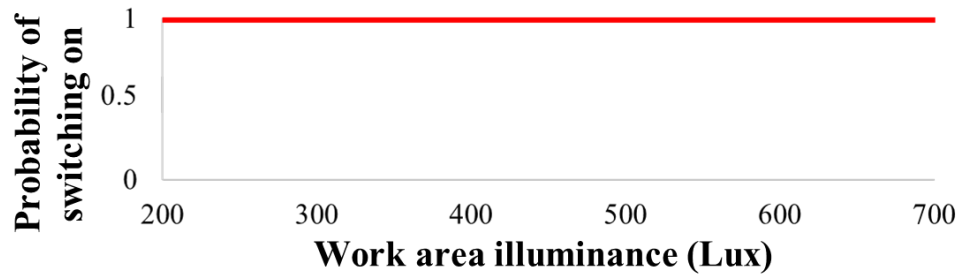
**Figure 2.5.** The IVE configuration.



**Figure 2.6.** The layout of the physical environment and the locations of the sensors.

The data collected from the physical environment were used to construct the IVE experiment. The author observed the major patterns of the occupant interactions (i.e., human-building interactions) with the office lighting system along with information of independent variable and contextual factors, namely work area illuminance, occupancy status, length of intermediate leaving, and outdoor illuminance. The major patterns of the occupant's interactions with the office lighting system were mapped into 141 events including (1) 25 event of arrival at the office, (2) 40 events of intermediate leaving, (3) 40 events of returning from the intermediate leaving, and (4) 36 events of departure. The IVE experiment was constructed based on the information of these events for data collection and validation.

The data also formed a baseline for evaluating the *augmented BPM*, i.e., the probabilities of switching on when the occupant arrived at the office. Figure 2.7 shows that the occupant always turned on the light regardless of the minimum work area illuminance based on the sensor data.



**Figure 2.7.** Probability of switching on under work area illuminance (physical environment).

### ***Immersive Virtual Environment (IVE) and Experiment***

The IVE configuration is illustrated in Figure 2.5. The IVE experiment was structured based on three main factors including (1) the considerations of the cost of developing IVE models and conducting experiments, (2) the obtained occupancy data from the physical environment, and (3) the Spatial-Temporal Event-Driven (STED) modeling approach [38]. The STED modeling approach designs an IVE experiment by modeling critical events during a day in chronological order, which comprises of four main components, namely states, contexts, events, and human-building interactions.

Based on the four main factors, states, contexts, events, and human-building interactions were defined as follows:

- States were light switch conditions, which included switched on and off.
- Contexts were conditions of the independent variable and the contextual factors in Table 2.1. The independent variable was work area illuminance considered in the Hunt model. The contextual factors were outdoor illuminance, occupancy, and intermediate leaving statuses. The occupancy statuses comprised of occupied and non-occupied. The intermediate leaving statuses were non-, short-, and long-leaving. The work area and outdoor illuminance were categorically defined. There were two

major constraints for illuminance to be designed as categorical. First, the STED modeling approach defined variables in IVE experiments as discrete. Although a small interval between minimum and maximum illuminance levels might have been defined to simulate continuous illuminance, it would have unnecessarily increased the number of IVE experiments. Second, due to the limitations of IVE technologies, an IVE experiment session typically only lasts for about 40 minutes. A participant might not have been able to tolerate the IVE for the long period of time necessary to simulate continuous illuminance found in the physical environment. Levels of work area illuminance were defined by applying the recommended lighting levels from the United States General Services Administration [101], which recommends 500 lux for the work area if performing office tasks. Accordingly, a darker level was defined as 200 lux, and a brighter level was defined as 700 lux based on the average of minimum and maximum natural light during office hours (8:00 a.m. to 5:00 p.m.). Although the levels of the work area illuminance were maintained as dark (200 lux), normal (500 lux), and bright (700 lux), the levels of the outdoor illuminance were determined with respect to the location of the Sun and direction of its light, which depended on the time of the day in the experiment. Therefore, if the outdoor illuminance was dark, normal, or bright, the work area illuminance without artificial lighting was assigned as dark, normal, or bright accordingly.

- Events were occurrences of contexts that triggered the occupant to change or maintain the states as shown in Table 2.3. There were five critical events considered during a day in the IVE experiments, including initial events before arrival at the office, on

arrival at the office, intermediate short leave or long leave, returning from the intermediate short leave or long leave, and departure.



- Human-building interactions were interactions of the occupant on light switching.

In each event, the situations of contextual factors and the independent variable included in the Hunt model (Table 2.1) were exposed to the participant. Visual (e.g., outdoor conditions) and auditory (e.g., reminders) cues were used to inform the participant about outdoor conditions and length of leaving or staying in the office, respectively. Examples of cues are shown in Table 2.2. The participant was asked to interact with the light switch, which he could switch on, switch off, or maintain. Then, data of occupancy status, work area and outdoor illuminance, and intermediate leaving status along with the light status in each event were collected throughout the sequence (Table 2.3). The participant was the same person who occupied the physical office. The participant used a HMD to experience the IVE and to participate in the experiment. The experiment was divided into two sessions, and each session lasted about 70 minutes in total. The study was approved by the local institutional review board.

**Table 2.1.** Contextual factors and the independent variable in the application.

Contextual factor	Status
Occupancy	Occupancy (True)
	Non-occupancy (False)
Outdoor illuminance	Dark
	Normal
	Bright
Intermediate leaving	Short intermediate leave
	Long intermediate leave
Independent variable	Status
Work area illuminance	Dark (200 lux)
	Normal (500 lux)
	Bright (700 lux)

**Table 2.2.** Examples of virtual and auditory cues used in the IVE experiment.

Event	Arrival at the office	Intermediate short leave	Intermediate long leave	Returning from the intermediate leave	Departure
Virtual cue	 (V1)	-	-	 (V2)	-
Auditory cue	You have arrived at your office. Please pick your most preferred your lighting choice for at least 2 hours of office work. (A1)	Your package is just delivered. Please go and quickly pick it up from the department office. (A2)	It is now 9:20 a.m., and you need to go and teach your class, which takes about 1 hour 30 minutes. You may or may not change the light status as you leave. (A3)	It is 11.15 a.m., and you have arrived at your office. Please pick your most preferred your lighting choice for at least 2 hours of office work. (A4)	It is 5:30 p.m. now, and you decide to go home. You may or may not change the lighting status of your office as you leave. (A5)



**Table 2.3.** The sequence of the IVE experiment.

Event	Sequence of IVE experiment in a day			
	<i>Light status before interaction</i>	<i>Virtual and auditory cues exposed to the participant</i>	<i>Interaction</i>	<i>Light status after interaction</i>
<b>Arrival at the office</b>	Initial light status	<i>V1</i> and <i>A1</i>	Participant interacts with light switch	Light status of the event
<b>Intermediate leave</b>	Light status of the previous event	Short ( <i>A2</i> ) Long ( <i>A3</i> )	Participant interacts with light switch	Light status of the event
<b>Returning from intermediate leave</b>	Light status of the previous event	<i>V2</i> and <i>A4</i>	Participant interacts with light switch	Light status of the event
<b>Departure</b>	Light status of the previous event	<i>A5</i>	Participant interacts with light switch	Light status of the event

Determinations of data points in the immersive virtual environment experiment

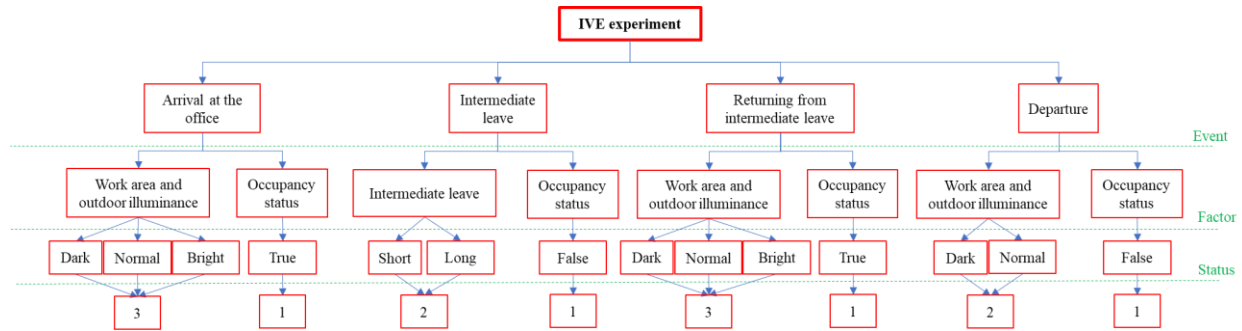
**Figure 2.8.** Diagram of factors included in the IVE experiment.

Figure 2.8 illustrates the diagram of factors included in the IVE experiment. Based on Figure 2.8, events of the arrival, intermediate leave, returning from intermediate leave, and departure included 3, 2, 3, and 2 alternatives, respectively, which led to  $3 \times 2 \times 3 \times 2 = 36$  unique combinations (called “sequences”). To construct the IVE experiment, 36 sequences along with the cost of developing the IVE and conducting experiment were taken into an account. First,

each sequence was extracted and assigned to four events (the 2<sup>nd</sup> to the 5<sup>th</sup> data points of each sequence in Table 2.4). Second, the initial events were appended to each sequence (the 1<sup>st</sup> data point of each sequence in Table 2.4). The initial events were not included in the sequence diagram in Figure 2.8 because it would have resulted in a two-fold increase in the number of total data points and excessive experimental time. Appending the initial events to the sequences relieved the number of total data points and excessive experimental time, and it maintained the uniqueness of the 36 sequences. Therefore, the total data points were 180 [i.e., 36 (sequences) x 4 (data points in each sequence) + 36 (data points from each initial event)].

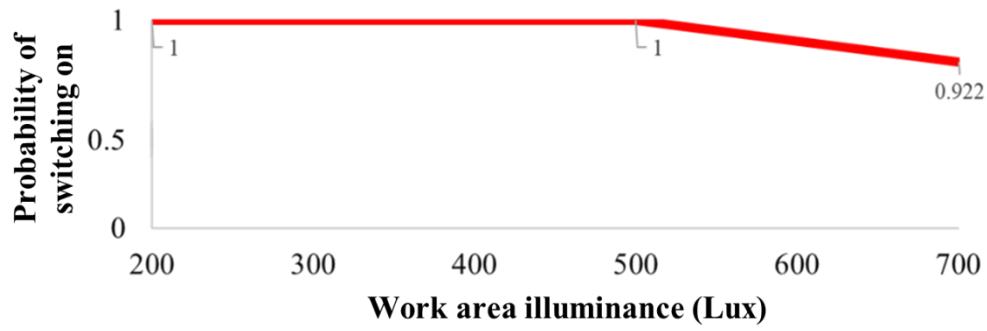
**Table 2.4.** Details of IVE experimental days.

Sequence	Initial (1 <sup>st</sup> data point of each sequence)		Arrival at the office (2 <sup>nd</sup> data point of each sequence)		Intermediate leave (3 <sup>rd</sup> data point of each sequence)		Returning from intermediate leave (4 <sup>th</sup> data point of each sequence)		Departure (5 <sup>th</sup> data point of each sequence)	
	Light switch	Occupancy	Illuminance	Occupancy	Intermediate leave	Occupancy	Illuminance	Occupancy	Illuminance	Occupancy
1	On	False	Bright	True	Long	False	Bright	True	Normal	False
2	Off	False	Bright	True	Long	False	Bright	True	Dark	False
3	On	False	Bright	True	Long	False	Normal	True	Normal	False
4	Off	False	Bright	True	Long	False	Normal	True	Dark	False
5	On	False	Bright	True	Long	False	Dark	True	Normal	False
6	Off	False	Bright	True	Long	False	Dark	True	Dark	False
7	On	False	Bright	True	Short	False	Bright	True	Normal	False
8	Off	False	Bright	True	Short	False	Bright	True	Dark	False
9	On	False	Bright	True	Short	False	Normal	True	Normal	False
10	Off	False	Bright	True	Short	False	Normal	True	Dark	False
11	On	False	Bright	True	Short	False	Dark	True	Normal	False
12	Off	False	Bright	True	Short	False	Dark	True	Dark	False
13	On	False	Normal	True	Long	False	Bright	True	Normal	False
14	Off	False	Normal	True	Long	False	Bright	True	Dark	False
15	On	False	Normal	True	Long	False	Normal	True	Normal	False
16	Off	False	Normal	True	Long	False	Normal	True	Dark	False
17	On	False	Normal	True	Long	False	Dark	True	Normal	False
18	Off	False	Normal	True	Long	False	Dark	True	Dark	False
19	On	False	Normal	True	Short	False	Bright	True	Normal	False
20	Off	False	Normal	True	Short	False	Bright	True	Dark	False
21	On	False	Normal	True	Short	False	Normal	True	Normal	False
22	Off	False	Normal	True	Short	False	Normal	True	Dark	False
23	On	False	Normal	True	Short	False	Dark	True	Normal	False
24	Off	False	Normal	True	Short	False	Dark	True	Dark	False
25	On	False	Dark	True	Long	False	Bright	True	Normal	False
26	Off	False	Dark	True	Long	False	Bright	True	Dark	False
27	On	False	Dark	True	Long	False	Normal	True	Normal	False
28	Off	False	Dark	True	Long	False	Normal	True	Dark	False
29	On	False	Dark	True	Long	False	Dark	True	Normal	False
30	Off	False	Dark	True	Long	False	Dark	True	Dark	False
31	On	False	Dark	True	Short	False	Bright	True	Normal	False
32	Off	False	Dark	True	Short	False	Bright	True	Dark	False
33	On	False	Dark	True	Short	False	Normal	True	Normal	False
34	Off	False	Dark	True	Short	False	Normal	True	Dark	False
35	On	False	Dark	True	Short	False	Dark	True	Normal	False
36	Off	False	Dark	True	Short	False	Dark	True	Dark	False

### *Generating Synthetic IVE Data*

Because the data sample from the IVE experiment was small, and the IVE data represented a sequence of events, the author decided to employ an HMM Baum-Welch algorithm [102] to generate a synthetic dataset based on data from the IVE experiment (i.e., increasing the number of independent and IID samples). The advantages of the HMM are that it has the ability to statistically learn information about observed parameters to estimate for non-observable parameters [103] and recognizes sequential patterns of provided data [104].

The HMM Baum-Welch algorithm [102] learnt the relationship of the participant's light switch interactions and the factors influencing the interactions. In general, the HMM assumes that the current state ( $S_t$ ) impacts the next state ( $S_{t+1}$ ). The hidden state happening at time  $t+1$  ( $S_{t+1}$ ) is dependent only on the hidden state happening at time  $t$  ( $S_t$ ) [105] [106]. The change in hidden states from time  $t$  to time  $t+1$  is called state transition. The probability of state transitions can be calculated and simplified as a transition probability matrix. The observations depend on the hidden state variables, and the probability density function of observations is therefore dependent on the hidden state variables [106]. The observation probabilities can be expressed in a matrix form as an observation probability matrix. The HMM is trained by using the distribution of hidden states and observations from the transition and observation probability matrices. After training, the HMM is executed to generate IID samples.

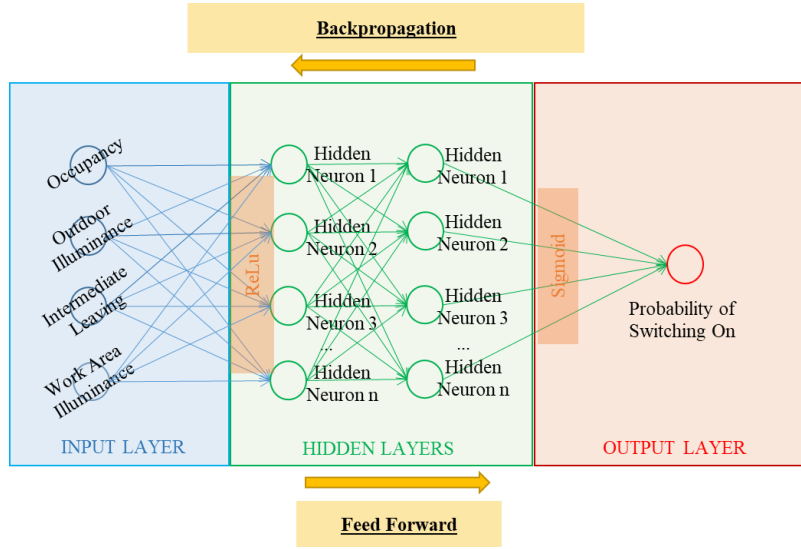


**Figure 2.9.** Probability of switching on under work area illuminance (HMM).

From the 180 data points obtained from the IVE experiment, the hidden states and the observations of events were classified. The statuses of the light switch were classified as the hidden states. The statuses of the other variables, namely occupancy status, intermediate leaving, as well as outdoor and work area illuminance were classified as observations. Each observation as a vector was encoded to an ordinal variable for training the HMM. For instance, non-occupancy, short intermediate leave, dark work area illuminance, and normal outdoor illuminance were represented as “no + short + dark + dark” and encoded by using a single value such as “1”. The transition and observation probabilities were calculated. The HMM was trained to learn the relationship between the hidden states and observations from the transition and observation probabilities. After training, the HMM was executed to generate the statuses of the light switch and variables in Table 2.1. A complete analysis of HMM for the case study can be found in [107]. To obtain the variables corresponding to the Hunt model, the probabilities of switching on upon arrival based on work area illuminance were computed (Figure 2.9). The probabilities of switching on upon arrival were calculated and paired with the IID samples of variables in Table 2.1 generated by the HMM, and this was called the *synthetic IVE dataset*.

### 2.4.3 Computation

#### *Artificial Neural Network (ANN)*



**Figure 2.10.** Scheme of ANNs of the computational framework.

The ANN (Figure 2.10) was a three-layered perceptron network including input, two hidden, and output layers. Input in the input layer included occupancy status, intermediate leaving, and minimum work area illuminance. Output in the output layer was the probability of switching on. The hidden layers of the model were generated using 300 hidden neurons with rectified linear unit activation function (ReLU) since it has been shown to have better fitting ability than the sigmoid function in similar applications [94]. To prevent overfitting, elastic net regularization [combination of L1 (Laplacian) and L2 (Gaussian) penalties] [108] was used. The sigmoid activation function was applied to the neuron at the output layer because the values of outputs were probabilities. The loss function of the model was binary cross entropy (logistic regression). The learning rate and regularization were  $10^{-6}$ .

Before input data could be used by the ANN, they were first normalized to ensure compatibility between the *existing BPM dataset* and the *synthetic IVE dataset*. Since the Hunt

model has only illuminance as an independent variable, contextual data for the Hunt model, e.g., occupancy, intermediate leaving, and outdoor illuminance were randomly generated according to the available occupancy, intermediate leaving, and outdoor illuminance of the *synthetic IVE dataset*. For instance, since occupancy status in the *synthetic IVE dataset* included non-occupied and occupied statuses, the data of occupancy in the *existing BPM dataset* were randomly generated with non-occupied and occupied statuses. Corresponding to the statuses of intermediate leaving in the *synthetic IVE dataset*, the data of intermediate leaving in the *existing BPM dataset* were randomly generated with non-leave, short intermediate leave, and long intermediate leave.

After normalization, inputs and outputs of the *existing BPM dataset* and the *synthetic IVE dataset* were defined as shown in Figure 2.10. The *existing BPM dataset* and the *synthetic IVE dataset* were divided into training datasets (i.e., the *existing BPM training dataset* and the *synthetic IVE training dataset*) and testing datasets (i.e., the *existing BPM testing dataset* and the *synthetic IVE testing dataset*) based on 80-20 splits. Five percent of the inputs of the *synthetic IVE training dataset* were changed to white Gaussian noise to prevent overfitting during training.

### ***Training Algorithm***

To initialize the ANN model, the *existing BPM training dataset* was used to train the ANN for 60,000 epochs to allow the ANN to accurately learn the probability distribution of the *existing BPM dataset*. After initializing, to train the ANN on a mixture of the *existing BPM training dataset* and the *synthetic IVE training dataset* with a mixture ratio ( $\alpha$ ), the efficient greedy heuristic algorithm (Figure 2.2) was used. The training continued for 400,000 epochs. To understand the impact of the mixture ratio on the prediction accuracy of *updated BPMs*, mixture ratios ( $\alpha$ ) from 0 to 1 with an interval of 0.1 were used to generate a sequence of *updated BPMs*.

After training, the *existing BPM testing datasets* were used as inputs to acquire outputs (the probabilities of switching on) through the trained ANN. The inputs (i.e., the inputs of the *existing BPM testing datasets*) and the obtained outputs were paired to construct *updated BPMs*. Among these *updated BPMs*, the best performing one defined as the one with the highest prediction accuracy relative to the actual data was selected as the *augmented BPM*.

### ***Feature Ranking***

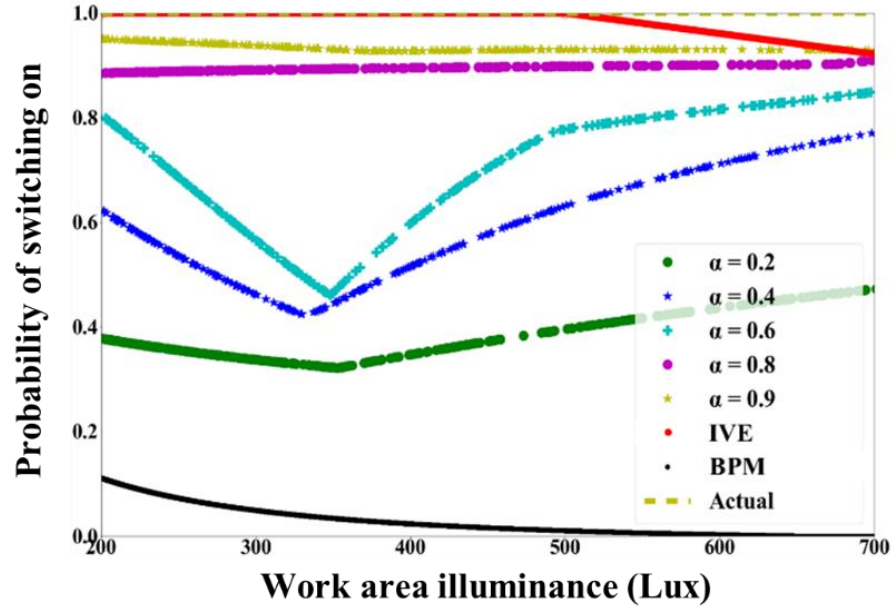
Factors, such as occupancy status, intermediate leaving, and work area illuminance, have different magnitudes of impact on predictions [38]. Thus, it is important to have a method to determine the relative importance of such factors. The feature ranking technique was used to evaluate the influence of factors on predictions. To perform feature ranking in the ANN, the ANN was trained using the *synthetic IVE training dataset*, which was modified by considering only one specific factor of interest, at the input layer. The probability of switching on was selected as the output in the output layer. For example, the ranking of occupancy status was analyzed by training an ANN using the *synthetic IVE training dataset* that was modified to have the occupancy status as the only input factor and the probability of switching on as the output in the ANN. The ANN was trained using the similar scheme as mentioned in Figure 2.10.

The coefficient of determination ( $R^2$ ) was used as a statistical measurement of the linear relationship between expected outputs (i.e., probability of switching on of the *synthetic IVE testing dataset*) and predicted output (i.e., the probability of switching on obtained from prediction by the ANN).  $R^2$  provides a measure of how accurate expected outputs are learned by the ANN [109]. The value of  $R^2$  ranges from 0 to 1, in which 1 means the probability of switching on can be predicted without error. Therefore, a higher  $R^2$  means a factor has a more influential impact on the prediction of switching on.



## 2.5 Results

### 2.5.1 Updated BPMs



**Figure 2.11.** Observations of updated BPMs obtained from the computational framework using various mixture ratio ( $\alpha$ )

\*\*\*IVE = the synthetic IVE dataset  
BPM = the existing BPM dataset  
Actual = the actual data from the physical environment

Figure 2.11 presents the *updated BPMs* (each with a different  $\alpha$ ) and plots the probability of switching on versus work area illuminance ranging from 200 lux to 700 lux. In addition, the *existing BPM dataset*, the *synthetic IVE dataset*, and the actual dataset obtained from the physical office are also presented. Some observations can be made from Figure 2.11:

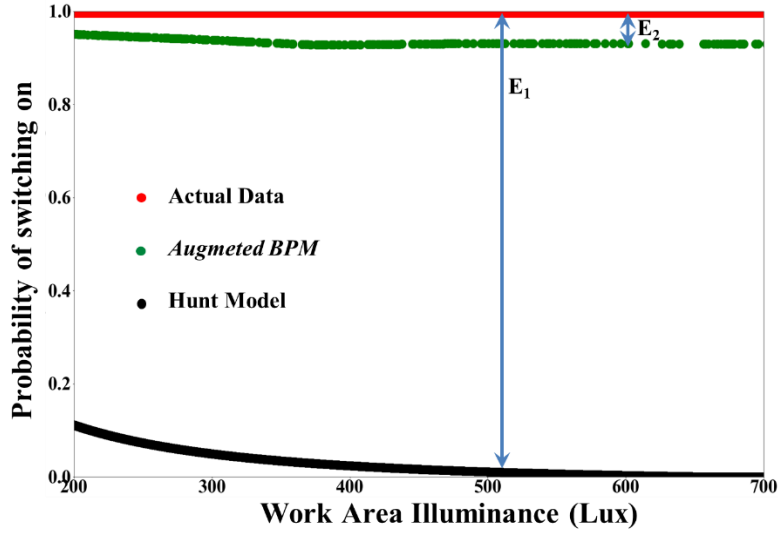
- The prediction accuracy of *updated BPMs* improved as  $\alpha$  increased.
- Significant improvements in terms of prediction accuracy of *the updated BPMs* occurred if  $\alpha$  was between 0.2 and 0.8. However, if  $\alpha$  was more than 0.8, the rate of the improvement was not as obvious as if  $\alpha$  was smaller than 0.8.
- If  $\alpha$  was 0.2, 0.4, or 0.6, the probability of switching on decreased when the work illuminance was lower than around 350 lux and then increased when the work

illuminance was higher than 350 lux. These behaviors occurred because of several reasons. One of the main reasons is that the *synthetic IVE dataset* was categorical, which included work area illuminance at 200, 500, and 700 lux. At  $\alpha = 0.2, 0.4$ , or  $0.6$ , the weight of the *existing BPM dataset* was stronger than that of the *synthetic IVE dataset*, especially in the region around 350 lux. The *existing BPM* biased the probability of switching on toward itself. At  $\alpha = 0.8$  and above, the weight of the *existing BPM dataset* became weaker than the *synthetic IVE dataset*. The *updated BPMs* tended to follow behaviors of the *synthetic IVE dataset*. However,  $\alpha$ 's between  $0.6$  and  $0.8$  were not observed in the study.

The observations demonstrate the potential of the proposed framework to generate *updated BPMs* that are better than the *existing BPM*. The observations also show that  $\alpha$ , a measure for mixing the two datasets, may have a relationship with the prediction accuracy of *updated BPMs*. Finding an optimal  $\alpha$  can help an application to reach a desired level of prediction accuracy.

### 2.5.2 Hypothesis Testing

The *updated BPM* with a mixture ratio  $\alpha$  of  $0.9$  was considered to be the *augmented BPM* since it had the best predictive ability among all the generated *updated BPMs*.



**Figure 2.12.** The hypothesis testing.

To validate the hypothesis, 500 data samples were randomly drawn from the *existing BPM testing dataset*, in which the occupancy status was “true”, and the intermediate leaving status was set to “non-leave” to be consistent with the Hunt model. The *augmented BPM* and the Hunt model were both tested on this dataset, and their predicted outputs were recorded. Five hundred samples were then drawn from the actual dataset under identical conditions (i.e., occupancy, intermediate leaving, and work area illuminance).

To test the hypothesis, Equations (2-1) and (2-2) were used to determine  $E_1$  and  $E_2$ , respectively, as shown below:

$E_1 =$  | The probability of switching on from the prediction of the *existing BPM dataset* –

The probability of switching on from the actual data |

$E_2 =$  | The probability of switching on from the prediction of the *augmented Hunt model* –

The probability of switching on from the actual data |

A one-tailed *t-test* was used to identify the statistically significant difference between the mean of two errors, i.e.,  $E_1$  and  $E_2$  (Figure 2.12). The hypothesis was:

$$H_0: \text{mean of } E_1 - \text{mean of } E_2 = 0$$

$$H_1: \text{mean of } E_1 - \text{mean of } E_2 > 0$$

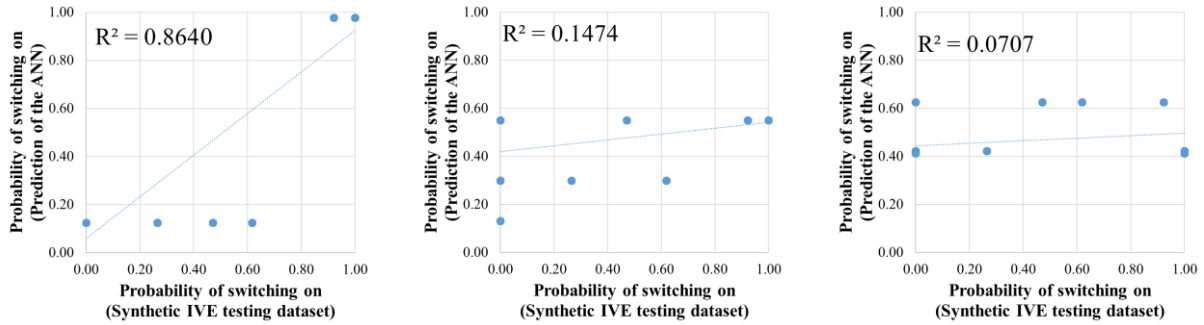
**Table 2.5.** The summary of *t-test* ( $\alpha = 0.05$ ) analysis

Absolute t-value	Degrees of freedom	P-value	H <sub>0</sub>
617.94	998	< 0.05	Reject

From Table 2.5, the result of the one-tailed *t-test* show that the p-value was smaller than 0.05; therefore, the null hypothesis was rejected. The result can be interpreted as the mean of  $E_1$  was significantly higher than the mean of  $E_2$ . It can be concluded that the probabilities of switching on estimated by the *augmented BPM* were significantly closer to the actual data than those estimated by the Hunt model. This result implies that combining data reflecting design-specific contextual factors with data from the Hunt model could generate an *augmented BPM* with higher prediction accuracy than the Hunt model.

### 2.5.3 Feature Ranking Analysis

Figure 2.13 shows the plot of the probability of switching on obtained from the *synthetic IVE testing dataset*, the prediction of the ANN, as well as the coefficient of determination ( $R^2$ ) for occupancy status (a), leaving status (b), and work area illuminance (c). It was assumed that occupancy status, leaving status, and work area illuminance were independent of each other. The results of  $R^2$  in Figure 2.13 show that the most influential factor was occupancy status ( $R^2 = 0.8640$ ), followed by leaving status and work area illuminance. This result is consistent with other studies (e.g., [110][111]), which suggests that the feature ranking analysis has the potential to identify influential factors.



a. Occupancy Status

b. Leaving Status

c. Work Area Illuminance

**Figure 2.13.** Plots of probability of switching on obtained from the synthetic IVE testing dataset and prediction of the ANN.

## 2.6 Limitations of the Study

Even though the potential of the framework is demonstrated through its application to the case, the major limitations of the study were as follows:

- At this stage, the framework did not have the capability to determine an optimal mixture of data from an *existing BPM* and context-aware design-specific data. Therefore, a series of mixtures were applied to show the impact of mixing data from two different sources. However, it is ideal to have an approach allowing designers or researchers to quickly determine an optimal mixture depending on the goal of building performance.
- The study used a single occupancy office as a case study. A participant's habitual behavior is unique, affecting the observational data. The case demonstrates the potential of the computational framework because it showed the deviation of human-building interactions from predictions and demonstrated the potential of the computational framework to bias a general model to fit a specific design. However, more cases, such as different types of occupants and multi-occupancy offices, need to be studied.

- The limitations of virtual reality technologies determine that it is difficult to continuously collect human-building interaction data in virtual environment for extended periods of time. Hence, data collection using IVEs are not longitudinal. That is, only a limited set of illuminance data is collected in an IVE experiment.

## 2.7 Conclusion and Future Work

In this chapter, a computational framework has been discussed. The framework combined design-specific contextual factors with an *existing BPM* to produce an *augmented BPM* with better prediction accuracy. An IVE was used to capture data related to design specific contextual factors. The framework was applied to a lighting use study, in which the Hunt model was chosen as the existing BPM. An ANN combined data from the Hunt model with the data obtained from the IVE experiment (i.e., context-aware design-specific data) to generate the *augmented BPM*. Results show that the *augmented BPM* produced better predictions than the Hunt model. Although the Hunt model was selected in this study, the framework was not designed specifically for the Hunt model.

Several conclusions can be made based on the application of the framework to the prediction of light switch status of a single occupancy office:

- Design-specific contextual factors play an important role in predicting human-building interactions. Other studies [9] [112] [113] have similar conclusions, which support the results of the present study.
- The framework demonstrates the potential of integrating design-specific contextual factors with an *existing BPM* to generate an *augmented BPM*, which produced better predictions than the *existing BPM*. However, it should be noted that this study did not

offer an approach to determine the  $\alpha$  of the *augmented BPM*. Future work is needed to determine such an approach.

- The framework relies on an IVE to collect data related to design-specific contextual factors. As pointed out by previous studies, using an IVE as a data collection tool has limitations [6]. Although visual stimulation is the most mature IVE capability and is mainly applied in this study, it alone cannot simulate all kinds of human-building interactions. Other capabilities, such as simulating the acoustics and thermal comfort of an indoor space, should be developed and incorporated in the future.
- Feature ranking has the potential to identify important factors influencing predictions. The proposed method effectively identified that occupancy status strongly affected the predictions of light switch status as reported by previous studies (e.g., [110] [111]). The ability to identify the most influential factors can help designs of IVE experiments to have better data collection.

The contributions of the study are as follows:

- The main contribution of the study is the computation framework that biased an *existing BPM* to better fit the context of a building under design. The case study demonstrates the potential of the framework to improve performance predictions. This approach is different from conventional approaches in which BPMs often developed using data of existing buildings are applied to buildings under design. Due to the uniqueness of each building and the context-dependent nature of the behaviors of occupants, *existing BPMs* developed using conventional approaches often fail to produce accurate predictions. Thus, the computational framework offers new

possibilities to assist designers or researchers to improve performance predictions during building design.

- An additional contribution of the framework is to assist designers or researchers in integrating contextual factors related to a new design with an *existing BPM*. To adopt the framework for a building under design, designers or researchers need to select an *existing BPM*, identify contextual factors that are relevant to the design, and then collect context-aware design-specific data addressing specific human-building interactions in the context of the design using IVEs. There is no restriction on the *existing BPM* or the contextual factors that can be considered. In most cases, it depends on the knowledge or experience of designers or researchers to make decisions. For a user, the computational framework is treated as a black box after the *existing BPM* and the contextual factors are determined. That is, a user only uses an augmented model produced by the framework to generate predictions, which better addresses the context of a building under design.
- The framework is intended for use during a design stage, especially if a designer has several design options and needs to determine the performance of a building under design.

In the future, the framework needs to be validated in different indoor environments. The data in this study were collected from a single occupancy office. Thus, other types of spaces, including homes and multi-occupancy offices along with other types of occupant needs and preferences should also be studied. Moreover, the framework needs to be improved to allow designers or researchers to use performance targets (e.g., building benchmarks, building standards, arbitrary building data, and energy consumptions) as the guide for combining data



from an *existing BPM* and context-aware design-specific data [114]. This step is important because it makes the framework practical. It will help designers or researchers to obtain appropriate mixtures without trying many mixture ratios. The framework will help designers or researchers to compare different design alternatives using performance targets as a guide. From the comparison, designers or researchers will be able to determine which design alternative should be selected in order to obtain an overall optimal design. In addition, uncertainties due to the limitations of IVE technologies need to be considered in the future improvement of the computational framework.

## **CHAPTER 3. AUGMENTING BUILDING PERFORMANCE PREDICTIONS DURING DESIGN USING GENERATIVE ADVERSARIAL NETWORK AND IMMERSIVE VIRTUAL ENVIRONMENTS**

### **3.1 Introduction**

The design stage of a building project is a critical step to make decisions and establish directions for engineering building components that affect the characteristics, functions, and performance of a building. To optimally translate design goals and objectives into the performance of a building, designers and engineers usually apply building performance models (BPMs) during the design stage, such as simulations of building energy consumptions and human-building interactions, to understand, investigate, and predict building performance as well as support decision-making. Nevertheless, the application of BPMs cannot eliminate significant performance discrepancies between the simulated and the actual performances that have been widely reported [9][115][116]. For example, studies have reported as much as a 150% difference between predictions and the actual performance of a building [14].

Many factors influence the simulations of building performance, especially human-building interactions, such as occupant responses to building contexts and occupant habitual behaviors [11]. Human-building interactions are highly context-dependent and sensitive to several contexts [18][117] that are described by situational factors that are not directly included in a model or simulation [118]. These situational factors are often assumed to remain constant across different applications of the model or simulation. For instance, “context” may be physical or natural factors (e.g., building characteristics, building surrounding and remain constant

---

This chapter was previously published as C. Chokwitthaya, Y. Zhu, S. Mukhopadhyay, and E. Collier, “Augmenting building performance predictions during design using generative adversarial networks and immersive virtual environments,” *Automation in Construction*, vol. 119, p. 103350, 2020, doi: 10.1016/j.autcon.2020.103350. Reprinted by permission of Elsevier.

across different applications of the model or simulation. For instance, “context” (environmental factors, and climate conditions), and socio-technical factors (e.g., a participant’s cultural background, race/ethnicity, and tasks to be performed), which may not be included as variables in a BPM. However, such factors can have an impact on analysis using the BPM during the design of a specific space that may have different situational factors from the assumptions of the BPM. That is, situational factors often cannot be treated as constant across different applications. In such cases, these situational factors in relation to any BPM need to be identified, analyzed, and integrated in building performance analyses. BPMs are often developed using data obtained from existing buildings, in which the contexts differ from the contexts of a building under design. Applying such BPMs to understand, investigate, and predict human-building interactions in a building under design may contribute to the discrepancy between predicted and actual performance. Therefore, being able to address human-building interactions responding to specific contexts in new designs (e.g., the context embodied) can potentially enhance the accuracy of BPMs, leading to reductions of the discrepancy between predicted and actual performance of a building.

Immersive virtual environments (IVEs) have demonstrated their potential in simulations and data collections in many disciplinary areas, especially engineering fields, such as emergency evacuations [76][77], building designs [80], and human-building interactions [35][36][37]. IVEs provide several advantages over other data collection methods, such as sensing, field studies, and surveys. For instance, IVEs can replicate certain contexts for buildings under design, especially if replication of the contexts in reality is not possible, cost-effective, or safe. Additionally, IVEs allow users to manipulate all experimental conditions and customize experimental models as desired. Human-building interactions in buildings under design may not be directly observed and

analyzed. As a result, the application of IVEs can be an alternative for generating and examining the context-aware design-specific data of a new design. Following Sowa's definition of context [118], "context-aware" refers to the capability of a method, simulation, or model to address the impact of identified contextual factors in analysis. Therefore, by using a method, simulation, or model, users are able to consider human-building interactions responding to contexts of a specific design. For example, in the application discussed in the present study, the context-aware design specific data of the proposed computation framework included contextual factors such as types of office task and locations of light switch.

To improve the accuracy of existing building performance models (*existing BPMs*), the author offers a framework for customizing *existing BPMs* to address contextual factors of a building under design. A framework using an artificial neural network (ANN)-based greedy algorithm was developed to combine an *existing BPM* with context-aware design-specific data obtained from IVE experiments [59]. The framework shows the potential to enhance the prediction accuracy of an *existing BPM*. However, its major limitation is lack of capability to determine the appropriate combination of an *existing BPM* and context-aware design-specific data by a principled approach rather than through trial and error, which can cause excessive resource use and is time consuming. Hence, the principal goal of this study is to improve the capability of the framework to be able to determine the appropriate combination without trial and error. The new computational framework applies generative adversarial networks (GANs) to combine an *existing BPM* with context-aware design-specific data obtained from IVE experiments and uses a *performance target* as a guide during computation to determine the appropriate mix without trial and error. The GAN-based framework produces an augmented

building performance model (an *augmented BPM*) representing the appropriate combination that satisfies the *performance target*.

In sections 3.2 to 3.9, the author compares the GAN-based framework and the ANN-based greedy algorithm framework, states the research objective, states an expression of the GAN-based framework, and describes the application of the framework to a single-occupancy office to validate the framework. The design and administration of the IVE experiment are explained in detail. Finally, results, discussion, limitations of the study, conclusions, and directions of future work are provided.

### 3.2 Comparison of the GAN-Based Framework and the ANN-Based Greedy Algorithm

This section discusses major differences between the GAN-based framework and the ANN-based greedy algorithm framework to clarify the relationship between both frameworks. In parametric approaches (e.g., Gaussian mixture model), mixture models mix datasets are derived from assumed probability distribution functions, such as normal, binomial, and exponential [119]. Datasets often do not fully comply with assigned distributions, leading to the generation of inaccurate mixture models. Consequently, an ANN-based greedy algorithm framework was proposed in the previous chapter [59]. The framework is non-parametric so that users do not have to assume that the distributions of the underlying datasets are mixed. The framework enhances an *existing BPM* by combining its dataset with context-aware design-specific data from IVE experiments. However, this framework still has several limitations. First, it only allows users to apply a linear combination using an assigned mixture ratio (number between 0 and 1) to mix data from the two datasets. Thus, if the probability distributions corresponding to the two datasets are termed  $f_1$  and  $f_2$ , the greedy algorithm would produce a mixture distribution,  $(1-\alpha)f_1 + \alpha f_2$ , in which  $0 \leq \alpha \leq 1$ . This is called a linear mixture with  $\alpha$  as the mixture ratio. The major

limitation of this approach is its inability to create a mixture distribution that is close to a *performance target* by non-linearly mixing the probability distributions of the two datasets. In addition, the mixture ratio is not directly related to any *performance target*. Even in the case of linear mixtures, the framework does not provide any algorithm to determine the appropriate mixture ratio that generates the best mix. Due to this limitation, the *augmented BPM* can only be constructed through trial and error. Users of the framework have to manually define the mixture ratios to combine datasets and need to perform several trials to obtain an appropriate combination for deriving the *augmented BPM*. In practice, it may sometimes not be possible to obtain appropriate combinations.

To overcome the disadvantages of the proposed ANN-based greedy algorithm framework, a GAN-based framework is proposed and compared with the ANN-based greedy algorithm framework. The GAN-based framework uses a GAN [50] to combine data of an *existing BPM* and context-aware design-specific data. Like the ANN-based greedy algorithm, the GAN allows a nonparametric approach to generate mixture models in which users do not have to assume any distribution (e.g., normal) for the underlying datasets getting mixed or for the mixture. In contrast to the ANN-based greedy algorithm, the GAN-based framework allows automatic determination of the appropriate mixture guided by a building performance target. This avoids the trial and error techniques required in the ANN-based greedy algorithm framework and allows users to obtain an appropriate combination of datasets in a single attempt.

Filtering based approaches, such as the Kalman filter [120], require manual determination of the filter type (e.g., linear, extended, and unscented) that result in an appropriate mixture. In contrast, the GAN-based framework allows automatic determination of the appropriate mixture guided by a building performance target.

### 3.3 Research Objective

The aim of this chapter is to create a new GAN-based framework that enables users to better perform building performance simulations during design. To achieve the goal, there are two objectives of this chapter: 1) to investigate efficacy of the GAN-based framework in enhancing the prediction accuracy of BPMs, and 2) to examine the reliability of the GAN-based framework using experiments.

To determine the reliability of the framework, the author conducted experiments on 30 college students to acquire data and statistically tested 30 comparisons between an *augmented BPM* and an *updated BPM*. The GAN-based framework generated an *augmented BPM* and the previous framework using the ANN-based greedy algorithm generated an *updated BPM* [59]. The comparison was based on the hypothesis that the *augmented BPM* without trial and error would be more accurate than the *updated BPM*. The absolute error measured discrepancy between the *updated BPM* and the *performance target* ( $E_1$ ) and the absolute error measured discrepancy between the *augmented BPM* and the *performance target* ( $E_2$ ) are shown in Table 3.1.  $E_1$  and  $E_2$  were calculated using Equation (3-1) and (3-2), respectively. They were used to test the hypothesis.

**Table 3.1.** The definition of the errors to prove the hypothesis.

Error	Measurement
$E_1$	The predicted outcome of an <i>updated BPM</i> – a <i>performance target</i>   (3-1)
$E_2$	The predicted outcome of an <i>augmented BPM</i> – a <i>performance target</i>   (3-2)

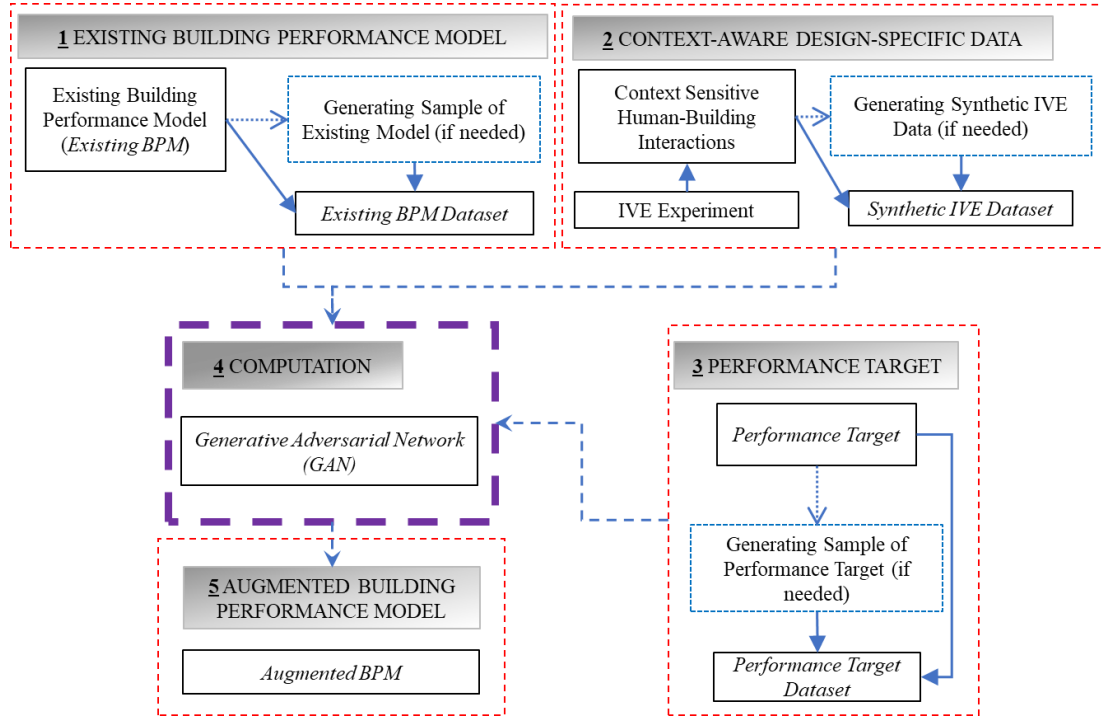
To test the performance of the *augmented BPM*, the hypothesis was defined as follows:

$$H_0: \text{mean of } E_1 - \text{mean of } E_2 = 0$$

$H_1$ : the null hypothesis is not true.

A *t-test* ( $\alpha = 0.05$ ) was applied to determine whether the performance of the *augmented BPM* was significantly different from that of the *updated BPM*.

### 3.4 Overview of the Computational Framework



**Figure 3.1.** The GAN-based framework.

The five main components of the GAN-based framework (Figure 3.1) were: 1) an existing building performance model (an *existing BPM*), 2) context-aware design-specific data obtained from an IVE experiment, 3) a *performance target*, 4) GANs, and 5) an *augmented BPM*.

In general, the term “BPMs” is used to describe models of building performance at different building scales. BPMs may include performance models from a small scale, such as specific building systems (e.g., lighting, blind, and window usage) to a large scale, such as whole buildings (e.g., whole building energy consumption). For example, at the building system level, Tahmasebi and Mahdavi [121] proposed a BPM for predicting window operations, and Keller, et al. [122] developed a BPM to estimate performance of building systems (e.g., gas, electricity,



and water). At the whole building level, Cho, et al. [123] developed a BPM to estimate the whole building energy performance. Indeed, the framework is parametric, i.e., taking three generic inputs, namely an *existing BPM*, context-specific data from IVE, and a *performance target*. These inputs are not related to a particular type of performance simulation. Therefore, the framework can be applied to different types of BPM at different scales, and it is not dependent on the nature of BPMs. In the following, details of the framework are discussed.

### 3.4.1 The Existing Building Performance Model

*Existing BPMs* describe historical events and observations and may not fully consider important contextual factors corresponding to a new building. That is, contextual factors influencing human-building interactions in a new building are ignored in the *existing BPM* [59]. In addition, the *existing BPM* for predicting human-building interactions may be in different forms, such as statistical models or synthetic datasets (generated by the models). When synthetic samples of an *existing BPM* are required, the GAN-based framework offers an approach to produce samples using a statistical approach (e.g., Monte Carlo simulation). The dataset associated with the samples is called an *existing BPM dataset* [59].

### 3.4.2 Context-Aware Design-Specific Data

Context-aware design-specific data describe key contextual conditions of a new design in which human-building interactions occur. For instance, the Hunt model [1] only models the relationship between the use of artificial lighting and work area illuminance. However, the types of task (e.g., reading, meeting, and drafting) and the locations of light switch (e.g., a switch by a door or on a desk) influence the preferences of occupants interacting with the light switch. In the present study, the types of task and the locations of light switch are contextual factors. When a design needs to explicitly consider such contextual factors to augment an *existing BPM*, IVEs

can be used as tools to obtain context-aware design-specific data for a new design. Nevertheless, conducting IVE experiments is sometimes time consuming, and each experiment session is often limited to 30-40 minutes. Therefore, IVE experiments usually result in small data samples with specific experiment conditions [38][84][85]. To overcome such limitations, the GAN-based framework uses a data synthesis technique, such as the Gaussian mixture model (GMM), to generate a large independent and identically distributed (IID) dataset, called a *synthetic IVE dataset* [59].

### 3.4.3 Performance Target

During design, designers often consider and balance multiple factors (e.g., code compliance, comfort, cost, energy, function, operation, occupancy characteristic, and sustainability) to satisfy design objectives [124]. Based on such objectives, various performance metrics do exist, such as the energy intensity of a building. However, for a new computational framework to work, building performance metrics need to be converted into operational *performance targets* to support computation. This process is still an open question, which requires further research attention.

In this chapter, the author assumes an operational *performance target*. Such a *performance target* may be created using empirical performance data of similar buildings and represented in the form of a statistical model or a set of data. However, since a *performance target* is used to evaluate with data generated by an *existing BPM*, the components in both the *performance target* and the *existing BPM* need to be comparable.

### 3.4.4 Computation

#### *Generative Adversarial Network (GAN)*

Since Goodfellow, et al. [50] proposed the GANs in 2014, the method has been successfully applied in various domains, especially deep learning-based studies [49][51][125][52] as well as image syntheses and analyses [54][126][127].

GANs have two parts: a generator and a discriminator. The generator is an ANN that attempts to learn a probability distribution and tries to generate an output that follows a target distribution. The discriminator is an ANN discriminating the output of the generator and the target distribution. Conceptually, the generator and the discriminator play a two-player minimax game during which they undermine each other. The undermining continues until an equilibrium point is reached at which the generator and the discriminator do not change their performance regardless of what the opposition may do. Theoretically, in each epoch, the generator tries to produce the output that follows the target distribution. The discriminator observes the output of the generator as well as the target distribution. It tries to accurately discriminate whether the output of the generator is from the target distribution. The feedback from the discriminator is used to train the generator through backpropagation. In every epoch, the generator keeps trying to produce an output that follows the target distribution while the weights of the discriminator are adjusted through backpropagation to accurately discriminate the generator outputs. The process continues until it reaches an equilibrium point at which the generator produces an output with a distribution close to the target distribution, and the discriminator accurately discriminates the generator's outputs and the target distribution [128].

In the GAN-based framework, the generator is trained using combinations of data associated with the *existing BPM* and context-aware design-specific data. The discriminator is

trained using the generator's predictions and the *performance target dataset*. The generator has the responsibility to produce an output that is close to the *performance target dataset*. The discriminator has responsibilities to discriminate the output of the generator and the *performance target dataset*. The prediction of the generator that is the closest to the *performance target* is considered as the *augmented BPM*.

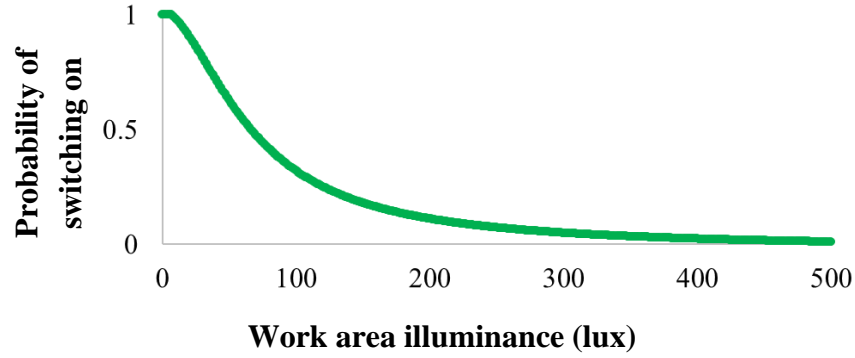
In GANs, the performance of the discriminator relatively relies on the complexity and dimension of the input. If the input is complex and has only one dimension, the discriminator may be inferior in discriminating the generator's predictions and the *performance target dataset* [114]. To avoid such circumstances and enhance the performance of the framework, the concept of conditional GANs [129] is applied to condition on the generator as well as the discriminator; that is, the inputs of the generator are fed into the input layer of the discriminator.

### 3.5 Application of the GAN-Based Framework

The application aimed to understand the efficacy and the reliability of the GAN-based framework by testing the hypothesis. The application used the lighting predictions in a single occupancy office as the studied case. An IVE configuration was created based on general recommendations of office designs [130], simulating situations related to variables of a BPM, and contextual factors [i.e., work area illuminance (lux), office tasks including reading, having a break, having a meeting, and drafting along with the location of a light switch such as by a door and on a desk]. Thirty people participated in the IVE experiment. The *existing BPM* was the Hunt model for predicting lighting usage [1]. The probability of switching on, a probit model provided in Da Silva et al. [98], was used as a *performance target*. After computation, *augmented BPMs* were compared with *updated BPMs* obtained using the ANN-based greedy

algorithm framework [59] to evaluate the efficacy and the reliability of the GAN-based framework.

### 3.5.1 Existing Building Performance Model



**Figure 3.2.** The Hunt model.

The selection of the light usage prediction model proposed by Hunt [1] was used for several reasons including 1) its citation and use as a baseline lighting BPM in lighting use studies [98][5][99][2], 2) its application in several building performance software packages (e.g., ESP-r, DAYSIM, and RADIANCE), and 3) it having only one independent variable (work area illuminance), allowing the author to study and demonstrate the potential to include other factors as contextual factors. One of the successful applications of Hunt model is that it was used as the underlying theory to develop an algorithm (Lightswitch-2002) for simulating dynamic daylight used in DAYSIM and RADIANCE software. Currently, both software packages have been widely used in not only the academic field but also in industrial ones. The selection of the Hunt model was only for demonstrating and testing the framework. In fact, the GAN-based framework is generic, which means it can take any BPM as an *existing BPM*.

Hunt collected data of human-building interactions with light switches using the time-lapse photography method in six rooms (e.g., multi-person offices, school classrooms, and open-plan teaching areas). The obtained data were fitted using a probit model as Equation (3-3) to

predict the probability of switching on using work area illuminance (lux) as an independent variable. Figure 3.2 shows the relationships of the probability of switching on and the work area illuminance of the Hunt model. The author applied an MC technique to generate samples of the Hunt model as an *existing BPM dataset*. MC simulation is a random process to mimic a behavior of real-life systems [131]. It is widely accepted as a standard technique to generate IID samples for models and has been used in several works [132][133][134]. The proper number of samples was determined by using a learning curve technique [135][136]. In general, the learning curve technique investigates the impact of the number of samples used to train the ANN on the accuracy of predictions of the trained ANN. The technique continuously increases the number of samples until additional samples do not significantly increase the accuracy of the trained ANN (i.e., the knee point is reached). The number of samples at the knee point is taken as the number of samples. The application excluded the analysis of the learning curve since it was demonstrated in Chokwitthaya, et al. [136], and the number of samples in the *existing BPM dataset* was carried over from the cited work.

$$\text{Probability of switching on} = -0.0175 + \frac{1.0361}{(1 + e^{4.0835(\log_{10}(\text{work area illuminance}) - 1.8223))}} \quad (3-3)$$

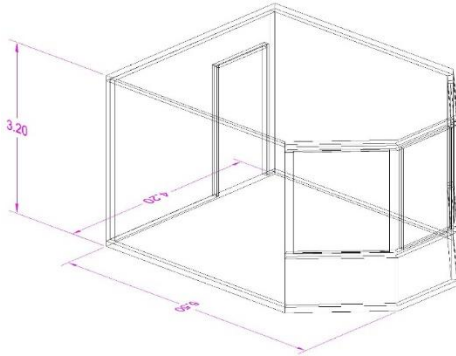
### 3.5.2 Context-Aware Design-Specific Data

#### *Immersive Virtual Environment (IVE) and Experiment*

##### The IVE Model

A single office was modeled in virtual reality based on the general recommendation of office designs [130]. The dimension of the office was 5.5 x 4.2 x 3.2 meters with a net area of 22 square meters as shown in Figure 3.3. AutoCAD software was used to create a 3D model of the office. Autodesk 3ds Max was used to assign and render materials of the model. It was also used to estimate the work area and the indoor illuminance in the application since Autodesk 3ds Max

showed the potential to simulate indoor illuminances for daylighting [137]. Unreal Engine 4 was applied to simulate the virtual 3D environment (Figure 3.4), allowing participants to interact with building components such as the light switches.



**Figure 3.3.** The model of the office.



**Figure 3.4.** The IVE configuration.

#### The Design of the IVE for the Experiments

The concept of the spatial-temporal event-driven (STED) modeling approach [38] was partially applied to the design of the IVE experiment. The STED modeling approach models critical events during a day in a chronological order representing longitudinal observations in reality to minimize the negative impact of IVE technologies on participants. The STED modeling approach uses categorical values to model such critical events in IVEs. For example, events during a day may be arrival, intermediate leaving, and departure. The STED model comprised of four main components (e.g., states, contexts, events, and human-building interactions). In the present application, the design of the IVE experiment ignored the chronological order so that events were discretely modeled and did not influence next adjacent events; that is, a finished state of an event was not transferred to be an initial state of a next adjacent event as described in the STED model.

According to the STED model, the four main factors (i.e., states, contexts, events, and human-building interactions) were defined as follows for the IVE experiment:

- States were the on/off status of the light switch, which was initially set to off in all scenarios.
- Contexts were conditional factors describing the state. There could be many contextual factors. In the present study, two commonly identified factors, office tasks and light switch locations, were selected based on previous literature discussing the impact of tasks [138][139][140] and, in particular, the location of light switches [35]. The level of work area illuminance was defined using the following recommendations: (1) the recommended lighting level from the United States General Services Administration [101], which suggests 500 lux as the light intensity at work areas for conducting office tasks and (2) previous studies that have shown a significantly low probability of switching on when the work area illuminance is higher than 200 lux [11][5][141]. Therefore, the IVE experiment was designed to use 500 lux as the maximum work area illuminance. The work area illuminance between 200 and 500 lux was assigned with 150-lux intervals. The smaller interval of illuminance (i.e., 50 lux) was assigned between 0 and 200 lux to capture more possible fluctuations of human-building interactions on light switching in this range. Table 3.2 describes details of the office tasks, light switch locations, and work area illuminance considered in the IVE experiment. However, it should be noted that although it would have been possible to define smaller intervals for the work area illuminance to simulate continuous illuminance, it would have resulted in a costly increase the duration of the IVE experiment. Thus, the choice of the interval was based on the consideration to finish an experiment for each participant within 60 minutes including a training session.



- Events were triggers that caused occupants to change or maintain the state. In this application, combinations of the contextual factors (4 office tasks and 2 light switch locations) and the dependent variable (6 work area illuminance) led to  $4 \times 2 \times 6 = 48$  events. Accordingly, each participant generated 48 data points with one data point for each event.
- Human-building interactions (dependent variables) were the likelihood of switching on or off (the dependent variable), and these are described in Table 3.2.

**Table 3.2.** Variables and their values considered in the IVE experiment.

<b>Contextual factor</b>		<b>Independent variable</b>	<b>Dependent variable</b>
<i>Office task</i>	<i>Light switch location</i>	<i>Work area illuminance (lux)</i>	<i>Likelihood of switching on</i>
Intensive reading	By the door	50	Very unlikely
Having a break	On the desk	100	Not likely
Having a meeting		150	Neutral
Accurate drawing		200	Likely
		350	Very likely
		500	
<i>Total = 4</i>	<i>Total = 2</i>	<i>Total = 6</i>	<i>Total = 5</i>

### The IVE Experiment and Data Collection

The experiment had two main sessions, namely the training session and the experiment session. In the following, the details of the IVE experiment are described.

The training session was designed to familiarize participants with the IVE experiment. This session allowed participants to explore in the virtual environment themselves along with practices of responding to questions in the experiment. The training session took around 10 minutes for each participant.

In the experiment session, participants initially sat on a chair that was at the desk in the virtual environment as if they were about to perform some tasks in the office. The participants

were told that they were the sole occupant of the office who could interact with a light switch freely. In each event, the participants were exposed to one of the work area illuminances presented in Table 3.2 and the audio cues presented in Table 3.3 that informed the participants about the current conditions of the office. After listening to audio cues, the participants were asked to determine the likelihood of switching on under the condition of the office at that moment. There were five available choices of the likelihood constructed based on a Likert scale [142], and the choices were mapped to the probability of switching on as shown in Table 3.4. The following example illustrates how the audio cues and questions were presented to the participants. If the task and the switch location were “*intensive reading*” and “*by the door*” in a particular event under any work area illuminance, the audio cue was “*You are going to intensively read research papers for at least an hour. If the light switch is by the door, and you have to walk to turn it on, please select your need of turning the light on under the provided situation.*” The 48 events mentioned previously were assigned to the participants randomly, and each experiment took around 40 minutes.

**Table 3.3.** Audio cues of office tasks and switch locations.

Office task		Switch location	
<i>Task</i>	<i>Audio cue</i>	<i>Location</i>	<i>Audio cue</i>
Intensive reading	You are going to intensively read research papers for at least an hour.	By the door	The light switch is by the door, and you have to walk to turn it on.
Having a break	You are going to have a break in your office for at least an hour.	On the desk	The light switch is on the desk, which is now reachable.
Having a meeting	You are going to have a meeting in your office for at least an hour.		
Drafting	You are going to work on a drafting task for at least an hour.		

**Table 3.4.** Choices of the likelihood of switching on and their interpretation.

Choice (likelihood)	<i>Very unlikely</i>	<i>Not likely</i>	<i>Neutral</i>	<i>Likely</i>	<i>Very likely</i>
Probability of switching on (%)	1	25	50	75	99

A total of 30 students (18 males and 12 females) participated in the research study.

Before conducting the experiment, the participants signed consent forms and completed a motion sickness screening questionnaire. The experiment was approved by the local institutional review board. The participants wore a HMD to conduct the experiment in the IVE during the experiment. Figure 3.5 shows a participant exploring the office and making a decision on the use of the lighting during the experiment.



**Figure 3.5.** A participant exploring the virtual office and selecting a need of turning the light on.

### 3.5.3 Generating Synthetic IVE Data

Since the sample from the IVE experiment was relatively small, the authors augmented the dataset by creating a synthetic dataset. The Gaussian mixture model (GMM) [143] was proven to have higher performance in clustering data than many other data clustering methods, such as k-means [144], k-nearest neighbor [145], and multivariate kernel density (MVKD) [146]. Moreover, it has been used to generate IID samples in research studies [147][148]. In the present application, a GMM was used to augment the IVE dataset by increasing the number of IID samples based on the data obtained from the IVE experiment, i.e., “the *synthetic IVE dataset*”. The GMM learnt the IVE experimental data by using mixtures of Gaussian distribution, in which the data were categorized into  $z$  groups based on the Gaussian (i.e., normal) distribution. Each group had its own mean ( $\mu$ ) and variance ( $\sigma^2$ ). It was assumed that each data point ( $x$ ) probabilistically belonged to  $z$ , and the distribution for each group was separately inferred.

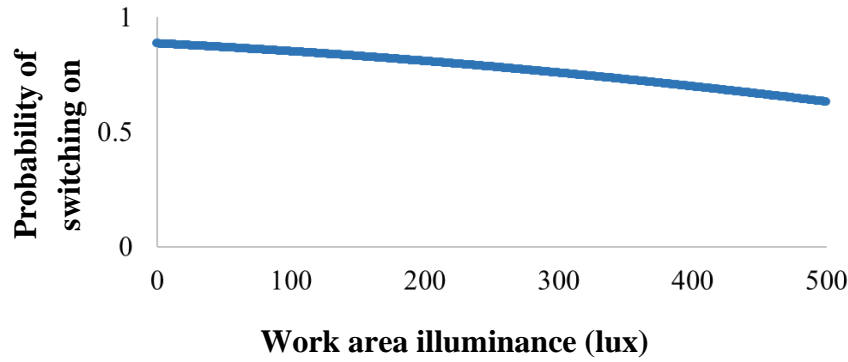
The GMM was constructed according to what was done by Chokwitthaya, et al. [149]. The context-aware design-specific data obtained from the IVE experiments were sourced from each individual participant, resulting in 30 datasets. To train the GMM, the  $K$ -mean algorithm was used to initialize the GMM parameters [150]. The full covariance type was applied. The convergence criterion for training the GMM was  $10^{-2}$ . After training, the GMM was executed to generate IID samples as the context-aware design-specific data, called the *synthetic IVE dataset*. Since 30 datasets were used to train the GMM individually, there was a total of 30 sets of the *synthetic IVE dataset*. Similar to the *existing BPM dataset*, the learning curve technique determined the number of samples in the *synthetic IVE dataset* [136].

### 3.5.4 Performance Target

In the present application, the probit regression model of light switch used proposed by Da Silva, et al. [98] described in Equation (3-4) was selected as a *performance target*. The model was selected for the following reasons: 1) data used to construct the model were obtained from eight single-occupancy offices that were similar to the design of the single occupancy office in the present application; 2) probit regression models have been accepted and applied to represent data associated with human-building interactions in many research studies [2][151][152]; 3) the model used work area illuminance as an independent variable similar to the *existing BPM*; and 4) the large discrepancy between the Hunt and Da Silva models set a high target, which increased the challenge for the framework to generate an *augmented BPM* that could meet the performance target. The challenge might prevent an *augmented BPM* from meeting the *performance target* in some cases, which provided the opportunity for the author to explore and discuss such cases. Indeed, the framework can take any performance model as a *performance target*. A target is used to guide the combination of an *existing BPM* and context-aware design-specific data in the framework. The framework attempts to generate an *augmented BPM* that is close to the performance target. Therefore, if a different *performance target* is used, the framework generates a different *augmented BPM* that is correlated with the characteristics of the new *performance target*.

$$\text{Probability of switching on} = \frac{1}{(1 + e^{-(2.035 + (-0.003) * \text{work area illuminance})})} \quad (3-4)$$

The authors applied MC simulation to generate IID samples of the Da Silva model, presented in Figure 3.6. The obtained IID dataset was called a *performance target dataset*.



**Figure 3.6.** The Da Silva model.

### 3.5.5 Computation

#### *Data Preprocessing*

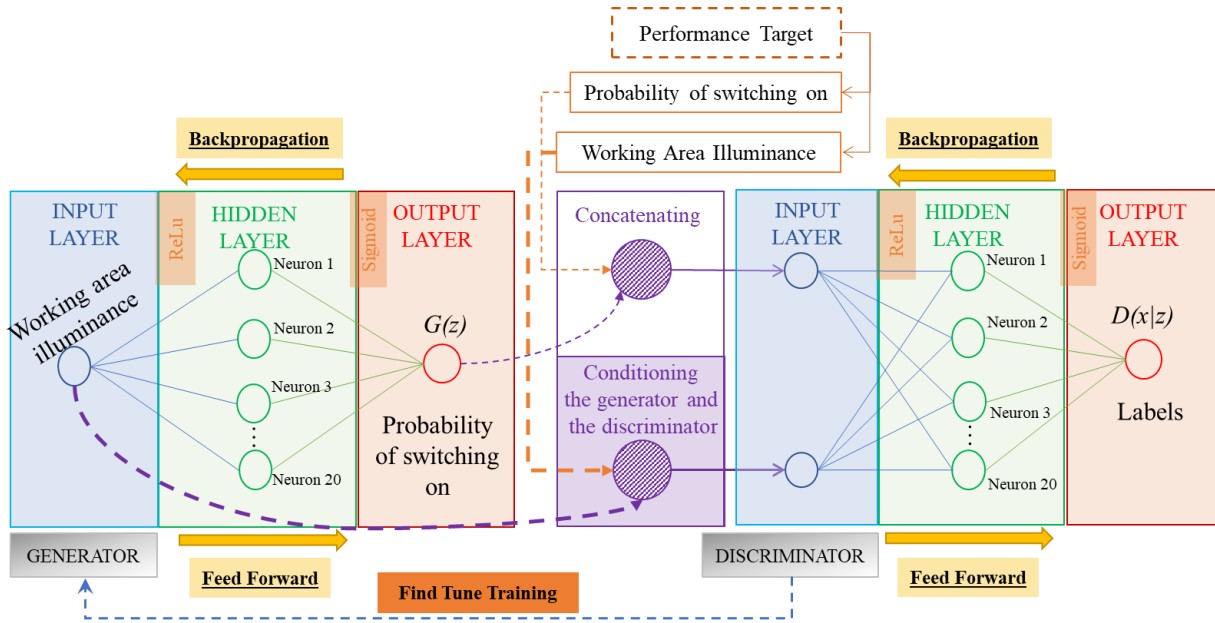
The *performance target dataset*, the *existing BPM dataset*, and the *synthetic IVE datasets* were normalized to maintain the compatibility and consistency of datasets during computation. The normalization was done by using Equation (3-5) [153]. To reduce the complexity of the *synthetic IVE datasets* during computations, each of the *synthetic IVE dataset* was split into eight groups of *sub-synthetic IVE datasets* based on the contextual factors (Table 3.5). The computation used each of the *sub-synthetic IVE datasets* to augment the *existing BPM*.

$$\text{Normalized data} = \frac{(\text{data} - \text{means of the synthetic IVE dataset})}{\text{standard deviation of the synthetic IVE dataset}} \quad (3-5)$$

**Table 3.5.** Groups of the *sub-synthetic IVE dataset*.

Group	Combination of the contextual factors
1	Intensive reading + Light switch by the door
2	Having a break + Light switch by the door
3	Having a meeting + Light switch by the door
4	Drafting + Light switch by the door
5	Intensive reading + Light switch on the desk
6	Having a break + Light switch on the desk
7	Having a meeting + Light switch on the desk
8	Drafting + Light switch on the desk

## Generative Adversarial Networks (GANs)



**Figure 3.7.** Scheme of the generative adversarial network (GAN) of this application.

A GAN was applied as the computational component for the framework. The GAN comprised of two major components, which were a generator and a discriminator (Figure 3.7). Prior to training the GAN, the *existing BPM dataset* and the *sub-synthetic IVE datasets* were split into training datasets (i.e., the *existing BPM training dataset* and the *sub-synthetic IVE training datasets*) and testing datasets (i.e., the *existing BPM testing dataset* and the *sub-synthetic IVE testing datasets*) with 70-30 splits.

The generator was an ANN with a three-layer perceptron, namely an input, a hidden, and an output layer. In each training epoch, the *existing BPM training dataset* and the *sub-synthetic IVE training dataset* were used to train the generator. The generator took data of the work area illuminance as the input in the input layer. The hidden layer included 20 hidden neurons with the rectified linear unit activation function (ReLU). Elastic net regularization, a combination of L1 (Laplacian) and L2 (Gaussian) penalties, was applied to prevent overfitting [108]. The input in

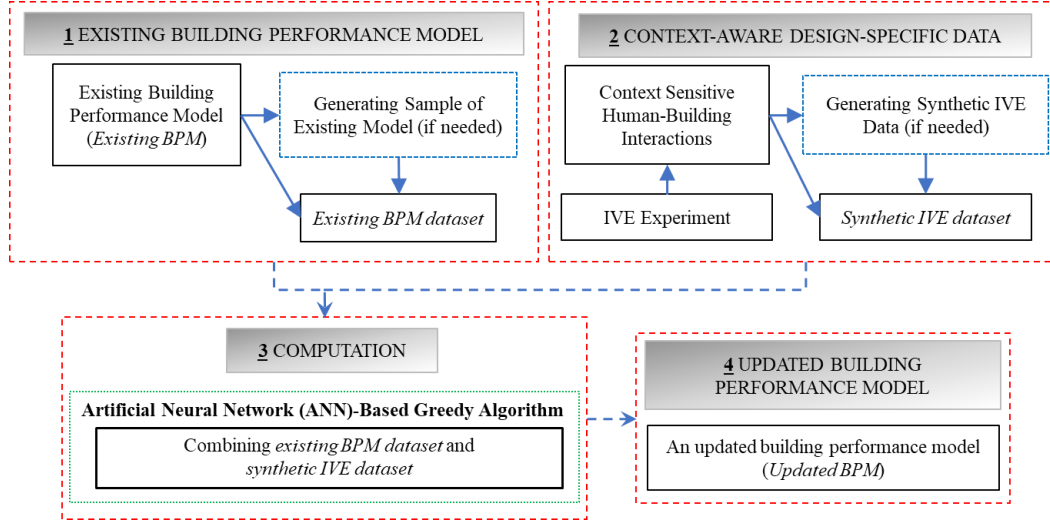
the output layer of the generator was the probability of switching on. The output layer applied the sigmoid activation function since the output was probability. The loss function, learning rate, and regularization were binary cross entropy (logistic regression),  $10^{-4}$ , and  $10^{-4}$  respectively. Before training the GAN, the generator was pre-trained using the *existing BPM training dataset* and the *sub-synthetic IVE training dataset* to initialize its weights and biases as well as to increase the robustness of its learning. The generator learned the relationship of work area illuminance and probability of switching on associated with the *existing BPM training dataset* and the *sub-synthetic IVE training dataset*. In every training epoch, the generator predicted the probability of switching on based on the work area illuminance of the *existing BPM testing dataset* and the *sub-synthetic IVE testing dataset*. The probability of switching on predicted by the generator that was closest to that in the *performance target* was used to construct an *augmented BPM*.

The discriminator was an ANN discriminating the outputs from the generator and the *performance target dataset*. The discriminator comprised of a three-layered ANN with similar structure as the generator but different input datasets. Before training the discriminator, the work area illuminance as the input in the generator were paired with the probability of switching on predicted by the generator. The paired data were labeled as 0. The *performance target dataset* was labels as 1. The labels were assigned to distinguish the dataset from the generator and the *performance target dataset*. Indeed, labels could be any numbers besides 0 and 1. The paired data were then concatenated with the *performance target dataset* (the purple box in Figure 3.7) and used as the input in the input layer of the discriminator. This step generated conditioning situation on to the generator and the discriminator. The labels were taken as the input in the output layer.



The *synthetic IVE dataset* of each participant was applied to train the GAN resulting in the *augmented BPM* of a participant. Therefore, a total of 30 *augmented BPMs* were obtained.

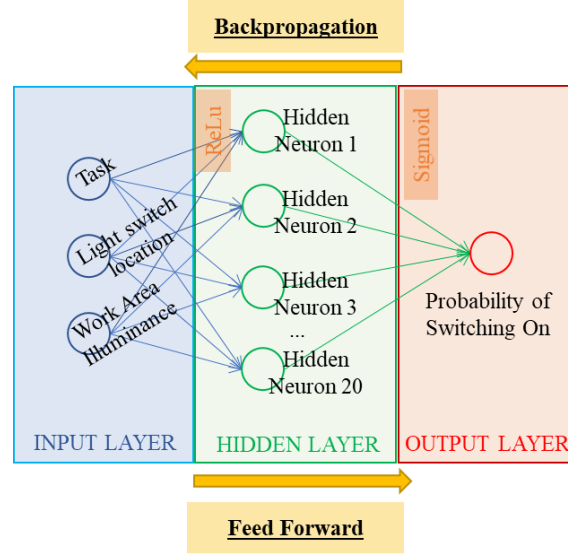
### 3.5.6 Overview of the ANN-Based Greedy Algorithm



**Figure 3.8.** Scheme of the ANN-based greedy algorithm framework.

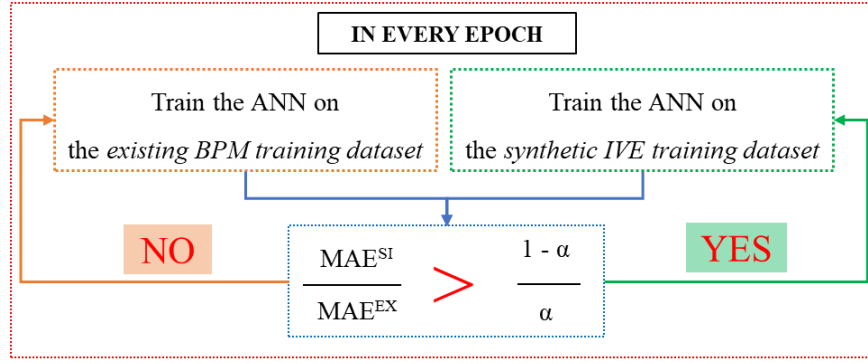
The ANN-based greedy algorithm framework [59][136] comprised of four major elements as shown in Figure 3.8: (1) *an existing BPM*, (2) context-aware design-specific data obtained from an IVE experiment, (3) computation, and (4) *an updated BPM*.

The *existing BPM*, the context-aware design-specific data, the *synthetic IVE datasets* and the data preprocessing were identical to those in the application of the new framework.



**Figure 3.9.** The scheme of the ANN of this application.

The computation comprised of the ANN combining the *existing BPM dataset* and the *synthetic IVE dataset* by using a mixture ratio ( $\alpha$ ) to guide the combination. The structure of the ANN is illustrated in Figure 3.9. In this application, the ANN was a three-layered perceptron with an input, a hidden, and an output layer. The ANN took the tasks, light switch locations, and work area illuminance as inputs in the input layer. The input in the output layer was the probability of switching on. The hidden layer had 20 neurons with ReLU and the elastic net regularization. The sigmoid activation function was applied to the output layer. The loss function, learning rate, and regularization were the binary cross entropy (logistic regression),  $10^{-4}$ , and  $10^{-4}$ , respectively. The *existing BPM dataset* and the *synthetic IVE datasets* were divided into training datasets (i.e., the *existing BPM training dataset* and the *synthetic IVE training datasets*) and testing datasets (i.e., the *existing BPM testing dataset* and the *synthetic IVE testing datasets*) with 70-30 splits.



**Figure 3.10.** Efficient greedy algorithm.

The training algorithm used an efficient greedy heuristic algorithm (Figure 3.10) proposed in Chokwitthaya, et al. [59] to determine whether the ANN was trained on the *existing BPM training dataset* or the *synthetic IVE training dataset*. In each training epoch, two values of the mean absolute error (MAE) were calculated: (1) the MAE measuring the difference between the predictions of the ANN and the *synthetic IVE testing dataset* ( $MAE^{SI}$ ) and (2) the MAE measuring the difference between the predictions of the ANN and the *existing BPM testing dataset* ( $MAE^{EX}$ ). During training the ANN, a mixture ratio ( $\alpha$ ) was used to maintain the proportion of mixture between the *synthetic IVE training dataset* and the *existing BPM training dataset* using Equation (3-6). If Equation (3-6) was true, the ANN was trained on the *synthetic IVE training dataset*. Otherwise, in that epoch, the ANN was trained on the *existing BPM training dataset*.

$$\frac{MAE^{SI}}{MAE^{EX}} > \frac{1 - \alpha}{\alpha} \quad (3-6)$$

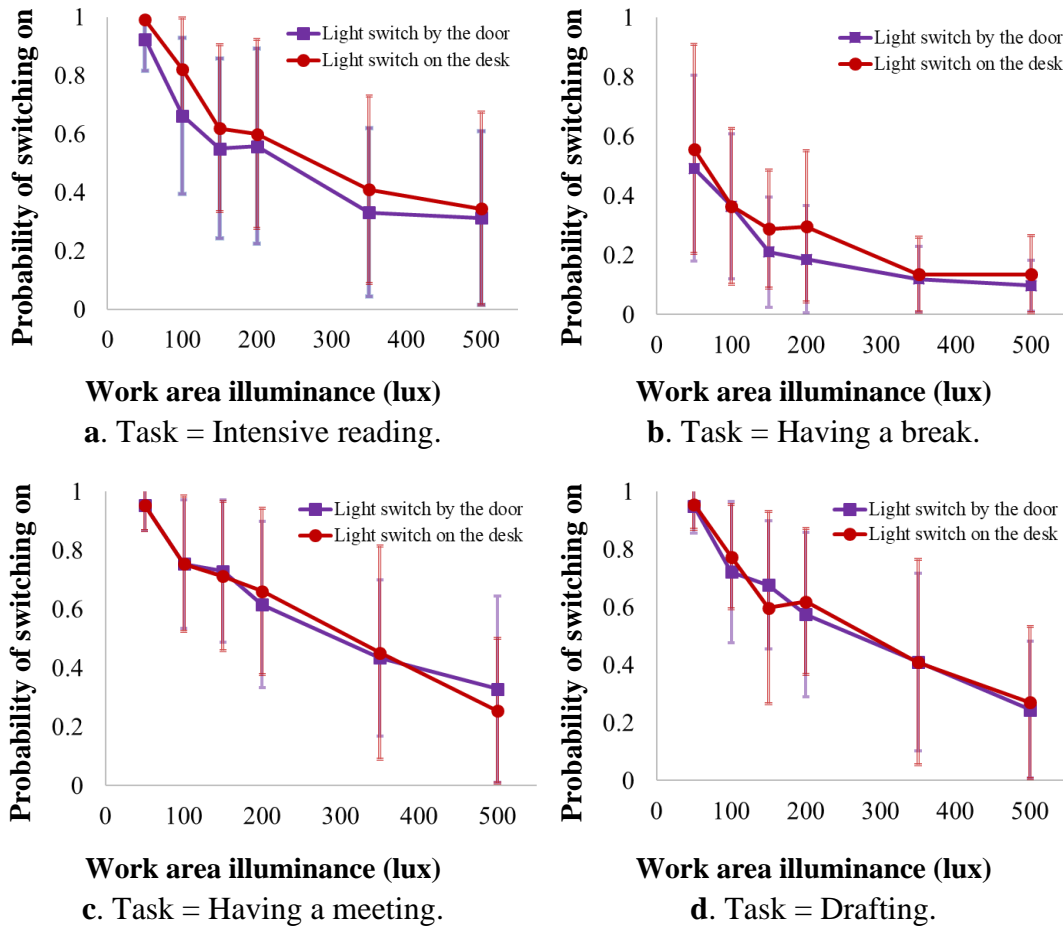
The computation was performed on each of the *synthetic IVE datasets* separately, resulting in 30 computational cases. A random number between 0 and 1 was generated and used as the mixture ratio ( $\alpha$ ) for each case as shown in Table 3.6. Accordingly, a total of 30 *updated BPMs* were generated. The *updated BPMs* were further used to test the hypothesis.

**Table 3.6.** The mixture ratio ( $\alpha$ ) of computational cases.

Case	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$\alpha$	0.73	0.75	0.18	0.45	0.64	0.59	0.13	0.85	0.34	0.85	0.36	0.84	0.96	0.33	0.61
Case	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
$\alpha$	0.83	0.40	0.97	0.57	0.54	0.21	0.60	0.20	0.41	0.29	0.05	0.08	0.35	0.14	0.38

### 3.6 Results

#### 3.6.1 The Context-Aware Design-Specific Data Obtained from the IVE Experiment



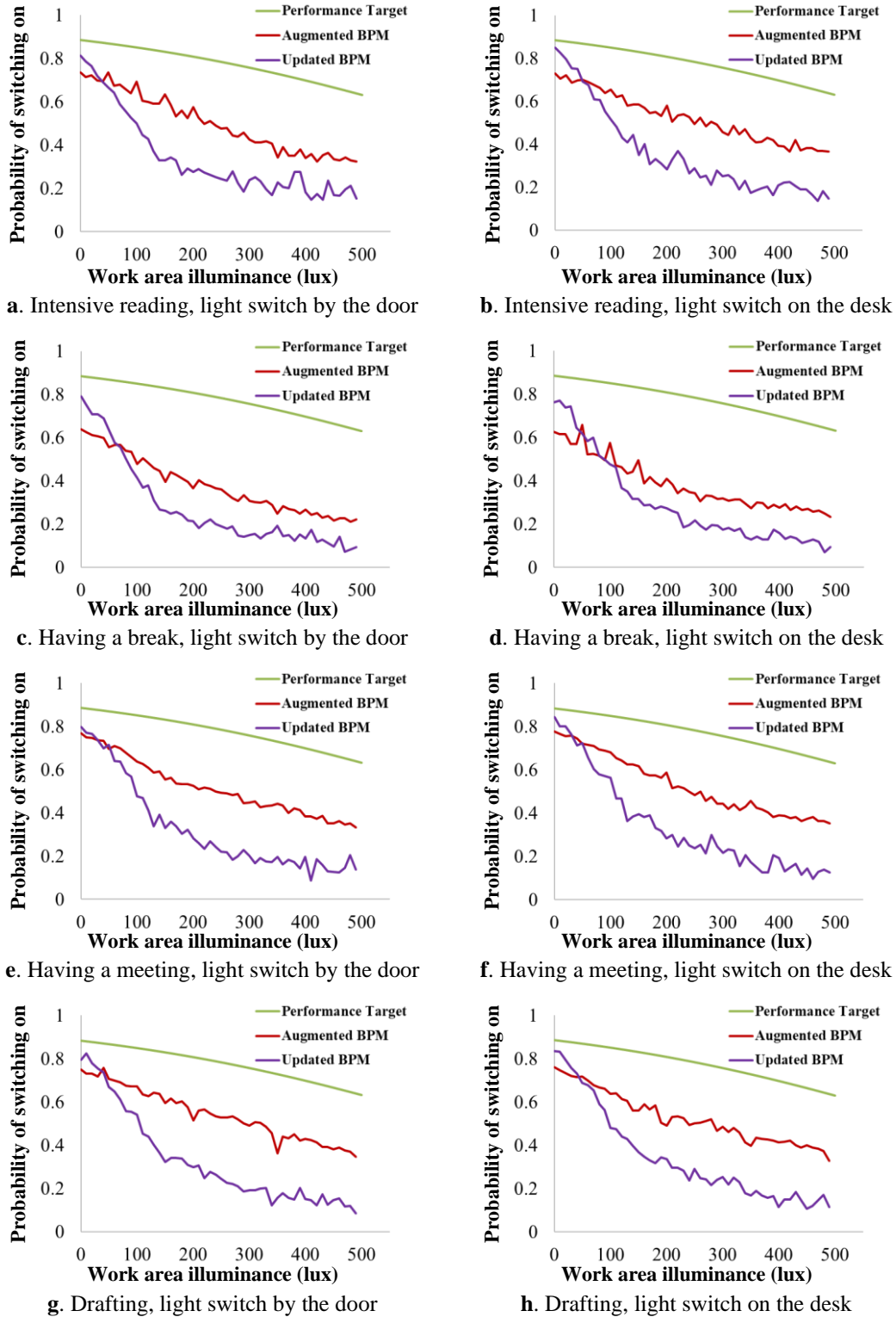
**Figure 3.11.** The context-aware design-specific data.

Figure 3.11 presents the means and the standard deviations of the context-aware design-specific data obtained from the IVE experiment of all participants classified by the office tasks and the light switch locations. Two observations show the qualitative effectiveness of the data:

- 1) The probability of switching on under low work area illuminance was higher than high work area illuminance regardless of the assigned tasks and the light switch locations. This general pattern is concordant with previous studies [5] [154] [100].
- 2) The probability of switching on with respect to the assigned task of “*having a break*” was slightly lower than the probability of switching on with respect to the other assigned tasks. Previous studies have shown that lighting needs are different based on task types [139] [140]; thus, the results show a certain level of qualitative validity.

However, the probability of switching on with respect to the location of light switches is visually similar under each plot in Figure 3.11. This observation is different from the literature, which generally suggests that location of switches influences use of switches by participants [35].

### 3.6.2 Descriptive Comparisons Between the Augmented BPMs and the Updated BPMs



**Figure 3.12.** Plots of the mean of the *augmented BPMs*, the mean of the *updated BPMs*, and the *performance target*.

Figure 3.12 presents plots of the probability of switching on versus the work area illuminance using the mean of the *augmented BPMs* and the mean of the *updated BPMs* classified by the office tasks and the switch locations. The mean of the *augmented BPMs* and the mean of the *updated BPMs* were calculated based on the 30 *augmented* and the 30 *updated BPMs*, respectively. The *performance target dataset* is also included in Figure 3.12. Several observations can be made based on Figure 3.12:

- The overall probability of switching on of the *augmented BPMs* was greater than the overall probability of switching on of the *updated BPMs* if the *augmented BPMs* and the *updated BPMs* are evaluated with the *performance target*.
- The probability of switching on of the *updated BPMs* between 0 and 50 lux were closer to the *performance target* than the probability of switching on of the *augmented BPMs*. One possible cause may be that the IVE experiment did not include illuminance between 0 and 50 lux, which may have impacted the computation procedure.
- The *augmented BPMs* had a lower probability of switching on than the *performance target*. This may imply that the *augmented BPMs* did not meet the *performance target*. Some contributing factors are discussed in the discussion section.

### 3.6.3 Hypothesis Testing

To test the hypothesis, two absolute errors were calculated including: 1) the absolute errors measured between the probability of switching on associated with the *updated BPM* and that associated with the *performance target dataset*, i.e.,  $E_1$  calculated using Equation (3-7) in Table 3.7 and 2) the absolute errors measured between the probability of switching on associated

with the *augmented BPM* and that associated with the *performance target dataset*, i.e.,  $E_2$  calculated using Equation (3-8) in Table 3.7.

**Table 3.7.** The absolute errors to prove the hypothesis.

Error	Measurement
$E_1$	$\frac{ \text{The probability of switching on from the } \textit{updated BPM} - \text{The probability of switching on from the } \textit{performance target dataset} }{\text{The probability of switching on from the } \textit{performance target dataset}}$ <div style="text-align: right;">(3-7)</div>
$E_2$	$\frac{ \text{The probability of switching on from the } \textit{augmented BPM} - \text{The probability of switching on from the } \textit{performance target dataset} }{\text{The probability of switching on from the } \textit{performance target dataset}}$ <div style="text-align: right;">(3-8)</div>

#### ***Tests on the Mean of $E_1$ and $E_2$ of All Cases***

A two-tailed *t-test* ( $\alpha = 0.05$ ) was first used to analyze whether  $E_1$  and  $E_2$  were significantly different. The hypothesis was defined as follows:

$$H_0: \text{mean of } E_1 - \text{mean of } E_2 = 0$$

$$H_1: \text{mean of } E_1 - \text{mean of } E_2 \neq 0$$

Table 3.8 shows the statistical test of significant differences between  $E_1$  and  $E_2$  of the 30 individuals. In 26 of the 30 cases, the differences between means of  $E_1$  and  $E_2$  were significant ( $p < 0.05$ ).



**Table 3.8.** The summary of tests of significant difference between  $E_1$  and  $E_2$ .

Case	Mean $E_1$	Std. $E_1$	Mean $E_2$	Std. $E_2$	$P$ -value	$H_0$
1	0.259	0.186	0.181	0.128	0.597	<i>Accept</i>
2	0.159	0.147	0.213	0.156	< 0.05	Reject
3	0.388	0.206	0.219	0.176	< 0.05	Reject
4	0.589	0.186	0.238	0.189	< 0.05	Reject
5	0.279	0.146	0.208	0.183	< 0.05	Reject
6	0.557	0.221	0.275	0.187	< 0.05	Reject
7	0.572	0.196	0.247	0.153	< 0.05	Reject
8	0.328	0.152	0.291	0.171	< 0.05	Reject
9	0.572	0.206	0.350	0.184	< 0.05	Reject
10	0.078	0.055	0.153	0.108	< 0.05	Reject
11	0.586	0.193	0.288	0.206	< 0.05	Reject
12	0.327	0.175	0.268	0.149	< 0.05	Reject
13	0.215	0.211	0.260	0.201	0.178	<i>Accept</i>
14	0.585	0.191	0.363	0.158	< 0.05	Reject
15	0.312	0.144	0.335	0.154	< 0.05	Reject
16	0.347	0.226	0.295	0.132	0.068	<i>Accept</i>
17	0.584	0.192	0.273	0.188	< 0.05	Reject
18	0.295	0.154	0.325	0.176	< 0.05	Reject
19	0.346	0.212	0.338	0.191	< 0.05	Reject
20	0.497	0.131	0.470	0.193	0.067	<i>Accept</i>
21	0.587	0.186	0.215	0.162	< 0.05	Reject
22	0.551	0.236	0.191	0.177	< 0.05	Reject
23	0.573	0.209	0.251	0.189	< 0.05	Reject
24	0.586	0.168	0.178	0.204	< 0.05	Reject
25	0.578	0.193	0.232	0.202	< 0.05	Reject
26	0.582	0.199	0.328	0.210	< 0.05	Reject
27	0.585	0.186	0.300	0.180	< 0.05	Reject
28	0.573	0.178	0.332	0.203	< 0.05	Reject
29	0.581	0.195	0.121	0.164	< 0.05	Reject
30	0.584	0.186	0.126	0.111	< 0.05	Reject

### ***Tests on the Mean of E<sub>1</sub> and E<sub>2</sub> in Cases 2, 10, 15 and 18***

Further tests were conducted to determine if the *updated BPM* was statistically more accurate than the *augmented BPM* in the four cases (i.e., cases 2, 10, 15, and 18) as suggested by their means. Tests were performed by a one-tailed *t-test* ( $\alpha = 0.05$ ) with the following hypothesis:

$$H_0: \text{mean of } E_1 - \text{mean of } E_2 = 0$$

$$H_1: \text{mean of } E_1 - \text{mean of } E_2 < 0$$

From Table 3.9, the null hypotheses were rejected in all cases, which show that the *updated BPM* was significantly more accurate than the *augmented BPM* in all four cases.

**Table 3.9.** The summary of tests of significant difference between E<sub>1</sub> and E<sub>2</sub>.

Case	Mean E <sub>1</sub>	Std. E <sub>1</sub>	Mean E <sub>2</sub>	Std. E <sub>2</sub>	P-value	H <sub>0</sub>
2	0.159	0.147	0.213	0.156	< 0.05	Reject
10	0.078	0.055	0.153	0.108	< 0.05	Reject
15	0.312	0.144	0.335	0.154	< 0.05	Reject
18	0.295	0.154	0.325	0.176	< 0.05	Reject

### ***Tests on the Remaining Cases***

The means of the remaining 22 cases showed that the *augmented BPMs* were more accurate than the *updated BPMs*. Since the *augmented BPMs* were expected to have significantly smaller errors than the *updated BPMs*, the hypothesis was defined as follows:

$$H_0: \text{mean of } E_1 - \text{mean of } E_2 = 0$$

$$H_1: \text{mean of } E_1 - \text{mean of } E_2 > 0$$

A one-tailed *t-test* ( $\alpha = 0.05$ ) was applied to test the hypothesis. The results of the hypothesis testing are shown in Table 3.10.

**Table 3.10.** The summary of the hypothesis testing.

Case	Mean E <sub>1</sub>	Std. E <sub>1</sub>	Mean E <sub>2</sub>	Std. E <sub>2</sub>	P-value	H <sub>0</sub>
3	0.388	0.206	0.219	0.176	< 0.05	Reject
4	0.589	0.186	0.238	0.189	< 0.05	Reject
5	0.279	0.146	0.208	0.183	< 0.05	Reject
6	0.557	0.221	0.275	0.187	< 0.05	Reject
7	0.572	0.196	0.247	0.153	< 0.05	Reject
8	0.328	0.152	0.291	0.171	< 0.05	Reject
9	0.572	0.206	0.350	0.184	< 0.05	Reject
11	0.586	0.193	0.288	0.206	< 0.05	Reject
12	0.327	0.175	0.268	0.149	< 0.05	Reject
14	0.585	0.191	0.363	0.158	< 0.05	Reject
17	0.584	0.192	0.273	0.188	< 0.05	Reject
19	0.346	0.212	0.338	0.191	< 0.05	Reject
21	0.587	0.186	0.215	0.162	< 0.05	Reject
22	0.551	0.236	0.191	0.177	< 0.05	Reject
23	0.573	0.209	0.251	0.189	< 0.05	Reject
24	0.586	0.168	0.178	0.204	< 0.05	Reject
25	0.578	0.193	0.232	0.202	< 0.05	Reject
26	0.582	0.199	0.328	0.210	< 0.05	Reject
27	0.585	0.186	0.300	0.180	< 0.05	Reject
28	0.573	0.178	0.332	0.203	< 0.05	Reject
29	0.581	0.195	0.121	0.164	< 0.05	Reject
30	0.584	0.186	0.126	0.111	< 0.05	Reject

The results of the hypothesis testing in Table 3.10 show that the  $p$ -values were smaller than 0.05 for all 22 cases; therefore, the null hypotheses were rejected. The tests suggest that the probability of switching on predicted by the *augmented BPMs* was significantly better than that predicted by the *updated BPMs*.

### 3.7 Discussion

The hypothesis testing at individual level showed mixed results. In the following, discussion regarding the context-aware design-specific data, the *augmented BPMs*, and the results of hypothesis tests are presented.

- The context-aware design-specific data involved variances associated with the probability of switching on (Figure 3.11). As mentioned in literature (e.g., [59] [6]), participants are clearly a source of the variances since different people may respond to the IVE experiment differently. In addition, factors such as the uses of virtual cues and experimental settings (e.g., the display quality, such as brightness and color, and the perception of participants about the IVE) may contribute to the variances of the context-aware design-specific data. Such variances can affect the accuracy of the *augmented BPMs*.
- Although the *augmented BPMs* did not underperform compared to the *updated BPMs* in most individual cases (i.e., 26 out of 30 cases), the means of the individual *augmented BPMs* were not close to the *performance target* (Figure 3.12). This result does not necessarily discount the effectiveness of the *augmented BPMs*. The issue may have resulted from the *existing BPM*. If the *existing BPM* significantly had lacked the ability to address the characteristics of a design, such as building configurations and occupant profiles, the *augmented BPMs* may have been significantly biased by data from the *existing BPM* and may not have reflected the target by using data from the *existing BPM*. Also, the IVE experiment may have failed to address contextual factors that the *performance target* addressed, which consequently affected the *augmented BPMs*. Finally, the *performance target* may

have been unrealistic and impossible to achieve. Therefore, a thorough investigation and evaluation of this issue is needed.

- In the four cases in which the *updated BPMs* had better performance than the *augmented BPMs*, the mixture ratios ( $\alpha$ ) were high (i.e., 0.75, 0.85, 0.61, and 0.97 corresponding to the cases 2, 10, 15, and 18, respectively). In those cases, the *updated BPMs* were constructed using more knowledge of the *synthetic IVE datasets* than the knowledge of the *existing BPM datasets*. In addition, the probability of switching on of the *synthetic IVE datasets* of those cases was close to or higher than the *performance target*. Consequently, high  $\alpha$ 's and the probability of switching on of the *synthetic IVE datasets* caused the *updated BPMs* to be closer to the *performance target* than the *augmented BPMs*. However, high  $\alpha$ 's contributed to heavy biases of the *updated BPMs* toward the *synthetic IVE datasets*, which therefore may not always generate better results if the *existing BPM datasets* are closer to the target. While the new framework tried to appropriately mix (i.e., with minimum bias) the *existing BPM dataset* and the *synthetic IVE dataset* toward the *performance target*, the *augmented BPM* incorporated balanced knowledge of the *existing BPM dataset* and the *synthetic IVE dataset*.
- Even though previous studies have stated that tasks [138] [139] [140] and switch locations [35] influence human-building interactions on lighting uses, the context-aware design-specific data in the present study showed a consistent pattern for tasks but not for switch locations. The situation with respect to light switch locations may be explained by several reasons: 1) the effectiveness of stimuli (e.g., visual and audio cues), which may not have been sufficient because the experiment assigned the office

tasks to the participants by using audio cues rather than the participants performing actual tasks, resulting in insufficient participant stimulation to realize how much lighting intensity they really needed, 2) the data collection procedure for the switch locations, which involved informing the participants about switch location by using audio and visual cues without actual interaction with the switch (e.g., walking to the switch at the door), so the participants might not have attempted to differentiate the difference in terms of access to the switch at different locations, and 3) the choice of locations, between which there may not have been any difference.

- The selection of the contextual factors (i.e., office tasks and locations of light switch) depended on the application and the knowledge of new contextual factors in literatures. In addition, the selection depended on the virtual reality technologies because the contextual factors needed to be model and experimented on in IVE. Therefore, to demonstrate and test the framework within our limitations of time and resources needed to develop IVEs, we chose the two contextual factors that have been commonly identified in literature [35][138][139][140].

The framework does not limit the number or the type of contextual factors that can be included. If other contextual factors are identified to have significant influence on human-building interactions related to lighting switch uses, they can be considered in the framework. Their inclusion will completely change the characteristics of context-aware design-specific data. Consequently, the *augmented BPM* will also change. Other types of contextual factors may increase the complexity of input parameters. However, the complexity of input parameters does not affect the application of the framework. With the current limitations of IVEs, some types of

contextual factors (e.g., senses and climates) may not be easy to effectively simulate in IVEs. Furthermore, including more contextual factors in IVE experiments increases experimental cost, resources used, and time consumed. To demonstrate the efficacy of the framework and due to limitations of IVE technologies, we only considered three parameters in the experiment. In summary, the framework does not preclude any additional contextual factors. Nevertheless, users of the framework need to consider the tradeoff between the desired number of contextual factors and the increase in time and resource needed along with the complexity.

- The framework is generic and parametric. It can take any BPM and context-aware design-specific data. For example, if there is a BPM modeling the performance of a multi-occupancy space along with an effective and reliable IVE to observe human-building interactions in the space, the framework can generate an *augmented BPM* for modeling the performance of the multi-occupancy space. Therefore, the characteristics of BPMs and the nature of human-building interactions do not affect the performance of the framework. However, the complexity in the development of BPMs and data collection of human-building interactions may vary, which influences the application of the framework. For instance, developing BPMs and collecting data in IVEs for analyzing building performance in multi-occupancy spaces are more complex. If such models and IVE data are available, the framework can produce an *augmented BPM*. Unfortunately, virtual reality technologies currently lack the capability to simulate human-building interaction scenarios in multi-occupancy spaces. Therefore, we chose a space with a single occupancy since the goal of the application is to show the efficacy of the framework.

### 3.8 Limitations of the Study

The major limitations of the study are discussed below:

- An approach to establish a *performance target* was not included in the framework. An approach to map design goals and objectives of buildings into a computational target is needed.
- As mentioned in the discussion, the most appropriate mixture may be obtained when context-aware design-specific data are relatively close to a *performance target*. If a target is unrealistic and context-aware design-specific data are relatively close to a *performance target*, an *augmented BPM* may be unrealistic as well. However, the framework does not yet have a method to assess whether a *performance target* is realistic.
- Since IVE experiments cannot be conducted for long periods of time, the capability of IVEs to collect longitudinal data is limited [38][155]. Accordingly, the IVE experiment was constructed using discrete events, which may not have thoroughly covered all possible situations.
- Currently, visual simulation is one of the most matured IVE capabilities. To simulate other sensations (e.g., Thermoception and olfaction), there is a need to integrate IVE with other equipment or devices (e.g., an external heating/cooling device to simulate thermal sensation). With limited resources (e.g., times, costs, and tools), we selected lighting performance to demonstrate the efficacy of the framework. Future work is needed to test the performance of the framework using other types of building performance models.



- Uncertainties of the components in the framework such as the *existing BPM*, the context-aware specific data, the *performance target*, and the computation procedures may have affected the development of the *augmented BPM*. The framework lacks uncertainty quantification and sensitivity analysis. Being able to analyze the uncertainty and sensitivity of the framework can contribute significantly to the improvement of the framework.

### 3.9 Conclusions and Future Work

The results of the hypothesis tests showed that the *augmented BPM* had higher accuracy than the *updated BPM* in most cases, which suggests that the GAN-based framework has generally better performance than the previous ANN-based greedy algorithm. However, in a few cases, the opposite was observed. Causes of the instability in performance of the framework require further research. In general, the selection of the *performance target* and the IVE experiments may have been the main causes. Therefore, further research is needed to create a technique that can analyze the uncertainty, sensitivity, and robustness of the framework including data from IVE experiments and a method to map *performance targets* between the design level and the computational level. Furthermore, methods to effectively and efficiently identify contextual factors (e.g., causality analysis [156], unsupervised approaches [157], and feature ranking [59]) need more research attention.

## **CHAPTER 4. ROBUSTNESS ANALYSIS OF A GAN-BASED FRAMEWORK FOR AUGMENTING BUILDING PERFORMANCE MODELS**

### **4.1 Introduction**

Building performance models (BPMs) are decision-support tools that designers often use to understand, analyze, and compare different design options based on design goals and objectives [59][123]. Significant research efforts have been devoted to improving the predictive accuracy of BPMs using several advanced techniques [158][159]. However, discrepancies between estimated building performance during design and actual measurements during operation have been extensively reported [9][115][116][14].

Recently, many studies have identified human-building interactions (e.g., occupant habitual behaviors and occupant responses to building contexts) as important factors influencing building performance [10][11], and these are sensitive to building contexts [6]. Traditionally, when BPMs are developed, this is often based on data of existing buildings with the inclusion of only a limited number of variables. Many other factors are considered as contextual factors, which are assumed to be constant across different application scenarios. However, if such an assumption no longer holds when designing a new building, contextual factors of the new building under design may have a significant impact on the application of *existing BPMs*. Consequently, applying the BPMs to the analysis of building performance may cause discrepancies between estimations during design and actual performance during operation. Therefore, being able to adjust existing BPMs with contextual factors such as human-building interactions in specific contexts can potentially enhance the accuracy and quality of BPMs [59].

Traditional methods (e.g., surveys [15][16], field studies [160][161], and laboratories [29][30]) have been effectively implemented to study human-building interactions in existing

buildings. However, capturing the interactions in non-existing buildings is challenging. Immersive virtual environments (IVEs) have applied to replicate situations that are unsafe, infeasible, or expensive to do in reality such as emergency situations [162][163], building designs [164], and occupant behaviors [165][38]. Furthermore, IVEs allow users to thoroughly control conditions and customize configurations of experiments. To that end, IVEs are used as tools for acquiring data of human-building interactions responding to specific the context of a new design, or called context-aware design-specific data.

The author proposed a GAN-based framework [114][166] to address the inclusion of contextual factors in *existing BPMs* for buildings under design. The framework used a generative adversarial network (GAN) to augment an *existing BPM* by using context-aware design-specific data acquired from IVE experiments. An IVE experiment simulated a building under design with specific contextual factors. The framework applied a *performance target* to guide the augmentation and generated an *augmented BPM*. Even though the efficacy of the GAN-based framework was demonstrated [166], the computational robustness of the framework was not assessed and investigated. Robustness is defined as the ability of a model to handle errors and uncertainty during execution [167]. Robustness analysis in the present study is defined according to the robust theorem in the work of Weisberg [167]. Robustness analysis is used to investigate whether the performance of a model remains robust when it is challenged by uncertainties that may occur during execution, including errors. Unavoidably, a model involves levels of uncertainty arising from different components, such as input parameters (e.g., input data) [168][169] and model structures (e.g., computational structures) [170]. For instance, if the uncertainty of input parameters is too large, the GAN as the computation component of the GAN-based framework may not be sufficiently robust, generating *augmented BPMs* with large

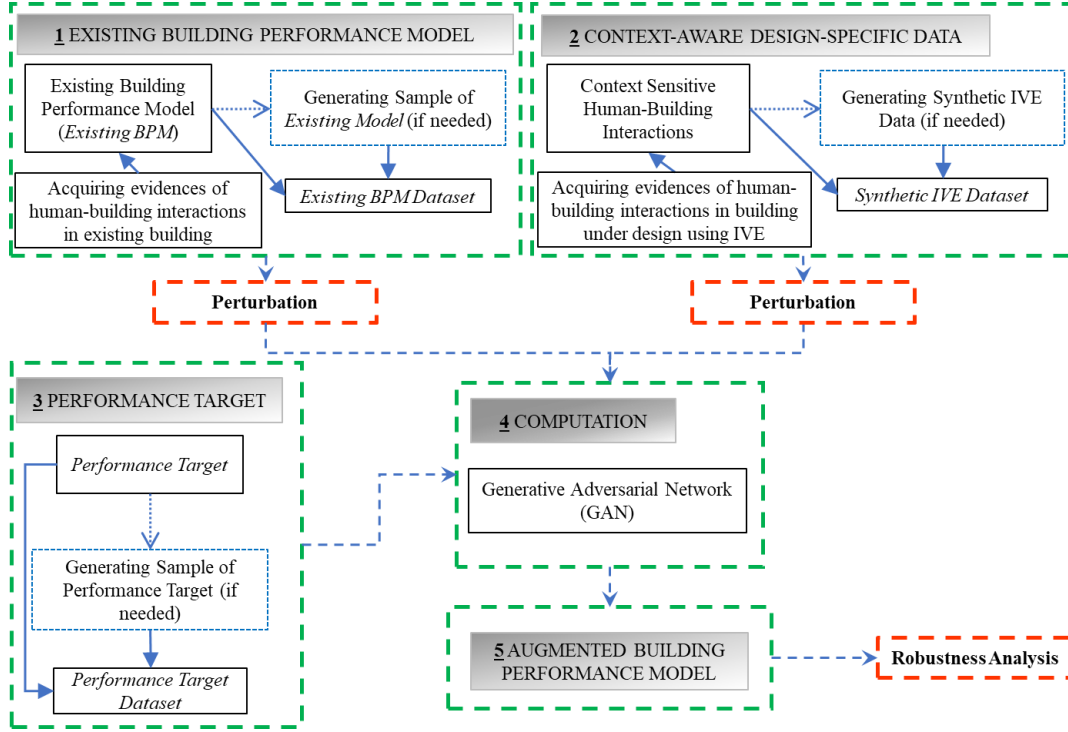
levels of uncertainty, and non-reliable extrapolation [171]. Thus, robustness analysis can assist users of the GAN-based framework to understand the impact of uncertainty, thereby gaining more confidence in using the framework for decision-making during design [172]. In general, uncertainties can be classified into two categories, namely aleatory uncertainty and epistemic uncertainty [173]. Aleatory uncertainty occurs due to the natural variability of a model system under study. It is also known as irreducible uncertainty and is thus ignored in the robustness analysis. Epistemic uncertainty occurs due to the lack of knowledge and information in analyses. It can be reduced if more information can be acquired [174]. For example, in the GAN-based framework, epistemic uncertainty may occur due to uncertainty related to the parameters of an *existing BPM* or context-aware design-specific data. Specifically, sources of such uncertainty may be errors in data, the varying degree of reliability of data collection tools, and the random nature of human-building interactions. Hence, this work focuses on analyzing the robustness of the GAN due to its epistemic uncertainty.

To analyze the robustness, knowledge of uncertainty of input parameters is required. One common approach to obtain uncertainty is using simulation. The perturbation method has been used to successfully simulate the uncertainty of input parameters for robustness analysis in several research studies related to machine learning, including image classifications [175][176], general classifications [177], and speech recognition [178][179]. There are many types of perturbation techniques, such as adding data noises, replacing data with random data, and altering data. To analyze the robustness of a model, the model runs on perturbed datasets, each of which represents a different level of uncertainty. The robustness of the model is assessed by comparing a baseline with the output generated using a perturbed input dataset. Typically, the output generated by a dataset without perturbation is the baseline.

The author proposes to complete and enhance the GAN-based framework by incorporating robustness analysis. The analysis applies the perturbation method to simulate uncertainty in data associated with the input parameters of the GAN. The robustness of the GAN is reflected by *augmented BPMs* obtained from the GAN trained on perturbed training datasets ( $A_{perturbation}$ ). The GAN is considered robust if it generates *augmented BPMs* that have comparable characteristics as the baseline, which is the *augmented BPM* generated by the GAN trained on non-perturbed training dataset ( $A_{non-perturbation}$ ).

This study focuses on the impact of uncertainty on computational models, an important issue discussed in several previous studies [168][169][170][171]. Extending to those studies, the robustness analysis can be a tool to enable better understanding and analyzing the impact of uncertainty. Moreover, robustness analysis further adds to applications of immersive virtual environments (e.g., [164][165][38]), where the uncertainty arising from such applications can be analyzed.

## 4.2 The GAN-Based Framework with Robustness Analysis



**Figure 4.1.** The GAN-based framework with the robustness analysis.

Figure 4.1 shows the workflow and components in the GAN-based framework supporting robustness analysis. In the following, a summary of the GAN-based framework is provided, and then the details of the robustness analysis are discussed. The complete documentation of the framework has been published in Chokwitthaya, et al. [166].

### 4.2.1 The GAN-Based Framework

There are five major components in the GAN-based framework (the green boxes in Figure 4.1) including: (1) an *existing BPM*, (2) context-aware design-specific data, (3) a *performance target*, (4) computation (i.e., the GAN), and (5) an *augmented BPM*.

### ***Existing Building Performance Model***

An *existing BPM* describes relationships of historical events (e.g., building environments and building characteristics) and observations (e.g., human-building interactions). Traditionally, data used to construct an *existing BPM* is acquired from existing buildings and human-building interactions with embedded contexts of existing buildings. Consequently, an *existing BPM* may not consider important contextual factors influencing human-building interactions in the context of a new building under design.

### ***Context-Aware Design-Specific Data***

Context-aware design-specific data describes human-building interactions influenced by contextual factors of a new design. For example, the Hunt model [1] uses work area illuminance as an independent variable to predict the status of light switch uses. However, other factors may also influence light switch uses such as the type of tasks (e.g., reading, meeting, and drafting) and the location of a light switch (e.g., a switch by a door or on a desk). For the Hunt model, the type of tasks and the location of a light switch are contextual since they are not included in the model. IVEs can be used to acquire such context-aware design-specific data [38][155].

### ***Performance Target***

A *performance target* is a performance metric (e.g., energy intensity of a space) that users define to satisfy the objectives of a building design [166]. Such a performance metric is converted into operational measures for computational purposes. The performance target is used to guide the combination of an *existing BPM* and the context-aware design-specific data so that the GAN can produce an *augmented BPM* with analytic results as close to the target as possible.

## ***Computation***

According to Chokwitthaya, et al. [166], a GAN [50] was implemented as the computation method in the framework. The GAN has a generator and a discriminator. The generator employs an ANN to learn a probability distribution and tries to predict outputs that follow a target distribution. The discriminator employs another ANN to discriminate outputs predicted by the generator and the target distribution. The GAN uses the concept of a two-player minimax game to train the generator and the discriminator. It generates an *augmented BPM* by augmenting an *existing BPM* using context-aware design-specific data guided by a *performance target*.

### **4.2.2 Robustness Analysis**

#### ***Introduction***

The objective of the robustness analysis in this chapter is to determine whether the GAN produces resilient *augmented BPMs*. If the GAN for particular assumptions about variability in inputs (e.g., uncertainty of the involved parameters) produces similar *augmented BPMs*, the GAN is considered robust for those assumptions. Robustness analysis identifies whether the GAN remains robust when input datasets are uncertain. An *augmented BPM* generated by the GAN trained on a non-perturbed training dataset ( $A_{non-perturbation}$ ) is considered as the baseline. The robustness analysis determines differences between  $A_{non-perturbation}$  and an *augmented BPM* generated by the GAN trained on perturbed training dataset ( $A_{perturbation}$ ). If  $A_{perturbation}$  is not significantly different from  $A_{non-perturbation}$ , the GAN is considered robust. Accordingly, the hypothesis is defined as follows:

$$H_0: A_{perturbation} - A_{non-perturbation} = 0$$

$$H_1: A_{perturbation} - A_{non-perturbation} \neq 0$$



## ***Perturbation***

The purpose of perturbation is to simulate variability in input datasets. In the present study, the author uses perturbations to add uncertainty to the input distributions represented by their respective training datasets [41]. The GAN acquires its knowledge using training datasets associated with the input parameters. To analyze the robustness of the GAN, perturbations have to be executed on the training datasets to make training datasets uncertain. Furthermore, various perturbations have to be modeled and considered. Perturbations may be performed using several techniques depending on the types of input parameters and purposes of studies. In image classification using machine learning, commonly applied perturbation techniques include injecting noises to images [180], changing information of images (e.g., watermarking, patching, and changing pixels) [181], and transforming image geometry [182][183]. In speech recognition using machine learning, perturbation techniques include adding noisy signal [184], making speech reverberated [185], and adding background noises [186]. Other examples to perturb datasets are inserting sentences in question answering systems [187] and using perturbation scale to alter data [188][189].

Overall, these perturbation techniques may be categorized into two main categories, namely additive perturbation (e.g., injecting noises to images, adding noisy signal, and adding background noises) and structured transformation (e.g., changing information of images, transforming image geometry, and making speech reverberated) [190]. The former adds additional unrelated data such as data noises to training datasets whereas the latter replaces data in training datasets with unrelated data or alter data in training datasets. The two categories serve different purposes and are meaningful in investigating and analyzing the robustness of GANs.

In addition, the selection of perturbation techniques depends on parameter types and practical circumstances that may cause uncertainty to the parameters. For instance, IVE experiments cannot simulate or include all possible scenarios that happen in the real world. Excluded scenarios may influence the uncertainty of training datasets and impact the robustness of a GAN. Adding data noises is an alternative to simulating additional uncertainty caused by the existence of excluded scenarios. Furthermore, human decisions, such as the choice of either switching a light switch on or off during experiments may be subjective or even involve wrong decisions, causing uncertainty in training datasets. Such uncertainty can be simulated by replacing parts of training datasets with unrelated data. Another example is that sensors used in experiments may involve uncertainty caused by unreliable measurements, which can be simulated by altering datasets associated with measurements.

One important factor in robust analysis is to determine which parameters or variables to study. As stated previously, the performance target guides the mix of two main input parameters of the GAN, namely the *existing BPM* and context-aware design-specific data. Therefore, it is assumed that the target is specific without uncertainty. Although the performance target can be perturbed in theory, there is no practical meaning in the scope of this study. On the other hand, the two main input parameters of the GAN are represented by various variables in this study, for instance, the Hunt model as an *existing BPM* with two variables, namely work area illuminance and the probability of switching on. Such variables are sources of uncertainty of the GAN because their data are often collected using instruments in experiments, field observations, or surveys, in which the occurrence of data uncertainty is unavoidable. However, it is possible that robustness analysis does not need to consider all variables. Although there is no specific criterion to determine the choice of variables, the decision is mainly based on the need of a particular

analysis. Another factor to consider when selecting the variables is the type of variables. For example, categorical variables (e.g., names and labels) are less likely to be subject to uncertainty. As a result, such variables may be excluded from perturbation in training datasets.

The level of perturbation is another important factor in robustness analysis. It helps to investigate the robustness of the GAN's response to different levels of uncertainty in training datasets. Generally, there is no standard and rule to define the level of perturbation. Most of previous studies defined the level of perturbation based on assumed amounts of uncertainty in variables that are believed to have an impact on the robustness of a study. For instance, Haghnegahdar and Raazavi [189] used a perturbation scale (e.g.,  $\pm 1\%$ ,  $\pm 5\%$ ,  $\pm 10\%$ , and  $\pm 20\%$ ) to distort datasets associated with input parameters and simulate uncertain input parameters for analyzing the robustness of earth and environmental system models.

### ***Perturbation Techniques***

#### ***Additive Perturbation***

Additive perturbation has been commonly used to perturb training datasets for analyzing the robustness of machine learning models in various studies [191][192][193][194]. It maintains the training data and adds additional unrelated data (e.g., data noises) to training datasets. The main purpose is to generate perturbed training datasets and investigate whether models have the ability to remain robust by maintaining the knowledge of training datasets even though the training datasets contain different levels of additive perturbations [195][196][197]. Furthermore, analysis of models' robustness regarding a different level of additive perturbation can be performed. For instance, Rolnick, et al. [198] generated data noises to investigate the robustness of their deep neural networks. They added noises up to 100 data for every training datum in

several experiments. The robustness of their deep neural networks was investigated across different levels of the data noises in the perturbed training datasets.

In the GAN-based framework, a major benefit of using additive perturbation is to investigate whether the GAN has the ability to remain robust when the GAN maintains the knowledge of training datasets associated with the *existing BPM* and the context-aware design-specific data even if the perturbed training datasets contain different levels of additive perturbation. Another benefit of using the additive perturbation is to explore whether the training datasets are sufficiently effective for the GAN to remain robust. If the GAN becomes un-robust when the training datasets involve a certain level of additive perturbation, revisions to the training datasets may need to be considered, such as acquiring more knowledge by conducting additional experiments to enhance the efficacy of the training datasets and robustness of the GAN.

Adding data noises is a traditional technique of additive perturbation. Among several categories of noise perturbations, one of the common forms is the additive white Gaussian noise (AWGN) [199]. AWGN allows direct control over the variance of noises, and data noises are generated using Gaussian (i.e., normal) distribution. As a result, if the variances of the training datasets are used to generate data noises, the data noises have similar variances as the training datasets, which makes data noises comparable to those in the training datasets. Furthermore, data noises are drawn from a Gaussian (i.e., normal) distribution, which is a common distribution applied to many experimental datasets [199]. Therefore, adding AWGNs is potentially an additive perturbation technique for analyzing the robustness of the GAN.

### Structured Transformation

Structured transformation investigates the robustness of models by reducing or distorting the knowledge of training datasets. Structured transformation has been widely applied to the robustness analysis in many research studies. For example, Liu, et al. [200] generated perturbed training datasets of traffic signs by scrawling and patching the signs in the original datasets. They used the signs with scrawls and patches to reduce the knowledge gained from the original datasets of signs and to re-train the classification model to investigate the classification accuracy of the model. The decrease in the accuracy of the model showed the decrease in its robustness. Engstrom, et al. [183] distorted the knowledge of images by rotating and transforming images, in which the normal images (i.e., images without rotation and transformation) were considered as training datasets. The rotations were performed by randomly rotating images between -30 and +30 degree and the transformations were performed by randomly transforming up to 10% of the pixels in images. They suggested that small rotations and transformations could significantly degrade accuracy and robustness of classifier models.

In the GAN-based framework, the main contribution of structured transformation is to inspect how reduced or distorted knowledge of training datasets impacts the robustness of the GAN. An additional advantage of using this technique is to explore the simulation of different levels of uncertainty involved in the input parameters that the GAN can tolerate and to which the GAN can remain robust.

### **4.3 Robustness Analysis of the GAN**

The analysis focused on understanding the robustness of the GAN and testing the hypothesis. The prediction of light switch uses in a single-occupancy office was used as an application case. It should be noted that this application case was fully reported in Chokwitthaya,

et al. [166]. To avoid unnecessary repetition, this section only provides a brief introduction of the major components, e.g., the *existing BPM*, the context-aware design-specific data, the *performance target*, the computation, and the *augmented BPM*. The training datasets associated with the *existing BPM* and context-aware design-specific data were perturbed by using the aforementioned perturbation forms. *Augmented BPMs* generated by the GAN trained on the perturbed training datasets were used to analyze the robustness of the GAN. The application is explained in detail in the following.

#### **4.3.1 Introduction of the Application Case**

According to Chokwitthaya, et al. [166], the light usage prediction model of Hunt [1] and Da Silva, et al. [98] were selected as the *existing BPM* and the *performance target*, respectively. Both models described the relationship between work area illuminance as an independent variable and the probability of switching on as a dependent variable. The datasets generated from the Hunt and Da Silva model were called “the *existing BPM dataset*” and “the *performance target dataset*”, respectively.

An IVE was used to simulate a single-occupancy office and acquire context-aware design-specific data corresponding to contextual factors. Contextual factors considered in the IVE experiment were office tasks (e.g., intensive reading, having a break, having a meeting, and drafting) and light switch locations (e.g., by the door and on the desk). Similar to the Hunt and Da Silva models, the independent and dependent variables included in the IVE experiment were the work area illuminance (lux) and the probability of switching on, respectively. Data corresponding to the contextual factors along with the independent and dependent variables were acquired from 30 students, including 18 males and 12 females, and was called the “*IVE dataset*”.

As discussed by Chokwitthaya, et al. [149], a GMM [143] was used to increase the number of IID samples based on the *IVE data* and generate a new dataset called the “*synthetic IVE dataset*”.

In data preprocessing, the *existing BPM dataset*, the *synthetic IVE dataset*, and the *performance target dataset* were normalized. The *existing BPM dataset* and the *synthetic IVE dataset* were split into training datasets and testing datasets with 70-30 splits, namely the *existing BPM training dataset*, the *existing BPM testing dataset*, the *synthetic IVE training dataset*, and the *synthetic IVE testing dataset*.

The GAN had a generator and a discriminator. The generator took the *existing BPM training dataset* and *synthetic IVE training dataset* as the input datasets. Before training the GAN, the generator was pre-trained on the combination of the *existing BPM training dataset* and the *synthetic IVE training dataset* to initialize its weights and biases. In every epoch, the generator gained knowledge by learning the *existing BPM training dataset* and *synthetic IVE training dataset* and making a prediction. The prediction that was closest to the *performance target* was considered as the *augmented BPM*. The discriminator determined differences between the prediction of the generator and the *performance target dataset*. The discriminator sent a feedback to the generator for improving its knowledge of mixtures (i.e., mixtures of the *existing BPM training dataset* and the *synthetic IVE training dataset*) and prediction in the next epoch.

### **4.3.2 Robustness Analysis**

#### ***Perturbation***

In this chapter, the contextual factors (i.e., task types and the location of light switches) are categorical variables. In addition, even if there had been uncertainty associated with them, the impact of uncertainty would have been reflected through the dependent variable (i.e., the probability of switching on). Therefore, these two contextual factors were not included in the

perturbation and the robustness analysis. The work area illuminance and the probability of switching on were subject to uncertainty. For example, the probability of switching on was subject to uncertainty because their data were obtained from human-building interactions, which tend to be sensitive to building contexts. Similarly, the work area illuminance was subject to uncertainty because it was measured using sensors for creating the *existing BPM* and simulated using an IVE to generate the context-aware design-specific data. Unavoidably, those experimental tools involved unknown levels of uncertainty. Therefore, the author perturbed the data of the work area illuminance and the probability of switching on in the training datasets using the two perturbation forms (i.e., additive perturbation and structured transformation).

Using additive perturbation, the author simultaneously perturbed the data of the probability of switching on and work area illuminance, the two variables in the *existing BPM training dataset* and the *synthetic IVE training dataset*, by adding data noises. It allowed the author to investigate and compare the overall impacts of the uncertain parameters, i.e., the *existing BPM* versus context-aware design-specific data on the robustness of the GAN. Using structured transformation, the author perturbed the two variables separately using two techniques, namely replacing the probability of switching on with random data and altering work area illuminance. This assisted the author to further investigate the impact of uncertainty of individual variables on the robustness of the GAN under specific circumstances. Table 4.1 summarizes parameters in the application along with their corresponding training datasets, variables subject to uncertainty, and the perturbation technique applied to each variable. The perturbed training datasets were called “*perturbed existing BPM training datasets*” and “*perturbed synthetic IVE training datasets*” when the *existing BPM training dataset* and the *synthetic IVE training dataset* were perturbed, respectively. It should be noted that this particular



design and administration of perturbation was performed only for understanding and demonstrating the impact of uncertain parameters and variables in this application. Other applications may have a different design and administration of perturbation depending on the purposes of the applications.

**Table 4.1.** The summary of the parameters and their corresponding components in the application.

Parameter	Training dataset associated with the parameter	Variables in the training dataset subject to uncertainty	Perturbation		
			Additive perturbation	Structured transformation	
			Adding data noises	Replacing probability of switching on with random data	Altering work area illuminance
Exiting BPM	<i>Existing BPM training dataset</i>	Probability of switching on	Yes	Yes	No
		Work area illuminance	Yes	No	Yes
Context-aware design-specific data	<i>Synthetic IVE training dataset</i>	Probability of switching on	Yes	Yes	No
		Work area illuminance	Yes	No	Yes

### ***Additive Perturbation***

To investigate the robustness of the GAN, AWGNs were added to the data of the probability of switching on and work area illuminance, the two variables in the *existing BPM training dataset* and the *synthetic IVE training dataset*. The AWGNs were added to the data of the variables. The simulation of AWGNs implemented a Gaussian (normal) distribution with zero means and specified variances ( $N(0, \sigma)$ ) to randomly generate the noisy data. The application used the variances of the probability of switching on and the work area illuminance as the variances of the Gaussian distribution when adding noises to their respective data. The author added various

amounts of AWGNs to the training datasets as shown in Table 4.2. For instance, the ratio of 10:1 denoted there were 10 actual datapoints to 1 AWGN in every 11 datapoints of the perturbed datasets. The perturbation ratios were used considering the limited resources available (e.g., computational costs and times) and the purpose of the application. The application preserved the actual data as the majority in the perturbed training datasets by limiting the ratio of the actual data to AWGN at 1:1. In other applications, more perturbation ratios may be used in the analysis. However, the trade-off between the resource needed and the number of perturbation ratios should be considered.

**Table 4.2.** The ratios of adding AWGNs to the data in the perturbed training datasets.

Case	The ratio of the data in the training datasets to AWGN	
1	10	1
2	10	3
3	10	5
4	10	7
5	10	10

### ***Structured Transformation***

Structured transformation perturbs the training datasets by using two techniques for different purposes. To investigate the robustness of the GAN due to the uncertain probability of switching on, portions of the training datasets with respect to the probability of switching on were replaced with random data. To analyze the robustness of the GAN on uncertain work area illuminance, the data with respect to the work area illuminance in the training datasets were altered using perturbation scales. In each perturbation, different levels of perturbation were assigned to investigate the responses of the GAN.

#### **Replacing Probability of Switching on with Random Data**

Ideally, the selection of the perturbation technique may reflect the practical circumstances causing uncertainty in the training datasets. For instance, a participant may be inconsistent at

different times when interacting with a light switch even if the lighting conditions are the same, which causes uncertainty in the probability of switching on. The structured transformation form is appropriate because the technique replaces a portion of the probability of switching on in the *existing BPM training dataset* and the *synthetic IVE training dataset* with unrelated data. Thus, it also reduces the knowledge of the training datasets. According to the definition of the robustness analysis, the robustness of the GAN was investigated in terms of the ability of the GAN to maintain robustness even if its knowledge of the probability of switching on was reduced in the training datasets.

The data points with respect to the probability of switching on in the training dataset were randomly replaced with random numbers between 0 and 1, the limits based on the nature of probability. According to Table 4.3, three perturbation ratios (i.e., 9:1, 7:3, and 5:5) were used to replace data of the probability of switching on in the *existing BPM training dataset* and *synthetic IVE training dataset*. For instance, the ratio of 9:1 denoted there were 9 actual data points to 1 randomized data point in every 10 data points in the perturbed training datasets. The selection of perturbation ratios was dependent mainly on the purpose of the study and the consideration of resource limitations (e.g., computational costs and time). The purpose of using different ratios was to assess the robustness of the GAN with respect to different amounts of knowledge about the probability of switching on in the training dataset. Furthermore, the application preserved the actual data as the majority in the perturbed training dataset by limiting the ratio of the actual data to replaced data at 1:1 (i.e., 5:5 in Table 4.3).

**Table 4.3.** The ratios of changing data of the probability of switching on to random data in the perturbed training datasets.

Case	The ratio of actual data to changed data	
1	9	1
2	7	3
3	5	5

#### Altering Work Area Illuminance

In this application, work area illuminance was one of the input parameters that was subject to uncertainty. Data of work area illuminance were often obtained from experimental tools (e.g., illuminance sensors and IVE simulation). The tools might alter the data of work area illuminance and thus induce uncertainty. To investigate the robustness of the GAN due to the uncertainty of work area illuminance and simulate the uncertainty of experimental tools, the author altered the data of work area illuminance in the training datasets using perturbation scales. The technique was based on the concept of the structured transformation and adaptations of previous perturbation techniques, namely transforming image geometry in Engstrom, et al. [183] and using the perturbation scales in Haghnegahdar and Raazavi [189].

The perturbation scales (i.e.,  $\pm 10\%$ ,  $\pm 30\%$ , and  $\pm 50\%$ ) were used to alter the data of work area illuminance in the training datasets. The perturbations were performed according to Equation (4-1). Even though a scale of 50% may appear to be impractical, it helps to assess the robustness of the GAN regarding extreme conditions of the experimental tools, for instance, if the illuminance sensors had been interrupted by external signals resulting in extreme errors in the measurements. The application used a perturbation interval at 20% and limited the perturbation at 50% because of resource limitations. Other applications may apply more and higher perturbation scales to assess the robustness of the GAN.

$$\text{Altered illuminance} = \text{illuminance} \pm (\text{illuminance} \times \text{perturbation scale}) \quad (4-1)$$

### ***Robustness Analysis, Hypothesis Testing, and Sensitivity Investigation***

The one-at-a-time technique [201][202] was applied to train the GAN using each perturbed training dataset once. A total of 23 *augmented BPMs* [i.e., non-perturbation + (5 cases of adding data noises + 3 cases of replacing probability of switching on with random data + 3 cases of altering work area illuminance) \* 2 input parameters] were generated.

The two-sample Kolmogorov-Smirnov test (*K-S test*) [203], a nonparametric test measuring the distance of two empirical distributions, was applied to test the hypothesis in the application. To test the hypothesis, a level of significance at  $\alpha = 0.05$  was applied to investigate the statistically significant difference between  $A_{\text{perturbation}}$  and  $A_{\text{non-perturbation}}$ . *P-values*  $\leq 0.05$  would have shown there was a significant difference between  $A_{\text{non-perturbation}}$  and  $A_{\text{perturbation}}$ , and the GAN was not sufficiently robust. On the other hand, *P-values*  $> 0.05$  would have shown there was no significant difference between  $A_{\text{non-perturbation}}$  and  $A_{\text{perturbation}}$ , and the GAN was robust.

Additionally, the *K-S statistic* obtained from the *K-S test* was used to assess the sensitivity of the GAN. The *K-S statistic* measured the distance between  $A_{\text{non-perturbation}}$  and  $A_{\text{perturbation}}$ . To determine the sensitivity of the GAN, the pairwise comparisons of the *K-S statistic* across  $A_{\text{perturbation}}$  generated from the GAN trained on the *perturbed existing BPM training dataset* and the *perturbed synthetic IVE training dataset* within the same level of perturbation (e.g., perturbation ratio and scale) were analyzed. For instance, if the *K-S statistic* associated with  $A_{\text{perturbation}}$  generated from the GAN trained on the *perturbed existing BPM training dataset* and the *synthetic IVE training dataset* at 10:1 perturbation ratio had been lower than that from the GAN trained on the *existing BPM training dataset* and the *perturbed synthetic IVE training*

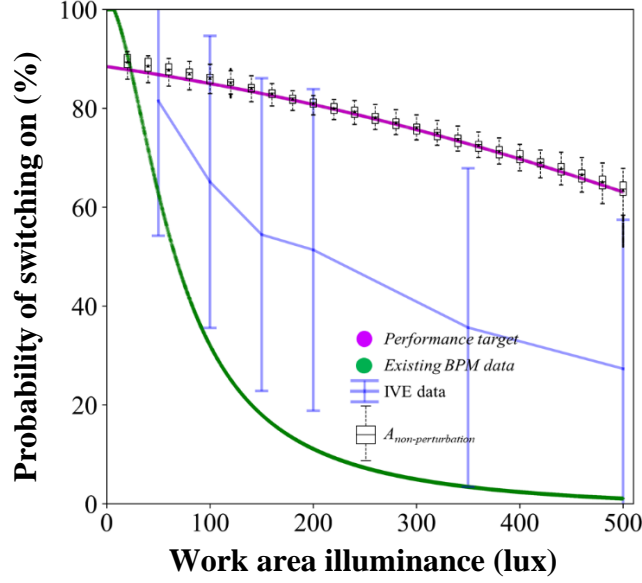
*dataset*, the GAN would have been less sensitive to the *existing BPM* than the context-aware design-specific data.

#### 4.4 Results and Discussion

Results and discussion are organized in three parts, 1) non-perturbation, 2) additive perturbation performed by adding data noises, and 3) structured transformation performed by replacing the probability of switching on with random data and altering work area illuminance.

Figure 4.2 illustrates the  $A_{non-perturbation}$ , the *existing BPM training dataset*, the *performance target dataset*, as well as the means and standard deviations of the *synthetic IVE training dataset* by plotting the probability of switching on versus their corresponding work area illuminance (lux). Figure 4.3, 4.5, and 4.7 demonstrate comparisons between  $A_{non-perturbation}$  and  $A_{perturbation}$  corresponding to each perturbation and its levels. In Figure 4.2, 4.3, 4.5, and 4.7, boxplots are used to demonstrate the variances representing the uncertainty of  $A_{non-perturbation}$  and  $A_{perturbation}$ . Table 4.4 to 4.6 contain the *p-values*, which were used to statistically evaluate the robustness of the GAN. Figure 4.4, 4.6, and 4.8 show plots of *K-S statistic* associated with levels of perturbation in each perturbation case supporting the sensitivity analysis of the GAN.

#### 4.4.1 Non-perturbation



**Figure 4.2.** *Augmented BPMs* corresponding to non-perturbed training dataset.

Figure 4.2 shows the efficacy of the GAN for generating *an augmented BPM* (represented by  $A_{non-perturbation}$ ) that reached the performance target. However, according to the boxplots, uncertainty existed in  $A_{non-perturbation}$  even though the input parameters were not perturbed. The finding agrees with the fact that uncertainty always exists in BPMs mentioned in the literature [172][204]. Several factors may contribute to the occurrence of uncertainty, such as the nature of the GAN (i.e., aleatory uncertainty), the structure of the GAN, and the completeness of the input parameters. Such factors may need attention in future research.

#### 4.4.2 The Additive Perturbation

##### *Adding Data Noises*

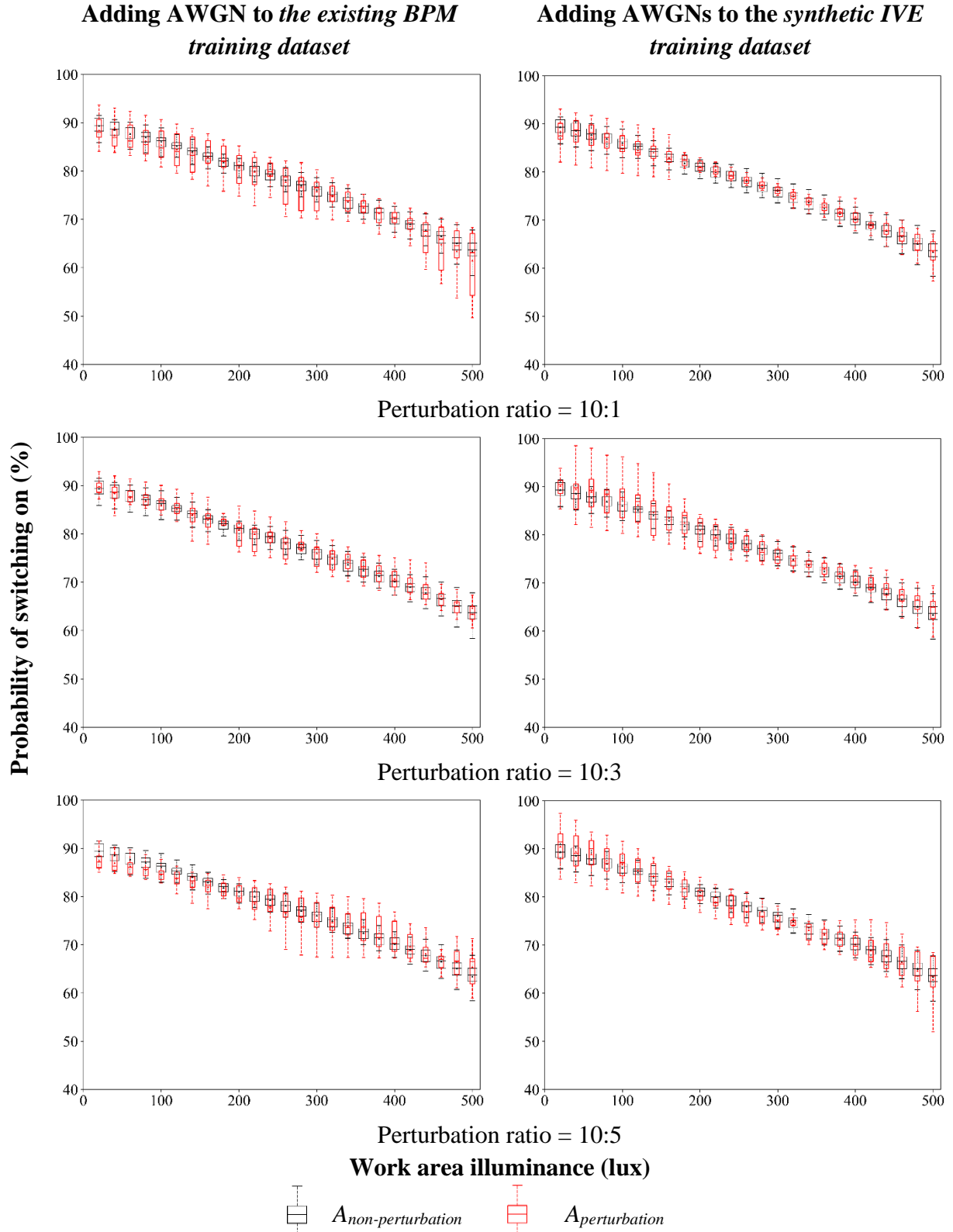
Figure 4.3 illustrates comparisons of  $A_{non-perturbation}$  and  $A_{perturbation}$  with respect to adding data noises to the probability of switching on and work area illuminance. The AWGNs were added according to the perturbation ratios described in Table 4.2, i.e., 10:1, 10:3, 10:5, 10:7, and 10:10. Figure 4.3 reveals that the uncertainty of  $A_{perturbation}$  was slightly higher than that of  $A_{non-}$

*perturbation* since the variances in the boxplots associated with  $A_{perturbation}$  were larger than those associated with  $A_{non-perturbation}$ . The results suggest adding noises marginally influenced the uncertainty of  $A_{perturbation}$ .

The influences of adding noises did not significantly impact the robustness of the GAN because the *p-values* were greater than 0.05 in all cases as shown in Table 4.4. Therefore, the GAN remained robust when the training datasets of the input parameters were perturbed by adding data noises. According to the results of the robustness analysis, the GAN remained robust in all perturbation ratios in both cases, the *perturbed existing BPM training datasets* and the *perturbed synthetic IVE training datasets*. The results suggest that the GAN was able to remain robust even if the level of noises was increased to 100% if the original knowledge of the training datasets was intact in the training datasets. However, this is only one case that shows the GAN remained robust when the training datasets were perturbed. The GAN may not remain robust in other applications using the adding data noise technique.

According to the *K-S statistics* in Figure 4.4 **Error! Reference source not found.**, the data pattern of *K-S statistics* was not consistent across the perturbation ratios. Hence, it is unclear whether the GAN is more sensitive to the *existing BPM training dataset* or the *synthetic IVE training dataset*. Consequently, the GAN was not more sensitive to the *existing BPM* or the context-aware design-specific data when the training datasets were perturbed by adding data noises.

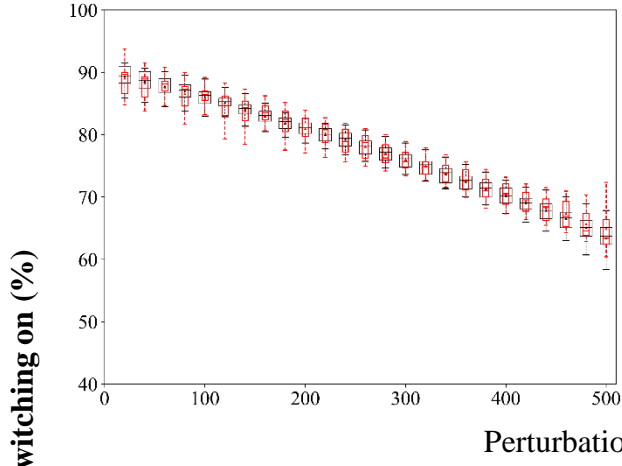




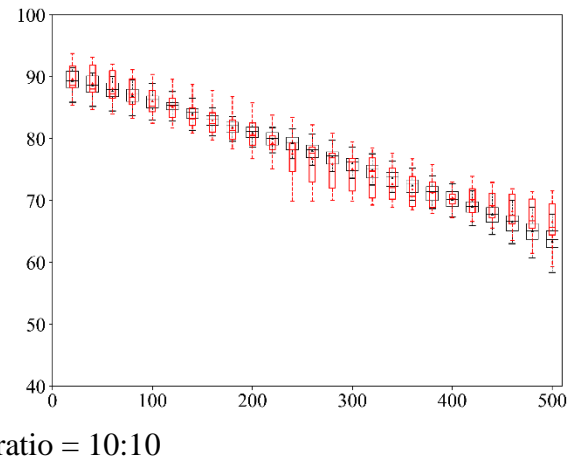
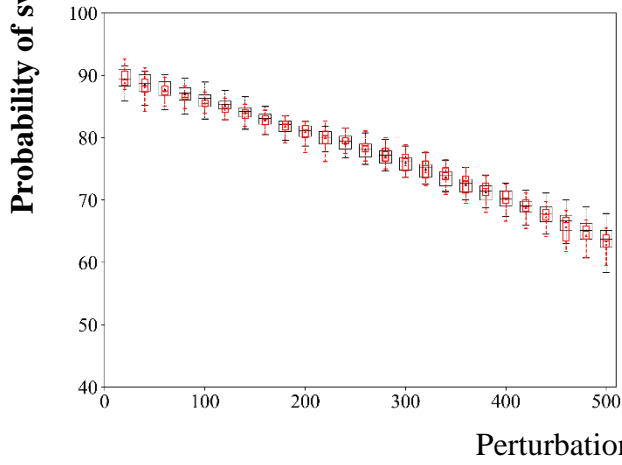
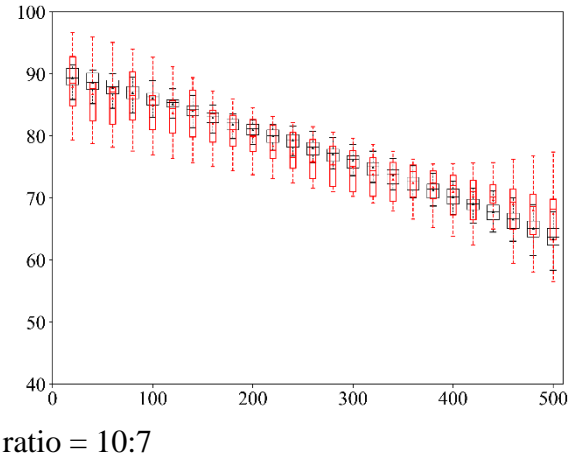
**Figure 4.3.** Augmented BPMs corresponding to adding data noises.

(figure cont'd.)

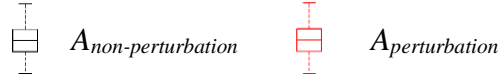
**Adding AWGN to the existing BPM training dataset**



**Adding AWGNs to the synthetic IVE training dataset**

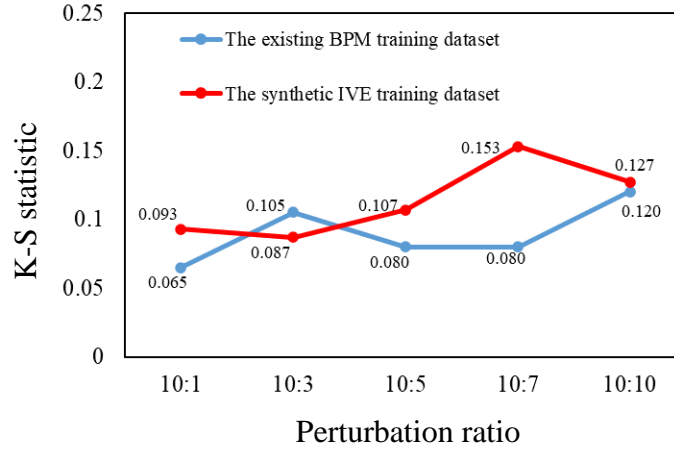


**Work area illuminance (lux)**



**Table 4.4.** *P-values* corresponding to adding data noises.

Perturbation ratio	The existing BPM training dataset	The synthetic IVE training dataset
	<i>P-value</i>	<i>P-value</i>
10:1	0.793	0.532
10:3	0.221	0.628
10:5	0.545	0.362
10:7	0.545	0.059
10:10	0.112	0.180

**Figure 4.4.** *K-S statistics* corresponding to adding data noises.

#### 4.4.3 The Structured Transformation

##### *Replacing the Probabilities of Switching on With Random Data*

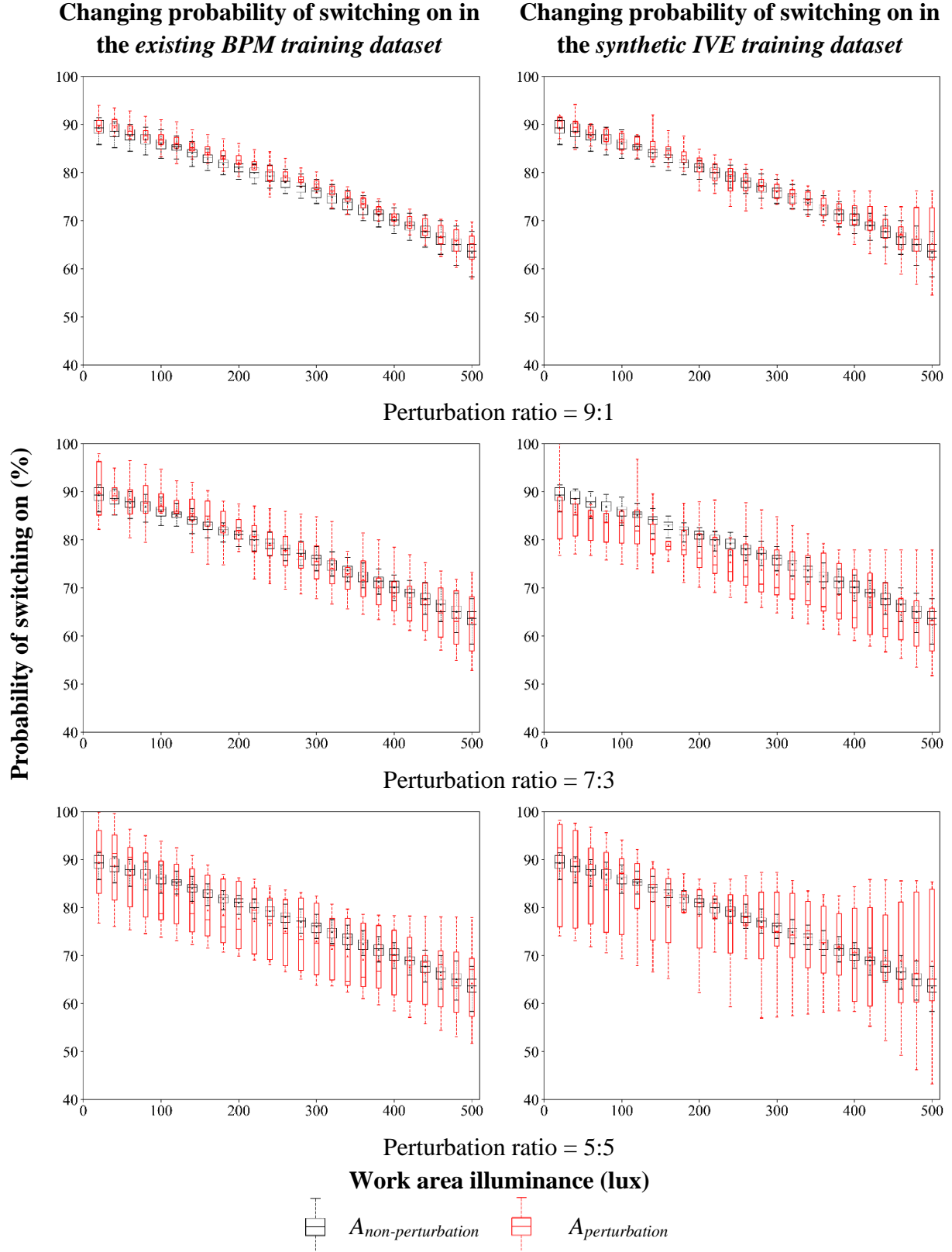
The comparisons of  $A_{non-perturbation}$  and  $A_{perturbation}$  with respect to replacing the probability of switching on with random data between 0 and 1 are illustrated in Figure 4.5. The perturbation ratios (i.e., 9:1, 7:3, and 5:5) defined the levels of perturbation. The results of the uncertainty, robustness, and sensitivity are explained in the following.

Figure 4.5 shows that the variances in the boxplots increased when the perturbation ratio increased, indicating the increases of the uncertainty of  $A_{perturbation}$ . This observation implies that changing the probability of switching on in the training datasets with random data contributed to the increases of the uncertainty of  $A_{perturbation}$ .

The results of hypothesis testing presented in Table 4.5 reveal that when the perturbation ratio increased, the number of rejected cases of the null hypothesis ( $p\text{-value} < 0.05$ ) also increased. They suggest replacing data in the training datasets with random number reduced the level of knowledge in the GAN about the training datasets and thus reduced the performance of the GAN, leading to decreases of the robustness.

According to Table 4.5, even though the increases of perturbation ratios reduced the robustness of the GAN, perturbing the probability of switching on in the *existing BPM training dataset* had less contribution to the reduction of the robustness than perturbing that in the *synthetic IVE training dataset*. When the probability of switching on in the *existing BPM training dataset* was perturbed, the null hypothesis was rejected in one case in which the perturbation ratio was set to 5:5. Nevertheless, the null hypothesis was rejected in two cases which the same variable was perturbed in the *synthetic IVE training dataset*.

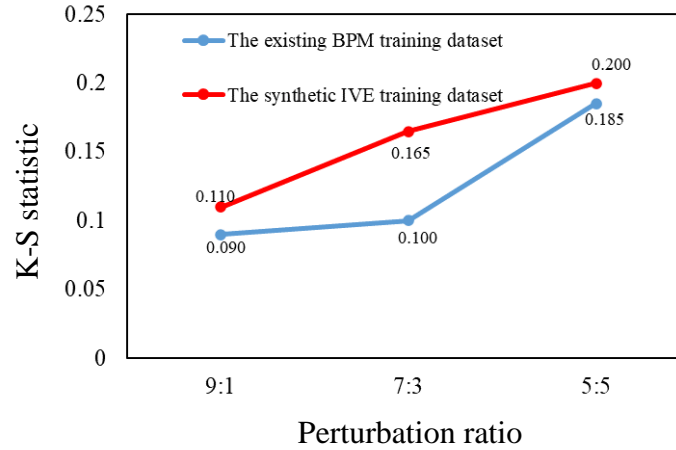
According to Figure 4.6, the *K-S statistics* associated with the *perturbed existing BPM training datasets* were lower than those associated with the *perturbed synthetic IVE training datasets* in all perturbation ratios. The finding implies that the GAN was less sensitive to the *existing BPM* than the context-aware design-specific data when the data of the probability of switching on were replaced by random data.



**Figure 4.5.** Augmented BPMs corresponding to replacing the probability of switching on with random data.

**Table 4.5.** *P-values* corresponding to replacing the probability of switching on with random data.

<b>Perturbation ratio</b>	<b>The existing BPM training dataset</b>	<b>The synthetic IVE training dataset</b>
	<i>P-value</i>	<i>P-value</i>
9:1	0.394	0.177
7:3	0.270	<0.05
5:5	<0.05	<0.05



**Figure 4.6.** *K-S statistic* corresponding to replacing the probability of switching on with random data.

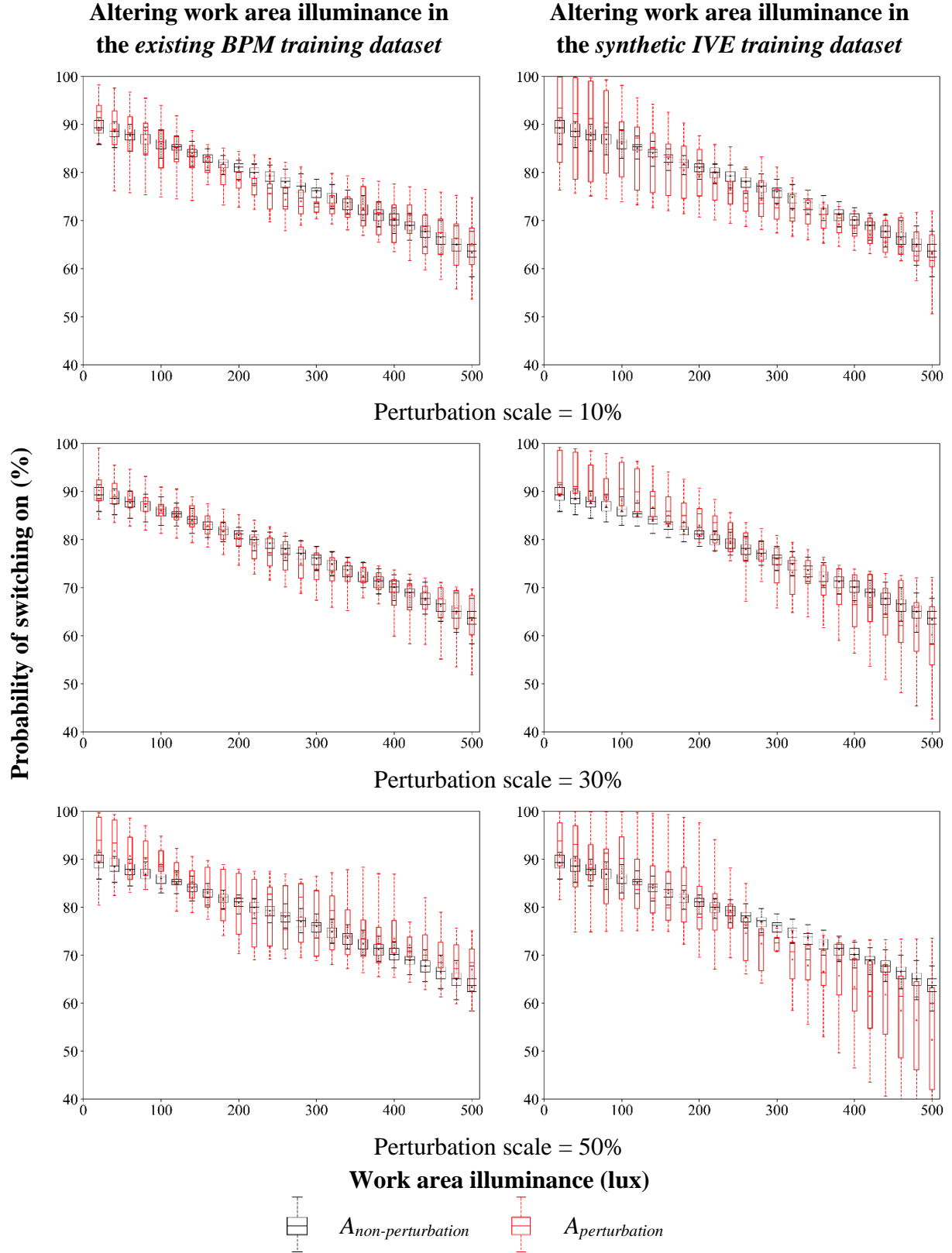
### ***Altering Work Area Illuminance***

According to the boxplots associated with  $A_{perturbation}$  in Figure 4.7, increases of the perturbation scale for altering work area illuminance in both *existing BPM training dataset* and *synthetic IVE training dataset* increased the variances and uncertainty of  $A_{perturbation}$ .

The *p-values* in Table 4.6 were less than 0.05 in 5 out of 6 cases, which shows that  $A_{perturbation}$  and  $A_{non-perturbation}$  were significantly different in most cases. The result suggests that altering the data associated with work area illuminance may have significantly impacted the robustness of the GAN. Similar to replacing the original training dataset with random data, altering work area illuminance in the *existing BPM training dataset* had less influence on the reductions of the robustness than altering that in the *synthetic IVE training dataset*. The null hypothesis was rejected in two cases in which the perturbation scale was 30% and 50% when the

work area illuminance in the *existing BPM training dataset* was altered. However, the null hypothesis was rejected in all cases when perturbing the same variable in the *synthetic IVE training dataset*.

According to Figure 4.8, the *K-S statistics* associated with the *perturbed existing BPM datasets* were lower than those associated with the *perturbed synthetic IVE datasets* throughout the perturbation scales. The result suggests that the GAN was less sensitive to the *existing BPM* than the context-aware design-specific data.

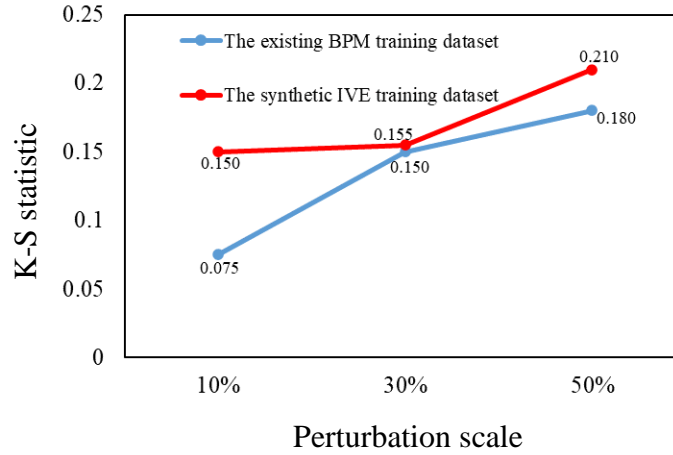


**Figure 4.7.** Augmented BPMs corresponding to altering work area illuminance.



**Table 4.6.** *P-values* corresponding to altering work area illuminance.

Perturbation scale	The existing <i>BPM</i> training dataset	The synthetic <i>IVE</i> training dataset
	<i>P-value</i>	<i>P-value</i>
10%	0.628	< 0.05
30%	< 0.05	< 0.05
50%	< 0.05	< 0.05



**Figure 4.8.** *K-S statistic* corresponding to altering work area illuminance.

#### 4.5 Limitations of the Study

Major limitations of the study include the following:

- The application only investigated the robustness of the GAN regarding input parameters (the *existing BPM* and the context-aware design-specific data). The robustness associated with other components, such as the structure of the GAN, was excluded in the application. Research attention on the robustness analysis of other components is needed in the future.
- The application limits the study of the robustness to three perturbation techniques. Other techniques that may impact on the robustness of the GAN should be investigated to be able to comprehensively discuss the robustness of the development of *augmented BPMs* using the GAN-based framework.

- More perturbation levels and smaller intervals should be considered to investigate the robustness of the GAN. Due to limitations of resources (e.g., computational costs and time) and purposes of the application, the perturbation was limited with large intervals between low and high perturbation levels.
- The application studied the robustness according to one specific *existing BPM*, one design, and one *performance target*. Practically, users do not rely only on only one of each. They may create several alternatives of *existing BPMs*, designs, and *performance targets*. However, the application was limited to only one of each. The robustness of the GAN regarding other *existing BPMs*, designs, and *performance targets* was excluded in the application.

## 4.6 Conclusions and Future Work

The robustness analysis shows effectiveness in identifying the robustness, uncertainty, and sensitivity of the GAN associated with the *augmented BPMs*. The results of the hypothesis tests show that adding noises slightly impacted the robustness of the GAN but not in any statistically significant manner. In addition, adding data noises marginally increased uncertainty of the *augmented BPMs*. Replacing data in the training datasets with noises and altering data in the training datasets caused significant reduction in the robustness of the GAN and increased the uncertainty of the *augmented BPMs*. The findings agree with previous works mentioning the impacts of perturbations causing reductions on the robustness of machine learning [190][205]. Furthermore, the GAN was more sensitive to the context-aware design-specific data than the *existing BPM*.

The uncertainty, robustness, and sensitivity are dependent on several factors such as input parameters, the computational structure, and the nature of the GAN. This study only investigated

the robustness of the GAN with respect to input parameters. Therefore, future research is needed to investigate the other factors that may have significantly impacts on the robustness of the GAN. A GAN-based framework with comprehensive robustness analysis could potentially reduce risks in design and enhance the quality of design.

## CHAPTER 5. CONTRIBUTIONS, CONCLUSIONS, AND FUTURE WORKS

### 5.1 Contributions

The contributions of the dissertation are as follows:

- The main contribution of the dissertation is a framework that biases an *existing BPM* to better reflect the contexts of a building under design to improve building performance simulation. Consequently, the framework could potentially reduce performance discrepancy between estimations during design and actual building as well as enhance the building design stage.
- The framework assists users in integrating an *existing BPM* with context-aware design-specific data addressing specific human-building interactions in contexts of a design.
- The framework offers an alternative approach using an IVE simulating a new design, allowing users to obtain occupancy data. The IVE overcomes conventional occupancy data collection approaches (e.g., sensing, field studies, and surveys) in many aspects. First, unlike the conventional approaches that allow users to collect occupancy data only in existing buildings, the IVE allows users to simulate and collect data in a specific building. Second, collected occupancy data were proven to be consistent with the actual data in a real environment. Last, it allows users to simulate a variety of building contexts.
- Robustness analysis is incorporated into the framework, which contributes to gains in confidence and supports users in making decision during using the framework.

## 5.2 Conclusions

The novel framework for augmenting *existing BPMs* was developed to increase the estimation performance during design. The framework customizes *existing BPMs* to address the contextual factors of a building under design. The aim of the research was to reduce performance discrepancy between estimations during design and the performance of the actual building. The framework uses IVEs to simulate buildings under design. Human-building interactions responding to contextual factors in a new design are acquired from the IVE experiments and are called context-aware design-specific data. Machine learning techniques assist computational operations to customize *existing BPMs* using context-aware design-specific data and generate *augmented BPMs* that have better estimation performance than *existing BPMs*. The framework preserves the general predictive power of *existing BPMs* while addressing specific human-building interactions in the context of a new design identified by users. As a result, the framework produces more representative BPMs (*augmented BPMs*) specific to the building under design to improve prediction accuracy rather than the generalized predictions from *existing BPMs*. The research was performed in three major stages including: 1) the potential of the framework, 2) the improvement of the framework, and 3) the robustness, uncertainty, and sensitivity of the framework.

The potential of the framework was investigated to ensure that it was effective and practical for *augmenting BPMs* to increase their performance during design. The application of the framework on a simulated single office confirmed the potential of the framework. To customize an *existing BPM*, an ANN-based greedy algorithm combined an *existing BPM* with context-aware design-specific data by using manually assigned mixture ratios. Based on the

results of the application as stated in chapter 2, the framework shows the potential to improve the prediction accuracy of *existing BPMs*.

Chapter 3 improves the framework to determine the appropriate combinations of an *existing BPM* and context-aware design-specific data. GANs assist the combinations (GAN-based framework). The framework allowed designers to guide the combinations using a *performance target*. According to the application, *augmented BPMs* obtained from the GAN-based framework were evaluated with *updated BPMs* obtained from the ANN-based greedy algorithm framework introduced in chapter 2. The evaluations show that the GAN-based framework produced *augmented BPMs* that had better performance than *updated BPMs* in most cases of the comparisons.

Chapter 4 completes and enhances the framework by incorporating robustness analysis on the computation of the framework (the GAN). It shows the efficacy in identifying the robustness, uncertainty, and sensitivity associated with the *augmented BPMs* that the framework produces. As a result, users of the GAN-based framework can have a better understanding of the impact of uncertainty and thus gain confidence in using the framework for decision-making during design.

### **5.3 Future Works**

Even though the framework was proven to effectively enhance BPMs toward specific designs, it can be improved in many aspects:

- According to chapter 3, the variance of the context-aware design-specific data obtained in the IVE experiment was relatively large. It may not have accurately represented the data of human-building interactions in the building under design. A criterion for selecting participants and designing cues in IVE experiments to properly serve the purposes of the

uses of the framework should be considered in the future research. The criterion would help users of the framework to reduce variance of context-aware design-specific data and obtain more accurate *augmented BPMs* toward specific designs.

- Recently, visual sensation is the most mature sensation in IVE experiments. When technologies for experimenting using other sensations in IVEs are available, combinations of sensations (e.g., visual, acoustic, and thermal) are needed to generate more accurate and complete context-aware design-specific data. Consequently, the performance of *augmented BPMs* could be further improved.
- The performance of the GAN would be improved by using a dynamically adaptive learning technique. The components of the GAN, such as the number of neurons, number of layers, number of iterations, and learning rate, could be adjusted dynamically based on situational training. The technique could potentially reduce computational costs (e.g., times and resources) and uncertainty during training as well as improve performance of the GAN, leading to producing more appropriate *augmented BPMs*.
- The robustness, uncertainty, and sensitivity analyses regarding to other components besides the GAN (e.g., *existing BPMs*, context-aware design-specific data, *performance targets*, and the system of the GAN) need comprehensive exploration to better understand the contributions of all the components as well as to allow users to gain confidence in using the framework.

## APPENDIX. COPYRIGHT INFORMATION

### Chapter 2 Copyright

1/12/2020

Rightslink® by Copyright Clearance Center



RightsLink®



Home



Help



Email Support



Sign in



Create Account



#### Combining context-aware design-specific data and building performance models to improve building performance predictions during design

Author: Chanachok Chokwitthaya,Yimin Zhu,Robert Dibiano,Supratik Mukhopadhyay

Publication: Automation in Construction

Publisher: Elsevier

Date: November 2019

Published by Elsevier B.V.

Please note that, as the author of this Elsevier article, you retain the right to include it in a thesis or dissertation, provided it is not published commercially. Permission is not required, but please ensure that you reference the journal as the original source. For more information on this and on your other retained rights, please visit: <https://www.elsevier.com/about/our-business/policies/copyright#Author-rights>

BACK

CLOSE WINDOW

### Chapter 3 Copyright

7/29/2020

Rightslink® by Copyright Clearance Center



RightsLink®



Home



Help



Email Support



Sign in



Create Account



#### Augmenting building performance predictions during design using generative adversarial networks and immersive virtual environments

Author: Chanachok Chokwitthaya,Yimin Zhu,Supratik Mukhopadhyay,Edward Collier

Publication: Automation in Construction

Publisher: Elsevier

Date: November 2020

© 2020 Elsevier B.V. All rights reserved.

Please note that, as the author of this Elsevier article, you retain the right to include it in a thesis or dissertation, provided it is not published commercially. Permission is not required, but please ensure that you reference the journal as the original source. For more information on this and on your other retained rights, please visit: <https://www.elsevier.com/about/our-business/policies/copyright#Author-rights>

BACK

CLOSE WINDOW



## REFERENCES

- [1] D. R. G. Hunt, "Predicting artificial lighting use- a method based upon observed patterns of behavior," *Lighting Research and Technology*, vol. 12, no. 1, pp. 7–14, 1980, doi: 10.1177/096032718001200102.
- [2] G. R. Newsham, "Manual control of window blinds and electric lighting: Implications for comfort and energy consumption," *Indoor and Built Environment*, vol. 3, no. 3, pp. 135–144, 1994, doi: 10.1177/1420326X9400300307.
- [3] F. Oldewurtel, A. Parisio, C. N. Jones, D. Gyalistras, M. Gwerder, V. Stauch, B. Lehmann, and M. Morari, "Use of model predictive control and weather forecasts for energy efficient building climate control," *Energy and Buildings*, vol. 45, pp. 15–27, 2012, doi: 10.1016/j.enbuild.2011.09.022.
- [4] D. Sturzenegger, D. Gyalistras, M. Morari, and R. S. Smith, "Model Predictive Climate Control of a Swiss Office Building," *Implementation, Results, and Cost–Benefit Analysis*, vol. 24, no. 1, pp. 1–12, 2016, doi: 10.1109/TCST.2015.2415411.
- [5] C. F. Reinhart, "Lightswitch-2002: A model for manual and automated control of electric lighting and blinds," *Solar Energy*, vol. 77, no. 1, pp. 15–28, 2004, doi: 10.1016/j.solener.2004.04.003.
- [6] Y. Zhu, S. Saeidi, T. Rizzuto, A. Roetzel, and R. Kooima, "Potential and challenges of immersive virtual environments for occupant energy behavior modeling and validation: A literature review," *Journal of Building Engineering*, vol. 19, pp. 302–319, 2018, doi: 10.1016/j.jobbe.2018.05.017.
- [7] V. Fabi, D. Energetica, P. Torino, T. Italia, S. P. Corgnati, D. Energetica, P. Torino, D. Energetica, and P. Torino, "Effect of Occupant Behaviour Related Influencing Factors on Final Energy End Uses in Buildings," in *Proceedings of Climamed*, 2011.
- [8] L. Barnard, J. S. Yi, J. A. Jacko, and A. Sears, "Capturing the effects of context on human performance in mobile computing systems," *Personal and Ubiquitous Computing*, vol. 11, no. 2, pp. 81–96, 2007, doi: 10.1007/s00779-006-0063-x.
- [9] G. Branco, B. Lachal, P. Gallinelli, and W. Weber, "Predicted versus observed heat consumption of a low energy multifamily complex in Switzerland based on long-term experimental data," *Energy and Buildings*, vol. 36, no. 6, pp. 543–555, 2004, doi: 10.1016/j.enbuild.2004.01.028.
- [10] J. F. Nicol, "Characterising occupant behavior in buildings: Towards a stochastic model of occupant use of windows, lights, blinds heaters and fans," in *the International Building Performance Simulation Association*, 2001, pp. 1073–1078, [Online]. Available: [http://www.ibpsa.org/proceedings/BS2001/BS01\\_1073\\_1078.pdf](http://www.ibpsa.org/proceedings/BS2001/BS01_1073_1078.pdf).
- [11] H. B. Gunay, W. O'Brien, I. Beausoleil-Morrison, and S. Gilani, "Development and

- implementation of an adaptive lighting and blinds control algorithm,” *Building and Environment*, vol. 113, pp. 185–199, 2017, doi: 10.1016/j.buildenv.2016.08.027.
- [12] T. Hong and H.-W. Lin, “Occupant Behavior: Impact on Energy Use of Private Offices,” *Asim IBSPA Asia Conference*, no. January, p. 12, 2013.
- [13] I. Gaetani, P.-J. Hoes, and J. L. M. Hensen, “Occupant behavior in building energy simulation: Towards a fit-for-purpose modeling strategy,” *Energy and Buildings*, vol. 121, pp. 188–204, 2016, doi: <http://dx.doi.org/10.1016/j.enbuild.2016.03.038>.
- [14] C. M. Clevenger and J. Haymaker, “The impact of the building occupant on energy modeling simulations,” in *Joint International Conference on Computing and Decision Making in Civil and Building Engineering*, 2006, pp. 1–10, doi: 10.1061/(ASCE)CP.1943-5487.0000323.
- [15] O. G. Santin, L. Itard, and H. Visscher, “The effect of occupancy and building characteristics on energy use for space and water heating in Dutch residential stock,” *Energy and Buildings*, vol. 41, no. 11, pp. 1223–1232, 2009, doi: 10.1016/j.enbuild.2009.07.002.
- [16] A. Kavousian, R. Rajagopal, and M. Fischer, “Determinants of residential electricity consumption: Using smart meter data to examine the effect of climate, building characteristics, appliance stock, and occupants’ behavior,” *Energy*, vol. 55, pp. 184–194, 2013, doi: 10.1016/j.energy.2013.03.086.
- [17] S. D’Oca, V. Fabi, S. P. Corngati, and R. K. Andersen, “Effect of thermostat and window opening occupant behavior models on energy use in homes,” *Building Simulation*, vol. 7, no. 6, pp. 683–694, 2014, doi: 10.1007/s12273-014-0191-6.
- [18] T. Hong, S. D’Oca, W. J. N. Turner, and S. C. Taylor-lange, “An ontology to represent energy-related occupant behavior in buildings. Part I: Introduction to the DNAs framework,” *Building and Environment*, vol. 92, pp. 764–777, 2015, doi: 10.1016/j.buildenv.2015.02.019.
- [19] V. Fabi, R. V. Andersen, S. Corngati, and B. W. Olesen, “Occupants’ window opening behaviour: A literature review of factors influencing,” *Building and Environment*, vol. 58, pp. 188–198, 2012, doi: <http://dx.doi.org/10.1016/j.buildenv.2012.07.009>.
- [20] J. F. Nicol, M. A. Humphreys, and B. Olesen, “A stochastic approach to thermal comfort - Occupant behavior and energy use in buildings,” *ASHRAE Transactions*, vol. 110, no. 2, pp. 554–568, 2004, [Online]. Available: [https://www.researchgate.net/publication/279888155\\_A\\_stochastic\\_approach\\_to\\_thermal\\_comfort\\_-\\_Occupant\\_behavior\\_and\\_energy\\_use\\_in\\_buildings](https://www.researchgate.net/publication/279888155_A_stochastic_approach_to_thermal_comfort_-_Occupant_behavior_and_energy_use_in_buildings).
- [21] T. Hong, S. D’Oca, S. C. Taylor-Lange, W. J. N. Turner, Y. Chen, and S. P. Corngati, “An ontology to represent energy-related occupant behavior in buildings. Part II: Implementation of the DNAs framework using an XML schema,” *Building and Environment*, vol. 94, pp. 196–205, 2015, doi: 10.1016/j.buildenv.2015.08.006.

- [22] S. Attia, A. Evrard, and E. Gratia, "Development of benchmark models for the Egyptian residential buildings sector," *Applied Energy*, vol. 94, no. 2012, pp. 270–284, 2012, doi: 10.1016/j.apenergy.2012.01.065.
- [23] X. Feng, D. Yan, C. Wang, and H. Sun, "A preliminary research on the derivation of typical occupant behavior based on large-scale questionnaire surveys," *Energy and Buildings*, vol. 117, pp. 332–340, 2016, doi: 10.1016/j.enbuild.2015.09.055.
- [24] P. R. Boyce, "Observations of the manual switching of lighting," *Lighting Research & Technology*, vol. 12, no. 4, pp. 195–205, 1980, doi: 10.1177/096032718001200403.
- [25] H. B. Rijal, P. Tuohy, M. A. Humphreys, J. F. Nicol, A. Samuel, and J. Clarke, "Using results from field surveys to predict the effect of open windows on thermal comfort and energy use in buildings," *Energy and Buildings*, vol. 39, no. 7, pp. 823–836, 2007, doi: 10.1016/j.enbuild.2007.02.003.
- [26] R. Andersen, B. Olesen, and J. Toftum, "Modelling occupants' heating set-point preferences," *12th Conference of International Building Performance Simulation Association, Sydney, 14-16 November*, pp. 151–156, 2011, [Online]. Available: [http://orbit.dtu.dk/fedora/objects/orbit:72610/datastreams/file\\_6369881/content](http://orbit.dtu.dk/fedora/objects/orbit:72610/datastreams/file_6369881/content).
- [27] Amir Seyed, Karava Panagiota, Konstantzos, and Iason, "Occupant interactions with shading and lighting systems using different control interfaces : A pilot field study," *Building and Environment*, vol. 97, pp. 177–195, 2016, doi: 10.1016/j.buildenv.2015.12.008.
- [28] H. B. Gunay, W. O'Brien, I. Beausoleil-Morrison, and S. Gilani, "Modeling plug-in equipment load patterns in private office spaces," *Energy and Buildings*, vol. 121, pp. 234–249, 2016, doi: 10.1016/j.enbuild.2016.03.001.
- [29] J. Wienold and J. Christoffersen, "Evaluation methods and development of a new glare prediction model for daylight environments with the use of CCD cameras," *Energy and Buildings*, vol. 38, no. 7, pp. 743–757, 2006, doi: 10.1016/j.enbuild.2006.03.017.
- [30] M. Schweiker and A. Wagner, "The effect of occupancy on perceived control, neutral temperature, and behavioral patterns," *Energy and Buildings*, vol. 117, pp. 246–259, 2016, doi: 10.1016/j.enbuild.2015.10.051.
- [31] T. Ming, D. Chen, S. N. Toudeshki, S. Talele, G. T. Checketts, N. Hasib, C. Wicaksono, G. Xiong, Y. Qiu, C. Peng, J. Mun, R. Rayegan, and Y. Tao, "A Zero Energy Lab as a validation testbed: Concept, features, and performance," *International Journal of Hydrogen Energy*, vol. 40, no. 37, pp. 12854–12867, 2015, doi: 10.1016/j.ijhydene.2015.07.010.
- [32] J. Whyte, N. Bouchlaghem, A. Thorpe, and R. McCaffer, "From CAD to virtual reality: Modelling approaches, data exchange and interactive 3D building design tools," *Automation in construction*, vol. 10, no. 1, pp. 43–45, 2000, doi: 10.1016/S0926-5805(99)00012-6.

- [33] K. D. Maldovan, J. I. Messner, P. Mera Faddoul, and M. Faddoul, "Framework for reviewing mockups in an immersive environment," in *CONVR. 2006*, 2006, vol. 2006, no. March, p. 6th.
- [34] P. S. Dunston, L. L. Arns, and J. D. McGlothlin, "an Immersive Virtual Reality Mock-Up for Design Review," in *7th International Conference on Construction Applications of Virtual Reality*, 2007.
- [35] A. Heydarian, J. P. Carneiro, D. Gerber, and B. Becerik-Gerber, "Immersive virtual environments, understanding the impact of design features and occupant choice upon lighting for building performance," *Building and Environment*, vol. 89, pp. 217–228, 2015, doi: 10.1016/j.buildenv.2015.02.038.
- [36] S. Saeidi, T. Rizzuto, Y. Zhu, and R. Kooima, "Measuring the effectiveness of an immersive virtual environment for the modeling and prediction of occupant behavior," in *Sustainable Human–Building Ecosystems*, 2015, pp. 159–167, doi: 10.1061/9780784479681.017.
- [37] S. Niu, W. Pan, and Y. Zhao, "A virtual reality integrated design approach to improving occupancy information integrity for closing the building energy performance gap," *Sustainable Cities and Society*, vol. 27, pp. 275–286, 2015, doi: 10.1016/j.scs.2016.03.010.
- [38] S. Saeidi, C. Chokwitthaya, Y. Zhu, and M. Sun, "Spatial-temporal event-driven modeling for occupant behavior studies using immersive virtual environments," *Automation in Construction*, vol. 94, pp. 371–382, 2018, doi: 10.1016/j.autcon.2018.07.019.
- [39] C. M. Bishop, *Neural networks for pattern recognition*. Oxford university press, 1995.
- [40] G. Zemella, D. De March, M. Borrotti, and I. Poli, "Optimised design of energy efficient building façades via Evolutionary Neural Networks," *Energy and Buildings*, vol. 43, no. 12, pp. 3297–3302, 2011, doi: 10.1016/j.enbuild.2011.10.006.
- [41] R. W. Da Fonseca, E. L. Didoné, and F. O. R. Pereira, "Using artificial neural networks to predict the impact of daylighting on building final electric energy requirements," *Energy and Buildings*, vol. 61, pp. 31–38, 2013, doi: 10.1016/j.enbuild.2013.02.009.
- [42] M. Aydinalp, V. I. Ugursal, and A. S. Fung, "Modeling of the space and domestic hot-water heating energy-consumption in the residential sector using neural networks," *Applied Energy*, vol. 79, no. 2, pp. 159–178, 2004, doi: 10.1016/j.apenergy.2003.12.006.
- [43] G. Escrivá-Escrivá, C. Álvarez-Bel, C. Roldán-Blay, and M. Alcázar-Ortega, "New artificial neural network prediction method for electrical consumption forecasting based on building end-uses," *Energy and Buildings*, vol. 43, no. 11, pp. 3112–3119, 2011, doi: 10.1016/j.enbuild.2011.08.008.
- [44] S. S. K. Kwok, "A study of the importance of occupancy to building cooling load in prediction by intelligent approach," *Energy Conversion and Management*, vol. 52, no. 7,

- pp. 2555–2564, 2011, doi: 10.1016/j.enconman.2011.02.002.
- [45] M. C. Leung, N. C. F. Tse, L. L. Lai, and T. T. Chow, “The use of occupancy space electrical power demand in building cooling load prediction,” *Energy and Buildings*, vol. 55, pp. 151–163, 2012, doi: 10.1016/j.enbuild.2012.08.032.
  - [46] R. Jing, M. Wang, R. Zhang, N. Li, and Y. Zhao, “A study on energy performance of 30 commercial office buildings in Hong Kong,” *Energy and Buildings*, vol. 144, pp. 117–128, 2017, doi: 10.1016/j.enbuild.2017.03.042.
  - [47] M. W. Ahmad, M. Mourshed, and Y. Rezgui, “Trees vs Neurons: Comparison between random forest and ANN for high-resolution prediction of building energy consumption,” *Energy and Buildings*, vol. 147, pp. 77–89, 2017, doi: 10.1016/j.enbuild.2017.04.038.
  - [48] H. D. Mehr, H. Polat, and A. Cetin, “Resident activity recognition in smart homes by using artificial neural networks,” *2016 4th International Istanbul Smart Grid Congress and Fair (ICSG)*, no. April, pp. 1–5, 2016, doi: 10.1109/SGCF.2016.7492428.
  - [49] Y. Lecun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553, pp. 436–444, 2015, doi: 10.1038/nature14539.
  - [50] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, “Generative adversarial nets,” in *Advances in Neural Information Processing Systems*, 2014, pp. 2672–2680, [Online]. Available: <http://papers.nips.cc/paper/5423-generative-adversarial-nets> (Accessed date: July 14, 2019).
  - [51] E. Collier, R. Dibiano, and S. Mukhopadhyay, “CactusNets: Layer applicability as a metric for transfer learning,” in *International Joint Conference on Neural Networks*, 2018, pp. 1–8, doi: 10.1109/IJCNN.2018.8489649.
  - [52] S. Basu, S. Mukhopadhyay, M. Karki, R. DiBiano, S. Ganguly, R. Nemani, and S. Gayaka, “Deep neural networks for texture classification—A theoretical analysis,” *Neural Networks*, vol. 97, pp. 173–182, 2018, doi: 10.1016/j.neunet.2017.10.001.
  - [53] S. Basu, S. Ganguly, S. Mukhopadhyay, R. DiBiano, M. Karki, and R. Nemani, “DeepSat - A learning framework for satellite imagery,” *GIS: Proceedings of the ACM International Symposium on Advances in Geographic Information Systems*, vol. 03-06-Nov, pp. 1–22, 2015, doi: 10.1145/2820783.2820816.
  - [54] E. Collier, K. Duffy, S. Ganguly, G. Madanguit, S. Kalia, G. Shreekanth, R. Nemani, A. Michaelis, S. Li, A. Ganguly, and S. Mukhopadhyay, “Progressively growing generative adversarial networks for high resolution semantic segmentation of satellite images,” in *International Conference on Data Mining Workshops*, 2018, pp. 763–769, doi: 10.1109/ICDMW.2018.00115.
  - [55] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi, “Photo-Realistic Single Image Super-

- Resolution Using a Generative Adversarial Network,” *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, pp. 4681–4690, 2017, doi: 10.1109/CVPR.2017.19.
- [56] A. Radford, M. Luke, and C. Soumith, “Unsupervised representation learning with deep convolutional generative adversarial networks.,” in *arXiv preprint arXiv:1511.06434*, 2015, pp. 1–15, doi: 10.1007/s11280-018-0565-2.
  - [57] X. Wang and A. Gupta, “Generative Image Modeling using Style and Structure Adversarial Networks,” in *European Conference on Computer Vision*, 2016, vol. 9908, pp. 318–335, doi: 10.1007/978-3-319-46493-0.
  - [58] R. J. Moffat, “Describing the uncertainties in experimental results,” *Experimental Thermal and Fluid Science*, vol. 1, no. 1, pp. 3–17, 1988, doi: 10.1016/0894-1777(88)90043-X.
  - [59] C. Chokwitthaya, Y. Zhu, R. Dibiano, and S. Mukhopadhyay, “Combining context-aware design-specific data and building performance models to improve building performance predictions during design,” *Automation in construction*, vol. 107, p. 102917, 2019, doi: 10.1016/j.autcon.2019.102917.
  - [60] G. Y. Yun, H. Kim, and J. T. Kim, “Effects of occupancy and lighting use patterns on lighting energy consumption,” *Energy and Buildings*, vol. 46, pp. 152–158, 2012, doi: 10.1016/j.enbuild.2011.10.034.
  - [61] A. Heydarian, E. Pantazis, D. Gerber, and B. Becerik-Gerber, “Defining Lighting Settings to Accommodate End-User Preferences while Reducing Energy Consumption in Buildings,” *Construction Research Congress 2016: Old and New Construction Technologies Converge in Historic San Juan - Proceedings of the 2016 Construction Research Congress, CRC 2016*, no. August, 2016, doi: 10.1061/9780784479827.113.
  - [62] L. Jia, S. Afshari, S. Mishra, and R. J. Radke, “Simulation for pre-visualizing and tuning lighting controller behavior,” *Energy and Buildings*, vol. 70, pp. 287–302, 2014, doi: 10.1016/j.enbuild.2013.11.063.
  - [63] A. Löffler, L. Marsalek, H. Hoffmann, and P. Slusallek, “Realistic Lighting Simulation for Interactive VR Applications,” *Proceedings of the 17th Eurographics Conference on Virtual Environments & Third Joint Virtual Reality*, no. 2011, pp. 1–8, 2011, doi: 10.2312/EGVE/JVRC11/001-008.
  - [64] R. K. Jain, K. M. Smith, P. J. Culligan, and J. E. Taylor, “Forecasting energy consumption of multi-family residential buildings using support vector regression: Investigating the impact of temporal and spatial monitoring granularity on performance accuracy,” *Applied Energy*, vol. 123, pp. 168–178, 2014, doi: 10.1016/j.apenergy.2014.02.057.
  - [65] I. Dincer, “Environmental impacts of energy,” *Energy policy*, vol. 27, no. 14, pp. 845–854, 1999, doi: 10.1016/S0140-6701(00)92845-5.
  - [66] V. Granadeiro, J. P. Duarte, J. R. Correia, and V. M. S. Leal, “Building envelope shape

- design in early stages of the design process: Integrating architectural design systems and energy simulation,” *Automation in Construction*, vol. 32, no. September, pp. 196–209, 2013, doi: 10.1016/j.autcon.2012.12.003.
- [67] W. Wang, R. Zmeureanu, and H. Rivard, “Applying multi-objective genetic algorithms in green building design optimization,” *Building and Environment*, vol. 40, no. 11, pp. 1512–1525, 2005, doi: 10.1016/j.buildenv.2004.11.017.
  - [68] P. Hoes, J. L. M. Hensen, M. Loomans, B. De Vries, and D. Bourgeois, “User behavior in whole building simulation,” *Energy and buildings*, vol. 41, no. 3, pp. 295–302, 2009.
  - [69] H. T. Pao, “Comparing linear and nonlinear forecasts for Taiwan’s electricity consumption,” *Energy*, vol. 31, no. 12, pp. 1793–1805, 2006, doi: 10.1016/j.energy.2005.08.010.
  - [70] B. Blocken and C. Gualtieri, “Ten iterative steps for model development and evaluation applied to Computational Fluid Dynamics for Environmental Fluid Mechanics,” *Environmental Modelling and Software*, vol. 33, no. 0, pp. 1–22, 2012, doi: 10.1016/j.envsoft.2012.02.001.
  - [71] J. A. Love, “Manual switching patterns in private offices,” *Lighting Research & Technology*, vol. 30, no. 1, pp. 45–50, 1998, doi: 10.1177/096032719803000107.
  - [72] a. I. Rubin, B. L. Collins, and R. L. Tibbott, “Window blinds as a potential energy saver-a case study,” *NBS Building Science Series*, vol. 112, p. 89, 1978, [Online]. Available: <https://www.ncjrs.gov/pdffiles1/Digitization/64368NCJRS.pdf>.
  - [73] D. Lindelöf and N. Morel, “A field investigation of the intermediate light switching by users,” *Energy and Buildings*, vol. 38, no. 7, pp. 790–801, 2006, doi: 10.1016/j.enbuild.2006.03.003.
  - [74] A. C. Kunduraci and T. Kazanasmaz, “Assessing Manual Lighting Control in Offices.”
  - [75] A. Mahdavi, A. Mohammadi, E. Kabir, and L. Lambeva, “Occupants’ operation of lighting and shading systems in office buildings,” *Journal of Building Performance Simulation*, vol. 1, no. 1, pp. 57–65, 2008.
  - [76] M. Kinateder, M. Müller, M. Jost, A. Mühlberger, and P. Pauli, “Social influence in a virtual tunnel fire - Influence of conflicting information on evacuation behavior,” *Applied Ergonomics*, vol. 45, no. 6, pp. 1649–1659, 2014, doi: 10.1016/j.apergo.2014.05.014.
  - [77] M. Kobes, I. Helsloot, B. De Vries, and J. Post, “Exit choice,(pre-) movement time and (pre-) evacuation behaviour in hotel fire evacuation - Behavioural analysis and validation of the use of serious gaming in experimental research,” *Procedia Engineering*, vol. 3, pp. 37–51, 2010, doi: 10.1016/j.proeng.2010.07.006.
  - [78] C. Mackie, J. Cowden, D. Bowman, and T. Walid, “Desktop and immersive tools for residential home design,” in *Construction Applications of Virtual Reality*, 2004, pp. 63–

- [79] D. Tutt and C. Harty, "Journeys through the CAVE: The use of 3D immersive environments for client engagement practices in hospital design," in *Annual Association of Researchers in Construction Management Conference*, 2013, pp. 111–121, [Online]. Available: <https://pdfs.semanticscholar.org/2cd3/603e77cafec17b1c2bfa77abe67a442fd606.pdf>.
- [80] Y. Hong and P. Michalatos, "LumiSpace: A VR architectural daylighting design system," in *SIGGRAPH ASIA 2016 Virtual Reality Meets Physical Reality: Modelling and Simulating Virtual Humans and Environments*, 2016, pp. 1–2, doi: 10.1145/2992138.2992140.
- [81] J. H. Kwon, J. Powell, and A. Chalmers, "How level of realism influences anxiety in virtual reality environments for a job interview," *International Journal of Human Computer Studies*, vol. 71, no. 10, pp. 978–987, 2013, doi: 10.1016/j.ijhcs.2013.07.003.
- [82] C. Chokwitthaya, S. Saeidi, Y. Zhu, and R. Kooima, "The Impact of Lighting Simulation Discrepancies on Human Visual Perception and Energy Behavior Simulations in Immersive Virtual Environment," in *Congress on Computing in Civil Engineering, Proceedings*, 2017, vol. 2017-June.
- [83] Y. Zhu, S. Saeidi, T. Rizzuto, A. Roetzel, and R. Kooima, "Potential and challenges of immersive virtual environments for occupant energy behavior modeling and validation: A literature review," *Journal of Building Engineering*, vol. 19, 2018, doi: 10.1016/j.jobbe.2018.05.017.
- [84] P. S. Bordnick, K. M. Graap, H. L. Copp, J. Brooks, and M. Ferrer, "Virtual reality cue reactivity assessment in cigarette smokers," *CyberPsychology and Behavior*, vol. 8, no. 5, pp. 487–492, 2005, doi: 10.1089/cpb.2005.8.487.
- [85] D. A. Das, K. A. Grimmer, A. L. Sparnon, S. E. McRae, and B. H. Thomas, "The efficacy of playing a virtual reality game in modulating pain for children with acute burn injuries: A randomized controlled trial," *BMC Pediatrics*, vol. 5, no. 1, pp. 1–10, 2005, doi: 10.1186/1471-2431-5-1.
- [86] I. El Farissi, M. Azizi, J.-L. Lanet, and M. Moussaoui, "Neural Network Vs. Bayesian Network to Detect Java Card Mutants," *AASRI Procedia*, vol. 4, pp. 132–137, 2013, doi: 10.1016/j.aasri.2013.10.021.
- [87] A. C. Comrie, "Comparing Neural Networks and Regression Models for Ozone Forecasting," *Journal of the Air and Waste Management Association*, vol. 47, no. 6, pp. 653–663, 1997, doi: 10.1080/10473289.1997.10463925.
- [88] M. S. El-Abbasy, A. Senouci, T. Zayed, F. Mirahadi, and L. Parvizsedghy, "Artificial neural network models for predicting condition of offshore oil and gas pipelines," *Automation in Construction*, vol. 45, pp. 50–65, 2014, doi: 10.1016/j.autcon.2014.05.003.



- [89] A. Taghi Nezhad Bilandi, "COSTS AND BENEFITS OF FLOOD MITIGATION IN LOUISIANA," Louisiana State University, 2019.
- [90] M. R. Mosavi, "Comparing DGPS corrections prediction using neural network, fuzzy neural network, and Kalman filter," *GPS Solutions*, vol. 10, no. 2, pp. 97–107, 2006, doi: 10.1007/s10291-005-0011-7.
- [91] H. Hruschka and M. Natter, "Comparing performance of feedforward neural nets and K-means for cluster-based market segmentation," *European Journal of Operational Research*, vol. 114, no. 2, pp. 346–353, 1999, doi: 10.1016/S0377-2217(98)00170-2.
- [92] K. Suzuki, H. Yoshida, J. Näppi, S. G. Armato, and A. H. Dachman, "Mixture of expert 3D massive-training ANNs for reduction of multiple types of false positives in CAD for detection of polyps in CT colonography," *Medical Physics*, vol. 35, no. 2, pp. 694–703, 2008, doi: 10.1118/1.2829870.
- [93] D. B. Sam, S. Surya, and R. V. Babu, "Switching convolutional neural network for crowd counting," *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, vol. 2017-Janua, pp. 4031–4039, 2017, doi: 10.1109/CVPR.2017.429.
- [94] Y. Sun, X. Wang, and X. Tang, "Deep Learning Face Representation From Predicting 10 000 Classes," *Cvpr*, pp. 1891–1898, 2014, doi: 10.1109/CVPR.2014.244.
- [95] P. Bermejo, J. A. Gámez, and J. M. Puerta, "A GRASP algorithm for fast hybrid (filter-wrapper) feature subset selection in high-dimensional datasets," *Pattern Recognition Letters*, vol. 32, no. 5, pp. 701–711, 2011, doi: 10.1016/j.patrec.2010.12.016.
- [96] Y. Saeys, T. Abeel, and Y. Van de Peer, "Robust Feature Selection Using Ensemble Feature Selection Techniques," in *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, 2008, pp. 313–325.
- [97] A. Jović, K. Brkić, and N. Bogunović, "A review of feature selection methods with applications," *2015 38th International Convention on Information and Communication Technology, Electronics and Microelectronics, MIPRO 2015 - Proceedings*, pp. 1200–1205, 2015, doi: 10.1109/MIPRO.2015.7160458.
- [98] P. C. Da Silva, V. Leal, and M. Andersen, "Occupants interaction with electric lighting and shading systems in real single-occupied offices: Results from a monitoring campaign," *Building and Environment*, vol. 64, pp. 152–168, 2013, doi: 10.1016/j.buildenv.2013.03.015.
- [99] D. Bourgeois, C. Reinhart, and I. Macdonald, "Adding advanced behavioural models in whole building energy simulation: A study on the total energy impact of manual and automated lighting control," *Energy and Buildings*, vol. 38, no. 7, pp. 814–823, 2006, doi: 10.1016/j.enbuild.2006.03.002.
- [100] D. Yan, X. Feng, Y. Jin, and C. Wang, "The evaluation of stochastic occupant behavior

- models from an application-oriented perspective: Using the lighting behavior model as a case study,” *Energy and Buildings*, vol. 176, pp. 151–162, 2018, doi: 10.1016/j.enbuild.2018.07.037.
- [101] “U.S. General Services Administration.” <https://www.gsa.gov/node/82715> (accessed Jul. 06, 2019).
- [102] L. E. Baum, “An Inequality and Associated Maximization Thechnique in Statistical Estimation for Probabilistic Functions of Markov Process,” *Inequalities*, vol. 3, pp. 1–8, 1972.
- [103] J. Picone, “Continuous speech recognition using hidden Markov models,” *ASSP Magazine, IEEE*, vol. 7, no. July 1990, p. 41, 1990, doi: 10.1109/53.54527.
- [104] S. Luhr, S. Venkatesh, G. West, and H. H. Bui, “Duration Abnormality Detection in Sequences of Human Activity,” *PRICAI*, no. February, pp. 983–984, 2004.
- [105] Z. Ghahramani, “An Introduction To Hidden Markov Models and Bayesian Networks,” *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 15, no. 01, pp. 9–42, 2001, doi: 10.1142/S0218001401000836.
- [106] I. Visser, “Seven things to remember about hidden Markov models: A tutorial on Markovian models for time series,” *Journal of Mathematical Psychology*, vol. 55, no. 6, pp. 403–415, 2011, doi: 10.1016/j.jmp.2011.08.002.
- [107] C. Chokwitthaya, Y. Zhu, R. Dibiano, and S. Mukhopadhyay, “A Machine Learning Algorithm to Improve Building Performance during Design,” *Manuscript submitted for publication*.
- [108] Zou, Hui, and H. Trevor, “Regularization and variable selection via the elastic net,” *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 67, no. 2, pp. 301–320, 2005, doi: 10.1111/j.1467-9868.2005.00503.x.
- [109] S. Kalogirou, “Artificial neural networks for the prediction of the energy consumption of a passive solar building,” *Energy*, vol. 25, no. 5, pp. 479–491, 2000, doi: 10.1016/S0360-5442(99)00086-9.
- [110] Y. Gu, “The Impacts of Real-time Knowledge Based Personal Lighting Control on Energy Consumption, User Satisfaction and Task Performance in Offices,” p. 157, 2011.
- [111] J. Yang, M. Santamouris, and S. E. Lee, “Review of occupancy sensing systems and occupancy modeling methodologies for the application in institutional buildings,” *Energy and Buildings*, vol. 121, pp. 344–349, 2016, doi: 10.1016/j.enbuild.2015.12.019.
- [112] C. Demanuele, T. Tweddell, and M. Davies, “Bridging the gap between predicted and actual energy performance in schools,” *World Renewable Energy Congress XI*, no. September, pp. 1–6, 2010.

- [113] A. C. Menezes, A. Cripps, D. Bouchlaghem, and R. Buswell, "Predicted vs. actual energy performance of non-domestic buildings: Using post-occupancy evaluation data to reduce the performance gap," *Applied Energy*, vol. 97, pp. 355–364, 2012, doi: 10.1016/j.apenergy.2011.11.075.
- [114] C. Chokwitthaya, E. Collier, Y. Zhu, and S. Mukhopadhyay, "Improving prediction accuracy in building performance models using generative adversarial networks (GANs)," in *International Joint Conference on Neural Networks*, 2019, pp. 1–9, doi: 10.1109/IJCNN.2019.8852411.
- [115] C. van Dronkelaar, M. Dowson, C. Spataru, and D. Mumovic, "A review of the energy performance gap and its underlying causes in non-domestic buildings," *Frontiers in Mechanical Engineering*, vol. 1, pp. 1–14, 2016, doi: 10.3389/fmech.2015.00017.
- [116] B. Frei, C. Sagerschnig, and D. Gyalistras, "Performance gaps in Swiss buildings : an analysis of conflicting objectives and mitigation strategies," *Energy Procedia*, vol. 122, pp. 421–426, 2017, doi: 10.1016/j.egypro.2017.07.425.
- [117] V. Fabi, S. Corgnati, Rune Vinther Andersen, B. W. Olesen, R. V. Andersen, S. Corgnati, and B. W. Olesen, "Occupants' window opening behaviour: A literature review of factors influencing occupant behaviour and models," *Building and Environment*, vol. 58, pp. 188–198, 2012, doi: 10.1016/j.buildenv.2012.07.009.
- [118] J. F. Sowa, "Syntax, semantics, and pragmatics of contexts," in *International Conference on Conceptual Structures*, 1995, vol. 954, pp. 1–15, doi: 10.1007/3-540-60161-9\_25.
- [119] A. Lijoi, R. H. Mena, and I. Prünster, "Controlling the reinforcement in Bayesian non-parametric mixture models," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 69, no. 4, pp. 715–740, 2007, doi: 10.1111/j.1467-9868.2007.00609.x.
- [120] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Journal of basic Engineering*, vol. 82, no. 1, pp. 35–45, 1960, doi: 10.1115/1.3662552.
- [121] F. Tahmasebi and A. Mahdavi, "An inquiry into the reliability of window operation models in building performance simulation," *Building and Environment*, vol. 105, pp. 343–357, 2016, doi: 10.1016/j.buildenv.2016.06.013.
- [122] M. Keller, J. O'Donnell, and K. Menzel, "Integrating the specification, acquisition and processing of building performance information," *Tsinghua Science and Technology*, vol. 13, no. 1, pp. 1–6, 2008, doi: 10.1.1.457.374.
- [123] C. S. Cho, D. Chen, and S. Woo, "Building information modeling (BIM) - based design of energy efficient buildings," in *International Symposium on Automation and Robotics in Construction*, 2011, pp. 1079–1084, doi: 10.22260/isarc2011/0198.
- [124] F. W. H. Wong, P. T. I. Lam, and E. H. W. Chan, "Optimising design objectives using the Balanced Scorecard approach," *Design Studies*, vol. 30, no. 4, pp. 369–392, 2009, doi: 10.1016/j.destud.2008.10.004.

- [125] S. Basu, S. Ganguly, S. Mukhopadhyay, R. DiBiano, M. Karki, and R. Nemani, “DeepSat - A learning framework for satellite imagery,” in *International Conference on Advances in Geographic Information Systems*, 2015, pp. 1–10, doi: 10.1145/2820783.2820816.
- [126] Q. Liu, S. Mukhopadhyay, M. X. B. Rodriguez, X. Fu, S. Sahu, D. Burk, and M. Gartia, “A one-shot learning framework for assessment of fibrillar collagen from second harmonic generation images of an infarcted myocardium,” in *International Symposium on Biomedical Imaging*, 2020, pp. 839–843, doi: 10.1109/ISBI45749.2020.9098444.
- [127] Q. Liu, E. Collier, and S. Mukhopadhyay, “PCGAN-CHAR: Progressively trained classifier generative adversarial networks for classification of noisy handwritten bangla characters,” in *International Conference on Asian Digital Libraries*, 2019, pp. 3–15, doi: 10.1007/978-3-030-34058-2\_1.
- [128] J. T. Springenberg, “Unsupervised and semi-supervised learning with categorical generative adversarial networks,” in *International Conference on Learning Representations*, 2016, pp. 1–20, [Online]. Available: <http://arxiv.org/abs/1511.06390> (Accessed date: August 12, 2019).
- [129] M. Mirza and S. Osindero, “Conditional generative adversarial nets,” *arXiv preprint arXiv:1411.1784*, pp. 1–7, 2014, [Online]. Available: <http://arxiv.org/abs/1411.1784> (Accessed date: September 23, 2019).
- [130] J. Voss, “Revisiting office space standards.” <https://www.brigholme.com/files/resources/Revisiting-office-space-standards-white-paper.pdf> (accessed Oct. 03, 2019).
- [131] D. P. Kroese, T. Brereton, T. Taimre, and Z. I. Botev, “Why the Monte Carlo method is so important today,” *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 6, no. 6, pp. 386–392, 2014, doi: 10.1002/wics.1314.
- [132] H. Ishwaran and G. Takahara, “Independent and identically distributed Monte Carlo algorithms for semiparametric linear mixed models,” *Journal of the American Statistical Association*, vol. 97, no. 460, pp. 1154–1166, 2002, doi: 10.1198/016214502388618951.
- [133] M. N. Almarshad, M. Barkat, and S. A. Alshebeili, “A Monte Carlo simulation for two novel automatic censoring techniques of radar interfering targets in log-normal clutter,” *Signal Processing*, vol. 88, no. 3, pp. 719–732, 2008, doi: 10.1016/j.sigpro.2007.09.013.
- [134] T. Homem-De-Mello, “On rates of convergence for stochastic optimization problems under non-independent and identically distributed sampling,” *SIAM Journal on Optimization*, vol. 19, no. 2, pp. 524–551, 2008, doi: 10.1137/060657418.
- [135] P. Cunningham, J. Carney, and S. Jacob, “Stability problems with artificial neural networks and the ensemble solution,” *Artificial Intelligence in Medicine*, vol. 20, no. 3, pp. 217–225, 2000, doi: 10.1016/S0933-3657(00)00065-8.
- [136] C. Chokwitthaya, Y. Zhu, R. Dibiano, and M. Supratik, “A machine learning algorithm to

- improve building performance modeling during design,” *MethodsX*, vol. 7, pp. 35–49, 2020, doi: 10.1016/j.mex.2019.10.037.
- [137] C. Reinhart and P. F. Breton, “Experimental validation of 3ds Max design 2009 and Daysim 3.0,” in *International Building Performance Simulation Association*, 2009, pp. 1514–1521, doi: 10.1.1.172.3251.
- [138] P. Boyce, “Lighting research for interiors: the beginning of the end or the end of the beginning,” *Lighting Research and Technology*, vol. 36, no. 4, pp. 283–293, 2004, doi: 10.1191/11477153504li118oa.
- [139] P. R. Boyce, N. H. Eklund, and S. N. Simpson, “Individual lighting control: Task performance, mood, and illuminance,” *Journal of the Illuminating Engineering Society*, vol. 29, no. 1, pp. 131–142, 2000, doi: 10.1080/00994480.2000.10748488.
- [140] J. A. Veitch and G. R. Newsham, “Lighting quality and energy-efficiency effects on task performance, mood, health, satisfaction and comfort,” *Journal of the Illuminating Engineering Society*, vol. 27, no. 1, pp. 107–129, 1998, doi: 10.1080/00994480.1998.10748216.
- [141] D. R. G. Hunt, “The use of artificial lighting in relation to daylight levels and occupancy,” *Building and Environment*, vol. 14, no. 1, pp. 21–33, 1979, doi: 10.1016/0360-1323(79)90025-8.
- [142] R. Likert, “A technique for the measurement of attitudes,” *Archives of psychology*, vol. 22, no. 140, pp. 1–55, 1932, [Online]. Available: [https://legacy.voteview.com/pdf/Likert\\_1932.pdf](https://legacy.voteview.com/pdf/Likert_1932.pdf) (Accessed date: August 17, 2019).
- [143] C. M. Bishop, *Pattern recognition and machine learning*. Springer, 2006.
- [144] C. Guestrin, “Unsupervised learning or clustering — K - means Gaussian mixture models.” <http://www.cs.cmu.edu/~guestrin/Class/10701-S07/Slides/clustering.pdf> (accessed Nov. 04, 2019).
- [145] R. B. Lanjewar, S. Mathurkar, and N. Patel, “Implementation and comparison of speech emotion recognition system using Gaussian mixture model (GMM) and K-nearest neighbor (K-NN) techniques,” *Procedia Computer Science*, vol. 49, no. 1, pp. 50–57, 2015, doi: 10.1016/j.procs.2015.04.226.
- [146] G. S. Morrison, “A comparison of procedures for the calculation of forensic likelihood ratios from acoustic-phonetic data: Multivariate kernel density (MVKD) versus Gaussian mixture model-universal background model (GMM-UBM),” *Speech Communication*, vol. 53, no. 2, pp. 242–256, 2011, doi: 10.1016/j.specom.2010.09.005.
- [147] B. Awwad, S. Hasan, and J. Q. Gan, “Machine learning and data mining in pattern recognition,” in *International Workshop on Machine Learning and Data Mining in Pattern Recognition*, 2005, vol. 3587, pp. 96–106, doi: 10.1007/b138149.

- [148] Q. Liu, S. Mukhopadhyay, Y. Zhu, and R. Gudishala, "Improving route choice models by incorporating contextual factors via knowledge distillation," in *International Joint Conference on Neural Networks*, 2019, pp. 1–8, doi: 10.1109/IJCNN.2019.8852482.
- [149] C. Chokwitthaya, Y. Zhu, and A. Jafari, "Applying the Gaussian mixture model to generate large synthetic data from a small data set," in *Construction Research Congress*, 2020, pp. 1–10, [Online]. Available: [https://www.researchgate.net/profile/Chanachok\\_Chokwitthaya/publication/335662593\\_Appling\\_the\\_Gaussian\\_Mixture\\_Model\\_to\\_Generate\\_Large\\_Synthetic\\_Data\\_from\\_a\\_Small\\_Data\\_Set/links/5d7283a84585151ee4a0f655/Applying-the-Gaussian-Mixture-Model-to-Generate-Lar](https://www.researchgate.net/profile/Chanachok_Chokwitthaya/publication/335662593_Appling_the_Gaussian_Mixture_Model_to_Generate_Large_Synthetic_Data_from_a_Small_Data_Set/links/5d7283a84585151ee4a0f655/Applying-the-Gaussian-Mixture-Model-to-Generate-Lar).
- [150] K. Y. Lee, "Local fuzzy PCA based GMM with dimension reduction on speaker identification," *Pattern Recognition Letters*, vol. 25, no. 16, pp. 1811–1817, 2004, doi: 10.1016/j.patrec.2004.07.006.
- [151] H. Rijal, "Investigation of comfort temperature and occupant behavior in Japanese houses during the hot and humid season," *Buildings*, vol. 4, no. 3, pp. 437–452, 2014, doi: 10.3390/buildings4030437.
- [152] G. Y. Yun and K. Steemers, "Time-dependent occupant behaviour models of window control in summer," *Building and Environment*, vol. 43, no. 9, pp. 1471–1482, 2008, doi: 10.1016/j.buildenv.2007.08.001.
- [153] P. V O'neil, *Advanced engineering mathematics*. Cengage learning, 1991.
- [154] C. F. Reinhart and K. Voss, "Monitoring manual control of electric lighting and blinds," *Lighting research and technology*, vol. 35, no. 3, pp. 243–260, 2003, doi: 10.1191/1477153503li064oa.
- [155] C. Chokwitthaya, R. Dibiano, S. Saeidi, S. Mukhopadhyay, and Y. Zhu, "Enhancing the prediction of artificial lighting control behavior using virtual reality (VR): A pilot study," in *Construction Research Congress*, 2018, pp. 216–233, doi: 10.1061/9780784481301.022.
- [156] A. Nabijiang, S. Mukhopadhyay, Y. Zhu, R. Gudishala, S. Saeidi, and Q. Liu, "Why do you take that route?," in *Cognitive Science Society*, 2019, pp. 2454–2460, [Online]. Available: [https://cogsci.mindmodeling.org/2019/papers/0424/0424.pdf?fbclid=IwAR0sYRRN2\\_DT Rm27DHvM0irXbCJ4D9PcWRJjJALiOMVWVjkx9JQq1eV9ZrY](https://cogsci.mindmodeling.org/2019/papers/0424/0424.pdf?fbclid=IwAR0sYRRN2_DT Rm27DHvM0irXbCJ4D9PcWRJjJALiOMVWVjkx9JQq1eV9ZrY) (Accessed date: October 12, 2019).
- [157] Q. Liu and S. Mukhopadhyay, "Unsupervised learning using pretrained CNN and associative memory bank," in *International Joint Conference on Neural Networks*, 2018, pp. 1–8, doi: 10.1109/IJCNN.2018.8489408.
- [158] A. Ahmed, N. E. Korres, J. Ploennigs, H. Elhadi, and K. Menzel, "Mining building performance data for energy-efficient operation," *Advanced Engineering Informatics*, vol.

- 25, no. 2, pp. 341–354, 2011, doi: 10.1016/j.aei.2010.10.002.
- [159] B. Si, J. Wang, X. Yao, X. Shi, X. Jin, and X. Zhou, “Multi-objective optimization design of a complex building based on an artificial neural network and performance evaluation of algorithms,” *Advanced Engineering Informatics*, vol. 40, no. March, pp. 93–109, 2019, doi: 10.1016/j.aei.2019.03.006.
  - [160] A. J. Sonta, P. E. Simmons, and R. K. Jain, “Understanding building occupant activities at scale: An integrated knowledge-based and data-driven approach,” *Advanced Engineering Informatics*, vol. 37, no. September 2017, pp. 1–13, 2018, doi: 10.1016/j.aei.2018.04.009.
  - [161] H. Lee, C. R. Ahn, and N. Choi, “Fine-grained occupant activity monitoring with Wi-Fi channel state information: Practical implementation of multiple receiver settings,” *Advanced Engineering Informatics*, vol. 46, no. January, p. 101147, 2020, doi: 10.1016/j.aei.2020.101147.
  - [162] J. Lin, R. Zhu, N. Li, and B. Becerik-Gerber, “Do people follow the crowd in building emergency evacuation? A cross-cultural immersive virtual reality-based study,” *Advanced Engineering Informatics*, vol. 43, no. July 2019, p. 101040, 2020, doi: 10.1016/j.aei.2020.101040.
  - [163] Z. Feng, V. A. González, C. Mutch, R. Amor, A. Rahouti, A. Baghouz, N. Li, and G. Cabrera-Guerrero, “Towards a customizable immersive virtual reality serious game for earthquake emergency training,” *Advanced Engineering Informatics*, vol. 46, no. June, p. 101134, 2020, doi: 10.1016/j.aei.2020.101134.
  - [164] J. M. Davila Delgado, L. Oyedele, P. Demian, and T. Beach, “A research agenda for augmented and virtual reality in architecture, engineering and construction,” *Advanced Engineering Informatics*, vol. 45, no. June, p. 101122, 2020, doi: 10.1016/j.aei.2020.101122.
  - [165] G. Ozcelik and B. Becerik-Gerber, “Benchmarking thermoception in virtual environments to physical environments for understanding human-building interactions,” *Advanced Engineering Informatics*, vol. 36, no. April, pp. 254–263, 2018, doi: 10.1016/j.aei.2018.04.008.
  - [166] C. Chokwitthaya, Y. Zhu, S. Mukhopadhyay, and E. Collier, “Augmenting building performance predictions during design using generative adversarial networks and immersive virtual environments,” *Automation in Construction*, vol. 119, p. 103350, 2020, doi: 10.1016/j.autcon.2020.103350.
  - [167] M. Weisberg, “Robustness analysis,” *Philosophy of Science*, vol. 73, no. 5, pp. 730–742, 2006.
  - [168] V. Asghari, Y. F. Leung, and S. C. Hsu, “Deep neural network based framework for complex correlations in engineering metrics,” *Advanced Engineering Informatics*, vol. 44, no. February, p. 101058, 2020, doi: 10.1016/j.aei.2020.101058.

- [169] X. Yang, W. Chen, A. Li, C. Yang, Z. Xie, and H. Dong, “BA-PNN-based methods for power transformer fault diagnosis,” *Advanced Engineering Informatics*, vol. 39, no. December 2018, pp. 178–185, 2019, doi: 10.1016/j.aei.2019.01.001.
- [170] G. H. Kruger, A. J. Shih, D. G. Hattingh, and T. I. Van Niekerk, “Intelligent machine agent architecture for adaptive control optimization of manufacturing processes,” *Advanced Engineering Informatics*, vol. 25, no. 4, pp. 783–796, 2011, doi: 10.1016/j.aei.2011.08.003.
- [171] R. Pasquier and I. F. C. Smith, “Robust system identification and model predictions in the presence of systematic uncertainty,” *Advanced Engineering Informatics*, vol. 29, no. 4, pp. 1096–1109, 2015, doi: 10.1016/j.aei.2015.07.007.
- [172] C. J. Hopfe and J. L. M. Hensen, “Uncertainty analysis in building performance simulation for design support,” *Energy and Buildings*, vol. 43, no. 10, pp. 2798–2805, 2011, doi: 10.1016/j.enbuild.2011.06.034.
- [173] J. C. Helton, J. D. Johnson, W. L. Oberkampf, and C. J. Sallaberry, “Representation of analysis results involving aleatory and epistemic uncertainty,” *International Journal of General Systems*, vol. 39, no. 6, pp. 605–646, 2010, doi: 10.1080/03081079.2010.486664.
- [174] H. Bae, R. V. Grandhi, and R. A. Canfield, “Epistemic uncertainty quantification techniques including evidence theory for large-scale structures,” *Computers and Structures*, vol. 82, no. 13-14(2004), pp. 1101–1112, 2004, doi: 10.1016/j.compstruc.2004.03.014.
- [175] S. M. M. Dezfooli, A. Fawzi, O. Fawzi, P. Frossard, and S. Soatto, “Robustness of classifiers to universal perturbations: A geometric perspective,” *6th International Conference on Learning Representations, ICLR 2018 - Conference Track Proceedings*, pp. 1–15, 2018.
- [176] A. Rozsa, M. Günther, E. M. Rudd, and T. E. Boulton, “Facial attributes: Accuracy and adversarial robustness,” *Pattern Recognition Letters*, vol. 124, pp. 100–108, 2019, doi: 10.1016/j.patrec.2017.10.024.
- [177] A. Fawzi, O. Fawzi, and P. Frossard, “Analysis of classifiers’ robustness to adversarial perturbations,” *Machine Learning*, vol. 107, no. 3, pp. 481–508, 2018, doi: 10.1007/s10994-017-5663-3.
- [178] V. Tyagi and C. Wellekens, “On desensitizing the mel-cepstrum to spurious spectral components for robust speech recognition,” *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, vol. I, pp. 1–9, 2005, doi: 10.1109/ICASSP.2005.1415167.
- [179] M. Cisse, Y. Adi, N. Neverova, and J. Keshet, “Houdini: Fooling deep structured visual and speech recognition models with adversarial examples,” *Advances in Neural Information Processing Systems*, vol. 2017-Decem, pp. 6978–6988, 2017.



- [180] A. Jungo, R. Meier, E. Ermis, M. Blatti-Moreno, E. Herrmann, R. Wiest, and M. Reyes, “On the effect of inter-observer variability for a reliable estimation of uncertainty of medical image segmentation,” in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2018, pp. 682–690, doi: 10.1007/978-3-030-00928-1\_77.
- [181] D. Karmon, D. Zoran, and Y. Goldberg, “LaVAN: Localized and visible adversarial noise,” *35th International Conference on Machine Learning, ICML 2018*, vol. 6, pp. 3903–3911, 2018.
- [182] D. Kang, Y. Sun, D. Hendrycks, T. Brown, and J. Steinhardt, “Testing Robustness Against Unforeseen Adversaries,” 2019, [Online]. Available: <http://arxiv.org/abs/1908.08016>.
- [183] L. Engstrom, B. Tran, D. Tsipras, L. Schmidt, and A. Madry, “Exploring the landscape of spatial robustness,” *36th International Conference on Machine Learning, ICML 2019*, vol. 2019-June, pp. 3218–3238, 2019.
- [184] N. Carlini and D. Wagner, “Audio adversarial examples: Targeted attacks on speech-to-text,” in *Proceedings - 2018 IEEE Symposium on Security and Privacy Workshops, SPW 2018*, 2018, pp. 1–7, doi: 10.1109/SPW.2018.00009.
- [185] Y. Qin, N. Carlini, I. Goodfellow, G. Cottrell, and C. Raffel, “Imperceptible, Robust, and targeted adversarial examples for automatic speech recognition,” *36th International Conference on Machine Learning, ICML 2019*, vol. 2019-June, pp. 9141–9150, 2019.
- [186] M. Alzantot, B. Balaji, and M. Srivastava, “Did you hear that? Adversarial Examples Against Automatic Speech Recognition,” no. Nips, 2018, [Online]. Available: <http://arxiv.org/abs/1801.00554>.
- [187] R. Jia and P. Liang, “Adversarial examples for evaluating reading comprehension systems,” *EMNLP 2017 - Conference on Empirical Methods in Natural Language Processing, Proceedings*, pp. 2021–2031, 2017, doi: 10.18653/v1/d17-1215.
- [188] W. Thomas and R. D. Cook, “Assessing influence on predictions from generalized linear models,” *Technometrics*, vol. 32, no. 1, pp. 59–65, 1990, doi: 10.1080/00401706.1990.10484593.
- [189] A. Haghnegahdar and S. Razavi, “Insights into sensitivity analysis of Earth and environmental systems models: On the impact of parameter perturbation scale,” *Environmental Modelling and Software*, vol. 95, pp. 115–131, 2017, doi: 10.1016/j.envsoft.2017.03.031.
- [190] A. Fawzi, S. M. Moosavi-Dezfooli, and P. Frossard, “The Robustness of Deep Networks: A Geometrical Perspective,” *IEEE Signal Processing Magazine*, vol. 34, no. 6, pp. 50–62, 2017, doi: 10.1109/MSP.2017.2740965.
- [191] Y.-C. Hsu and Z. Kira, “Neural network-based clustering using pairwise constraints,” pp. 1–12, 2015, [Online]. Available: <http://arxiv.org/abs/1511.06321>.

- [192] H. Hosseini, B. Xiao, and R. Poovendran, “Google’s cloud vision API is not robust to noise,” *Proceedings - 16th IEEE International Conference on Machine Learning and Applications, ICMLA 2017*, vol. 2017-Decem, pp. 101–105, 2017, doi: 10.1109/ICMLA.2017.0-172.
- [193] G. J. Barlow and C. K. Oh, “Robustness analysis of genetic programming controllers for unmanned aerial vehicles,” *GECCO 2006 - Genetic and Evolutionary Computation Conference*, vol. 1, pp. 135–142, 2006, doi: 10.1145/1143997.1144023.
- [194] B. Li, C. Chen, W. Wang, and L. Carin, “Certified Adversarial Robustness with Additive Noise,” *Advances in Neural Information Processing Systems*, pp. 9459–9469, 2019, [Online]. Available: <http://arxiv.org/abs/1809.03113>.
- [195] R. Munir, “Robustness Analysis of Selective Image Encryption Algorithm Based on Arnold Cat Map Permutation,” in *Proceedings of 3rd Makassar International Conference on Electrical Engineering and Informatics*, 2012, pp. 1–5.
- [196] J.-H. Huang, C. D. Dao, M. Alfadly, C. H. Yang, and B. Ghanem, “Robustness Analysis of Visual QA Models by Basic Questions,” *arXiv preprint arXiv:1709.04625*, 2017, [Online]. Available: <http://arxiv.org/abs/1709.04625>.
- [197] M. Ahmadlou and H. Adeli, “Enhanced probabilistic neural network with local decision circles: A robust classifier,” *Integrated Computer-Aided Engineering*, vol. 17, no. 3, pp. 197–210, 2010, doi: 10.3233/ICA-2010-0345.
- [198] D. Rolnick, A. Veit, S. Belongie, and N. Shavit, “Deep Learning is Robust to Massive Label Noise,” 2017, [Online]. Available: <http://arxiv.org/abs/1705.10694>.
- [199] W. Liu and W. Lin, “Additive white gaussian noise level estimation in SVD domain for images,” *IEEE Transactions on Image Processing*, vol. 22, no. 3, pp. 872–883, 2013, doi: 10.1109/TIP.2012.2219544.
- [200] A. Liu, X. Liu, J. Fan, Y. Ma, A. Zhang, H. Xie, and D. Tao, “Perceptual-Sensitive GAN for Generating Adversarial Patches Aishan,” in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2019, vol. 33, pp. 1028–1035, doi: 10.1609/aaai.v33i01.33011028.
- [201] X. Y. Sun, L. T. H. Newham, B. F. W. Croke, and J. P. Norton, “Three complementary methods for sensitivity analysis of a water quality model,” *Environmental Modelling and Software*, vol. 37, pp. 19–29, 2012, doi: 10.1016/j.envsoft.2012.04.010.
- [202] B. Moaveni, J. P. Conte, and F. M. Hemez, “Uncertainty and sensitivity analysis of damage identification results obtained using finite element model updating,” *Computer-Aided Civil and Infrastructure Engineering*, vol. 24, no. 5, pp. 320–334, 2009, doi: 10.1111/j.1467-8667.2008.00589.x.
- [203] F. J. Massey, “The Kolmogorov-Smirnov Test for Goodness of Fit,” *Journal of the American statistical Association*, vol. 46, no. 253, pp. 68–78, 1951.

- [204] L. Wang, P. Mathew, and X. Pang, “Uncertainties in energy consumption introduced by building operations and weather for a medium-size office building,” *Energy and Buildings*, vol. 53, no. September, pp. 152–158, 2012, doi: 10.1016/j.enbuild.2012.06.017.
- [205] D. Hendrycks and T. Dietterich, “Benchmarking neural network robustness to common corruptions and perturbations,” in *International Conference on Learning Representations*, 2019, pp. 1–16, [Online]. Available: <https://arxiv.org/pdf/1903.12261.pdf> (Accessed date: June 17, 2020).

## **VITA**

Chanachok Chokwitthaya was born in Bangkok, Thailand. He received his bachelor's degree in Civil Engineering in March 2007 from Kasetsart University, Bangkok, Thailand. Upon graduation, Chanachok worked as a structural design engineer and office/field construction engineer for 4 years. In August 2011, Chanachok joined the Department of Civil Engineering at Lamar University to pursue his Master of Science in Civil Engineering. In August 2013, he continued his studies at Louisiana State University (LSU) to pursue a Ph.D. in Engineering Science with an emphasis on construction management. Chanachok has worked as a research assistant at LSU from 2013 to the present. He expects to obtain his doctoral degree in May 2020.