Towards Improved Thermal Comfort Predictions and Building Energy Savings: Bayesian Modelling of Indoor Environmental Design Conditions

by

Sarah Crosby

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF

Doctor of Philosophy

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES
(Mechanical Engineering)

The University of British Columbia
(Vancouver)

January 2023

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the thesis entitled:

Towards Improved Thermal Comfort Predictions and Building Energy Savings: Bayesian Modelling of Indoor Environmental Design Conditions

submitted by **Sarah Crosby** in partial fulfillment of the requirements for the degree of **Doctor of Philosophy** in **Mechanical Engineering**.

Examining Committee:

Dr. Adam Rysanek, Professor, School of Architecture and Landscape Architecture, UBC
Supervisor

Dr. Steven Rogak, Professor, Mechanical Engineering, UBC
Co-Supervisor

Dr. Sheryl Staub-French, Professor, Civil Engineering, UBC
University Examiner

Dr. Amanda Giang, Professor, Institute for Resources, Environment and Sustainability, UBC
University Examiner

Dr. Elie Azar, Professor, Civil and Environmental Engineering, Carleton University
External Examiner

Additional Supervisory Committee Members:

Dr. Naomi Zimmerman, Professor, Mechanical Engineering, UBC
Supervisory Committee Member

Abstract

The judgment of thermal comfort is a cognitive process that is influenced, not only by measurable indoor environmental conditions but also by less tangible aspects of an occupant's well-being and overall satisfaction. Recent studies have examined the multi-domain nature of thermal comfort to bridge the performance gap between model-predicted and measurements of thermal comfort. This thesis seeks to inform a well-known research gap with respect to standard models of thermal comfort: that seminal data-informed models have not always accurately predicted true thermal comfort observations from independent field studies. This thesis presents a novel approach that involves the use of Bayesian inference to predict thermal comfort as a function of both thermal and *non-thermal* metrics of indoor environmental quality.

Bayesian regression was performed on a large field dataset to investigate whether perceived thermal comfort can be attributed in a measurable and/or significant manner to one or several non-thermal parameters of indoor environmental quality. Posterior results revealed that higher CO_2 concentrations are independently correlated with lower incidences of thermal satisfaction in open-plan offices. At indoor temperatures of 23.5 °C, the probability of an occupant feeling thermally

satisfied at measured CO_2 levels of 550 ppm was 0.62 [0.54 - 0.69, 95% CrI]. This decreased to 0.28 [0.17-0.42, 95% CrI] at 750 ppm. Further, this is the first work to demonstrate that predictions of thermal comfort can be improved upon adding measurements of indoor CO_2 concentrations.

The new data-driven thermal comfort model is integrated into a building energy model framework to predict occupants' thermal satisfaction based on thermal indoor environmental conditions and ventilation rates. Four different post-COVID-19 occupancy schedules were investigated to reflect and compare different occupancy profiles for post-COVID-19 hybrid work models. The simulation results showed that it might be possible to increase the ventilation rates with minimal building heating energy demand increase while maintaining the levels of occupants' thermal comfort. This thesis presented a solution for building managers that have been under pressure to increase the current amounts of fresh air to lower the risk of spreading the COVID-19 virus, and other diseases, indoors.

Lay Summary

Prior studies have suggested that occupants who are generally satisfied with many non-thermal indoor environmental conditions are more likely to be satisfied with thermal conditions as well. This thesis takes advantage of the emerging awareness in research on the multidomain nature of thermal comfort and presents a novel approach to investigate whether perceived thermal comfort can be attributed in a measurable and/or significant manner to one or several non-thermal indoor environmental quality parameters. Posterior results suggested that predictions of thermal comfort can be improved by adding measurements of indoor CO_2 concentrations. Building energy simulation results revealed that it may be possible to increase the ventilation rates with minimal building heating energy demand increase while maintaining the levels of thermal comfort. This thesis presents a solution for building managers that have been under pressure to increase the current amounts of fresh air to lower the risk of spreading the COVID-19 virus.

Preface

This dissertation is an original intellectual product of the author, Sarah Crosby.

This dissertation is an integration of published manuscripts in scholarly journals and conference proceedings as follows. Manuscripts are slightly modified for the dissertation formatting style and coherence.

Various results from chapters 2, 3, 4, and 5 of this dissertation have been presented as oral presentations at Comfort at the Extremes (CATE) 2022 Conference, Indoor Air 2022 Conference, Healthy Buildings America 2021 conference, ACM BuildSys 2021 conference, Building Simulation (BS21) 2021 conference, IBPSA eSIM 2020 conference, IAQVEC 2019 conference.

A version of Chapter 2 has been published [Crosby, S., Newsham, G., Veitch, J., Rogak, S., Rysanek, A. (2019). "Bayesian Inference of Thermal Comfort: Evaluating the Effect of "Well-Being" on Perceived Thermal Comfort in Open-Plan Offices", IOP Conference Series: Materials Science and Engineering, vol. 609, No. 4, p. 042028, September 2019. doi:10.1088/1757-899X/609/4/042028]. This author completed all the analytical work and all the data analysis and wrote the manuscript under the supervision of Dr. Adam Rysanek. Dr. Guy Newsham and Dr. Jennifer Veitch contributed in providing the COPE field IEQ dataset and re-

viewing the manuscript. Dr. Steven Rogak contributed in the paper's revisions phases. A version of Chapter 2 has also been published [Crosby, S., Rysanek, A. (2020).“Correlations between thermal satisfaction and non-thermal conditions of indoor environmental quality: Bayesian inference of a field study of offices”, Journal of Building Engineering,35, 102051. doi:10.1016/j.jobe.2020.102051]. This author completed all the analytical work and all of the data analysis and wrote the manuscript under the supervision of Dr. Adam Rysanek.

A version of Chapter 3 has been published [Crosby, S., Rysanek, A. (2022). “Predicting thermal satisfaction as a function of indoor CO2 levels: Bayesian modelling of new field data”, Building and Environment.108569, ISSN 0360-1323. doi:10.1016/j.buildenv.2021.108569] and [Crosby, S., Rysanek, A. (2021). “Extending the Fanger PMV model to include the effect of non-thermal conditions on thermal comfort”, In Proceedings of eSIM 2020: Building Simulation meets a global pandemic, Vancouver, Canada] and [Crosby, S., Rysanek, A. (2021). “Towards Improved Thermal Comfort Predictions for Building Controls: Hierarchical Bayesian Modelling of Indoor Environmental Design Conditions”. In The 8th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation (BuildSys), Coimbra, Portugal.doi:10.1145/3486611.3491128]. This author completed all the experimental work and all of the data analysis and prepared the manuscript under the supervision of Dr. Adam Rysanek.

A version of Chapter 4 has been published [Crosby, S., Rysanek, A. (2022). “Towards improved thermal comfort predictions and higher energy savings: building energy model of an open-plan office based on indoor CO2 and temperature controls”. Proceedings of Indoor Air 2022.] and [Crosby, S., Rysanek, A. “A Novel Multi-Domain Model for Thermal Comfort which Includes Building In-

door CO2 Concentrations". Proceedings of Building Simulation 2021, Bruges, Belgium. doi:10.26868/25222708.2021.30760]. This author completed all the analytical work and all the building modelling and simulation work, and wrote the manuscript under the supervision of Dr. Adam Rysanek.

A version of Chapter 5 has been published [Crosby, S., Rysanek, A. (2022). "On Higher Ventilation Rates and Energy Efficiency in Post-COVID-19 Buildings: A New Thermal Comfort Model based on Indoor CO2 Levels and Temperature". CATE 2022 proceeding]. This author completed all the analytical work, data analysis, and all the building modelling and simulation work, and wrote the manuscript under the supervision of Dr. Adam Rysanek.

Ethics approval for this work was sought and received from the UBC Office of Research Ethics (Ethics approval certificate Ref: H19-01364).

Contents

List of Tables

List of Figures

List of Abbreviations

ACH Air changes per hour

ASHRAE American Society of Heating, Refrigerating, and Air-Conditioning Engineers

BEM building energy model

CIRS Centre for Interactive Research on Sustainability

COLAB Google Collaboratory

COPE Cost-effective Open Plan Environments

CRI Credible Interval

ELPD log predictive density

HRV Heating Recovery Ventilation

IAQ Indoor Air Quality

IEQ Indoor Environmental Quality

LOO-CV Leave-One-Out Cross-Validation

MAPE Maximum a posteriori estimates

MCMC Markov chain Monte Carlo

MLE Maximum likelihood Estimation

MPC model predictive control

MRT mean radiant temperature

NRC National Research Council of Canada

PDF Probability Density Function

PPC Predictive Posterior Checks

PSIS pareto-smoothed importance sampling

SMC Sequential Monte Carlo

TRNSYS Transient System Simulation Tool

TVOC total volatile organic compounds

UBC University of British Columbia

VAV Variable Air Volume

WAIC Watanabe-Akaike Information Criteria

WHO World Health Organization

WWR window-to-wall ratio

List of Symbols

Symbol	Unit	Description
C	ppm	Indoor CO_2 concentrations
C_{out}	ppm	Outdoor CO_2 concentrations
D	-	Thermal Dissatisfaction (binary variable)
E	-	Speech intelligibility index
F	-	Thermal IEQ parameters
L	Lux	Measured desktop illuminance
M	ºC	Mean radiant temperature
N	dBA	A-weighted indoor noise levels
P	cm	Partition height
p(S)	-	Probability of Thermal Satisfaction
p(D)	-	Probability of Thermal Dissatisfaction
R	%	Indoor relative Humidity
S	-	Thermal Satisfaction (binary variable)
T	ºC	Indoor air temperature
TS	-	Thermal Satisfaction Responses (on a Likert scale)
V	m/s	Indoor air velocity
$V^{\cdot}$	L/s	Ventilation rate
V_{inf}	L/s	Infiltration rate
VC	L/s	Volume of indoor CO_2
VC_{gen}	L/s	Occupants' CO_2 generated rate
W	-	Non-thermal IEQ parameters

Greek Letters

β	Model parameter
Δt	Time difference between CO_2 measurements
μ	Probability distribution mean
σ	Probability distribution variance

Acknowledgments

I would like to start by thanking Dr. Adam Rysanek, my research supervisor. Adam has been an exceptional and inspiring supervisor, his guidance and support not only made my PhD journey possible, but also enjoyable. He is brilliant, approachable and caring. His encouragement and positive attitude kept me motivated and inspired during the most challenging times. Not to mention, his significant role in developing my fascination with Bayesian statistics. Adam is not only my PhD supervisor, but also my mentor, friend and role model. For everything he has done for me, I am and will always be grateful.

I would like to express my sincere gratitude to my co-supervisor, Dr. Steven Rogak, for his continuous guidance, encouragement and support. My thanks also go to Dr. Naomi Zimmerman for the thoughtful suggestions and productive discussions throughout the different PhD phases.

I am also very grateful to my two university examiners, Dr. Sheryl Staub-French and Dr. Amanda Giang, for their valuable time reviewing my work, reading my thesis, and for the insightful discussion during my PhD defence. I would like to thank the external examiner of my PhD thesis, Dr. Elie Azar, for the time and effort he took to review my thesis and for his thoughtful review of my work.

I have to also thank Dr. Jennifer Veitch and Dr. Guy Newsham from the National Research Council of Canada (NRC) for their significant contribution in making the 'Cost Effective Open-plan Environment' (COPE) dataset accessible for this thesis. The COPE dataset played a key role in the establishment of this work.

I would also like to thank my former colleagues at BDRG, Zoe LeHong, Merriam Vahedi, and Karen Lai, for their direct support of the experimentation undertaken in the development of the UBC field study.

To my friends at the Building Decisions Research Group, Denon Sheppard, Justin McCarty, Sanyogita Manu, Claude Demers-Belanger, Brendan Buchanan Dee, Donna Vakalis, Jeremy Wong, Maverick Chan, Julian Donges, and Kyle Gerrard, I could not have wished for better friends and office mates. Thank you for your support and the amazing company.

To my friends at the Aerosols lab and the Clean Energy Research Centre (CERC), Dr. Amin Engarnevis, Dr. Pooyan Kheirkhah, Dr. Michael Karpinski-Leydier, René Zepeda, Alex Sylvester, Keyhan Babaee, Dr. Alberto Baldelli, Dr. Mahdiar Khosravi, Aditya Singh, Dr. Jeremy Rochusen and my friends at the Centre for Interactive Research on Sustainability and the Mechanical Engineering Department, Christina Draeger, Rowan Waldron, Tim Herron, Dr. Athena Liu, Somesh Bhatia, Miguel Villalba, Harsh Pokharna, Faisal Al-Qurooni, Dr. Juuso Heikkinen, Dr. Alondra Renteria, Ratul Das, Claire Grégoire, Filipe Broetto, you have been amazing friends. Thank you for the wonderful time we have had together.

To my parents, Salwa and Wael, and my brother, Tamer, thank you for all the unconditional love, encouragement, and support, for that I am very grateful.

Finally, to Moustafa and Nadine, my family and my support system, I could not have done this without you and I am forever grateful to have you in my life.

Dedication

To Nadine, my smart, brave, and determined daughter.

To Moustafa, my motivation, inspiration, my partner, and my best friend.

To Adam, my supervisor, role model, and my very good friend.

Chapter 1

Introduction

1.1 Background

From its origins in vernacular architecture to modern-day environmental systems control, the regulation of the indoor thermal environment in support of occupant comfort is one of the principal functions of a building [68]. In recent years, the perception of the thermal environment has become one of the important aspects of the built environment [124]. The dissatisfaction with indoor thermal conditions is known to be one of the most common sources of complaints by occupants of commercial buildings with respect to indoor environmental quality [32]. These complaints are not without merit, either. When building occupants are found to be dissatisfied with their thermal environment, it has been observed that their overall health, productivity, well-being, and satisfaction with the workplace are adversely affected [40, 58, 60, 61, 130].

Accurate predictions of thermal comfort in a built environment has great impact on building energy saving [142]. In studies that have aimed to compare the

economic costs and benefits of occupant thermal comfort vs. building energy consumption, it has been found that employee-compensation costs attributed to poor indoor environmental quality are more than 130 times greater the cost of providing heating and air-conditioning services, concluding that occupant comfort should not be compromised in the pursuit of energy-efficient building systems [58]. In a specific study on this issue, Jensen et al. [61] found that when space cooling systems are adjusted to improve energy-efficiency, the cost of losses in building occupant productivity was higher than the reduced energy cost (i.e., investing in more cooling energy resulted in a 5 times higher benefit-to-cost ratio caused by increased employee performance).

Several recent research studies have focused on the implications of occupants' thermal dissatisfaction on cognitive performance, health, well-being, productivity, and workplace satisfaction [2, 40, 41, 58, 61, 130]. Thermal comfort, defined as "the condition of mind that expresses satisfaction with the thermal environment" [7], is not a state of a condition, as its judgment is a cognitive process influenced, not only by measurable indoor environmental conditions but also - in theory - by less tangible aspects of an occupant's own well-being and overall satisfaction [35, 82]. In recent years, it has been increasingly suggested that an occupant's thermal comfort can be demonstrably influenced by changes to the occupant's mood, well-being, and overall satisfaction with the built environment [2, 63, 82, 102, 109]. It has been argued that, if it is accepted that an occupant's judgment of thermal comfort is a cognitive process, then perceived thermal comfort may be affected by the psychological effect of many physical conditions that occupants encounter in the built environment, not only thermal conditions [35, 82]. Emerging studies observed correlations between an occupant's overall mood and thermal comfort [62],

and others have found similar correlations with measured parameters of indoor environmental quality (Indoor Environmental Quality (IEQ)), pointing to the need for a more holistic interpretation of thermal comfort [60, 132].

In field studies, these arguments have been supported by observed relationships between acoustic and visual comfort, biophilia, indoor air quality, office layout, and overall occupant comfort and satisfaction with the indoor environment [2, 54, 63, 107]. Field analyses carried out in these studies, which includes both long-term and short-term exposure analyses [9], found independent relationships between occupant thermal acceptability and acceptability of other conditions, such as noise levels and air quality. Overall, evidence in support of the multi-perceptual, or multi-domain, nature of thermal comfort as well as the existing correlations between thermal comfort and non-thermal indoor environmental metrics of building performance is growing [110].

Although prevailing thermal comfort models have been used in building international codes for some time, it has been found that they have many limitations when applied to buildings' control systems [65] and show poor predictive performance when applied to individuals [8].

Research has shown that it has been possible to find discrepancies between standard model-predicted thermal comfort and thermal comfort observations, which affect building operations and controls [77]. It has been suggested that this performance gap could be filled by including new parameters in thermal comfort models, such as measurements of CO_2 concentrations [44].

Filling this performance gap by improving predictions of thermal comfort in office buildings will not just reduce building energy use [61], but also will contribute to optimizing the use of this energy to provide better indoor environmental

conditions, Indoor Air Quality (IAQ), and healthier indoor environments [58]. This is becoming more critical, particularly nowadays, as global issues like the climate crisis and the emergence of the COVID-19 pandemic are rising. Recently, there has been a universal movement towards more energy efficient buildings and adequate indoor ventilation [3, 90]. It is becoming increasingly necessary to increase the amount of indoor fresh air to mitigate the risk of airborne diseases transmission like SARS-CoV-2 [138]. This leads to a reduction in indoor CO_2 concentration setpoint values [3, 12, 103].

The added cost of moving more air as well as heating a larger volume of air poses a considerable challenge to building managers and needs to be taken into consideration when modelling buildings [3]. To address these issues, model predictive control (MPC) has been recently adapted into indoor thermal environment control systems [24, 81]. Research has found that MPC can reduce building energy use while maintaining occupant's thermal comfort levels [81]. Aiming to mitigate the performance gap between observations and predictions of thermal comfort that affect building control systems, this thesis seeks to inform a well-known research challenge with respect to standard models of thermal comfort: that seminal data-informed models have not always accurately predicted true thermal comfort observations from independent field studies [77]. Specifically, this thesis aims to take advantage of an emerging awareness in research of the inter-dependencies between perceived thermal comfort and overall IEQ, and will evaluate an approach to predicting thermal comfort as a function of both thermal and *non-thermal* metrics of IEQ. The thesis will develop a novel predictive model of thermal comfort to improve prediction accuracy of occupants' thermal satisfaction. Moreover, the thesis will look at methods to implement model predictive control to control for

both heating and ventilation setpoints so that fresh air amounts can be increased with no increase in energy demand and without compromising occupants' thermal satisfaction levels.

1.1.1 Identified Research Questions

This PhD thesis seeks to find answers to the following four main research questions:

1. Can thermal comfort, as perceived by occupants in a field dataset, be attributed in a measurable and/or significant manner to one or several non-thermal parameters of IEQ?

2. Which non-thermal metrics of IEQ are most significantly correlated with perceived thermal comfort, if any?

3. Can the prediction accuracy of thermal satisfaction be improved upon the addition of measurements of non-thermal IEQ parameters?

4. Can we formulate a new predictive model of thermal comfort and integrate it into building control systems to increase the ventilation rates with minimal energy demand increase and without compromising occupants' thermal satisfaction levels?

1.2 Literature Review

1.2.1 Existing Models of Thermal Comfort

The American Society of Heating, Refrigerating, and Air-Conditioning Engineers (ASHRAE) defines thermal comfort as "the condition of the mind in which satis-

faction is expressed with the thermal environment" [7]. In the past half-century, ASHRAE, like other building standard authorities, have adopted standard mathematical models for predicting thermal comfort as a function of measurable physical parameters.

The "PMV-PPD" model of thermal comfort, developed by Ole Fanger from climate chamber experiments of 1,386 test subjects in the 1960s [38], allows building designers and operators to attribute the predicted perceived thermal sensation and thermal satisfaction of building occupants to be functions of indoor thermodynamic conditions such as temperature, humidity, and air speed as well as occupants' metabolic rate, and their clothing levels. Though the adoption of the PMV-PPD model use has been widespread for some time, it has also been contested [21, 137]. Laboratory data originally collected by Fanger to define the original PMV-PPD model has not always represented accurately the thermal sensation of real-world occupants across different regional, cultural, and/or climatic contexts [55]. Besides, the model is not applicable to nonuniform environment, sleep environment, and not applicable for elderly as the data used to develop Fanger's model was collected only from adults [142]. Several adaptations, improvements, or recalibrations of the original standard models, such as the PMV-PPD model, have been proposed to address these shortcomings, with some suggesting adoption of new parameters to this particular model [21], [92], [77]. There is, and will continue to be, ongoing discussions regarding the appropriateness of the PMV-PPD model and whether the formulation of altogether new thermal comfort prediction models is warranted given the evidence collected since Ole Fanger's seminal work.

An adaptive model of thermal comfort was developed by Brager and de Dear [32] in an attempt to consider likely adaptive actions of building occupants [140].

Based on the RP-884 database, which contained approximately 21,000 thermal comfort field data from a wide range of climatic zone, the adaptive model proposes that contextual factors and past thermal history affect one's thermal preferences and thus thermal satisfaction [92]. The model relates the acceptable indoor air temperature of naturally-ventilated buildings to a function of prevailing outdoor air temperature only. In a similar manner, while the adaptive comfort model provides an improved representation of true building occupant behaviour with respect to thermal comfort in naturally-ventilated buildings, studies such as the work of Carlucci et al. [17], de Dear et al. [34], and Yau et al. [140] have found that the model fails to be amenable to buildings with mechanical ventilation systems. Additionally, in a previous study by Sourbon et al., they found that adopting an adaptive thermal comfort model encourages unnecessary cooling energy in summertime [112]. Furthermore, the adaptive model has some limitations when applied to real field setups: for instance, the model can only be applied to naturally ventilated spaces, occupant's metabolic rates should be between 1 and 1.3, the outdoor temperature should fall between 10°C and 33.5°C, and clothing insulation should have values between 0.5 and 1 clo [17].

Data-driven thermal comfort models are emerging. Taking advantage of building data, the models enhance the prediction accuracy of thermal comfort [142] and contribute to the development of personalized thermal comfort [35].

More recently, the development of *personal* thermal comfort models have emerged [37, 83]. A personalized prediction model of thermal comfort takes an individual subject as the unit of data analysis and uses direct feedback for development instead the average response of a large population [39]. This approach makes tailored use of machine learning, indoor sensor measurements, and gather-able occupant data

in order to generate bespoke thermal comfort characteristics of individual occupants [5, 39, 84]. Several personalized models and algorithms for thermal comfort predictions have been developed to support personalized decision-making in healthcare [39]. All existing models of thermal comfort have been developed and continue to be evaluated through the rigorous lens of research, experimentation and validation. In the worst-case conditions, it has been observed that inaccurate representations of occupant's thermal comfort requirements, as dictated by current standard thermal comfort models, have encouraged unnecessary heating in cold weather and unnecessary cooling in warm weather, which in turn increase building energy consumption [99, 112, 121, 140].

1.2.2 Bayesian Modelling of Thermal Comfort

Several modelling approaches have been applied in prior thermal comfort studies. Both Frequentist and Bayesian methods have been used in thermal comfort modelling, yet most of the previous studies made use of existing Frequentist approaches. Bayesian modelling of thermal comfort has well-known advantages over Frequentist methods; for example, Bayesian inference of thermal comfort allows the integration of our prior knowledge of the modelled parameters [98]. This contributes to more accurate predictions while accounting for all existing uncertainties [79]. Further, Bayesian regression analysis treats the regression parameters as uncertain and probabilistic which leads to appropriately expressing the inherent uncertainty of a regression model's fit.

Bayesian processing, which refers to the computational modelling of Bayesian problems, has been used effectively in recent years to improve the characterization of thermal comfort probability distributions using new observational data. For

example, Langevin et al.[77] developed an updated curve for the PMV-PPD relationship using Bayesian analysis of datasets from both laboratory and field settings. They used previous data from the field-based ASHRAE RP-884 database as priors for their Bayesian model [33].

Jensen et al.[61] developed a model that correlated indoor air temperature with the mental performance of office employees using a Bayesian Network approach. They used data from ASHRAE's RP-884 database to develop a correlation between thermal sensation votes and indoor air temperature.

In the field of personalized thermal comfort, although most of the prevailing techniques use frequentist methods, some Bayesian inference techniques have been applied to personal comfort models as well. For instance, Bayesian ridge regression has been used to find a personal thermal comfort model using data collected from an IoT based system and wearable devices [75]. Auffenberg et al. [8] used Bayesian networks to develop a personalized thermal model that can learn individual thermal comfort by predicting thermal sensation of occupants within a specific area. Lee et al. [79] developed a Bayesian modelling approach for learning individual occupants' thermal preferences in office buildings using real-time feedback from occupants. The results showed that the developed approach improved the prediction accuracy for personalized thermal preference profiles. Wong et al. [137] proposed a Bayesian approach to quantify the discrepancies between Fanger's model prediction of thermal satisfaction and observations collected from field studies. The results show that the Bayesian estimation is close to the actual percentage of occupants' thermal dissatisfaction.

1.2.3 Multicontextual Modelling of Thermal Comfort

Until recently, systematic performance criteria for IEQ, which takes into account both psychological and physical indicators of IEQ, have not been explicitly expressed in prevailing building codes and standards. However, in 2014, the International WELL Building Institute launched the WELL building standard, a building performance accreditation scheme focusing solely on the health, comfort and wellness of building occupants [119]. The emergence of the WELL standard occurred in tandem with recent research exploring the interdependencies between thermal comfort, IEQ, and building design.

While there does not appear to be a dispute in literature regarding whether thermal comfort is correlated significantly to thermo-physical parameters of the human body and indoor environment, it has been found that the occupants' satisfaction with several aspects of the indoor environment appear to be correlated with their perception of the thermal environment [35, 60, 109, 132].

In field studies, these arguments have been supported by observed relationships between acoustic and visual comfort, biophilia, indoor air quality, office layout, and overall occupant comfort and satisfaction with the indoor environment [2, 54, 63, 107].

Further, the evidence in support of the existing inter-dependencies between thermal comfort and non-thermal metrics of building performance and indoor environmental quality is growing. Wagner et al. [132] conducted a field study of 50 occupants on the subject of thermal comfort and one's general satisfaction with the workplace (which included satisfaction with air quality, noise and daylight) in office buildings. Results from the IEQ survey revealed that occupants who are

generally satisfied with many non-thermal metrics of IEQ seem more likely to be satisfied with thermal conditions as well.

Jokl et al. [62] found that the optimal operative temperature for thermal comfort is correlated with the general mood and feelings of human beings in addition to thermal conditions. Frontczak et al.[41] found that some arbitrary psychological factors, such as overall job stress, bears influence on occupants' perception of indoor air quality.

In a prior proof of this concept, Jamrozik et al. [60] studied the effects of six well-being factors on occupant's satisfaction in a living lab experiment. They found that building occupants' perceptions of environmental conditions is holistic: dissatisfaction with one type of environmental condition appears to affect occupants' perception of the whole environment and result in dissatisfaction with an unrelated set of environmental conditions. For example, perceived dissatisfaction with air quality may very well influence perceived thermal sensation.

A prior study observed correlations between an occupant's overall mood and thermal comfort [62], and others have found similar correlations with measured parameters of indoor environmental quality (IEQ), pointing to the need for a more holistic interpretation of thermal comfort [132]. Field analyses carried out in these studies, which includes both long-term and short-term exposure analyses [9], found independent relationships between occupant thermal acceptability and acceptability of other conditions, such as noise levels and air quality. Overall, evidence in support of the multi-perceptual, or multi-domain, nature of thermal comfort as well as the existing correlations between thermal comfort and non-thermal indoor environmental metrics of building performance is growing [110].

With regards to conditions of air quality, and its potential impact on perceived

thermal comfort, there remains divergent views in literature. In a prior experimental study, Allen et al. [3] found that modest increases in indoor CO_2 concentrations between 500 and 900 ppm can noticeably reduce cognitive performance of office workers. A different explanation of these results was disputed by Zhang et al. [141], who argued that indoor air bio-effluents attributed to low ventilation rates, which would subsequently cause indoor air CO_2 concentrations to rise, may better represent the direct cause of these effects. Even further, a recent work by Pantelic et al. [51], determined that the concentrations of CO_2 which building occupants inhale are significantly higher than concentrations in the ambient air due to their significant re-breathing of exhaled air. Pantelic's work would suggest that modest variations in ambient CO_2 are less likely to significantly affect the concentration of CO_2 that occupants breath.

1.2.4 Quantifiable Correlations Between Thermal Comfort and IEQ

Though aforementioned studies and several others [110], have helped to establish the view that thermal comfort is a multi-perceptual and multi-domain condition, studies that provide a quantifiable and/or generalizable predictive correlation between thermal comfort and multi-domain conditions are still relatively few. In this subsection, several are reviewed.

Recent works that examined the cross-modal effects of acoustical environmental quality on thermal perception have found that occupant's perceived thermal dissatisfaction increases with increased noise levels [54, 94, 101, 102, 139]. In a prior experimental study conducted on 18 subjects in a climate chamber, Pellerin et al. [102] found that a 1 °C temperature deviation from thermoneutral indoor conditions leads to the same change in thermal sensation as an increase of 2.6 dBA

(short-term exposure). The authors strongly suggest that this effect on the thermal perception stems from the occupant's perception of an IEQ metric being affected by other environmental components which the authors argue might explain some combined effects of the indoor environmental conditions on cognitive performance. A similar finding is found in the work of Huang et al. [54].

Likewise, Alm et al. found that a 1°C change in operative temperature has the same impact on the human perception of the overall environment as a change of 3.8 dBA in sound pressure level or a change of 7 dp (units of perceived air quality) in air pollution [4].

Nagano et al. [94] examined the combined effect of thermal and acoustic exposures. They concluded that the operative temperature has a slight effect on perceived acoustic comfort, while the auditory condition has a significant effect on the thermal comfort and discomfort sensations.

Geng et al. [50] conducted a controlled IEQ study on 21 subjects to investigate the effects of thermal environment on occupants' IEQ perception. The indoor air temperature varied from 16 °C to 28 °C and the subjective perceptions of thermal comfort, IAQ, lighting and acoustic environment were collected from participants through a questionnaire. The authors found that the thermal environment affected occupant perception of indoor air quality, lighting and acoustic environment. For instance, they found that 70% of participants felt dissatisfied with the overall environment at 16 °C, while the percentage of overall environment dissatisfaction fell down to zero at 24 °C. They argued that this was caused by a comparative impact, where the uncomfortable thermal conditions weakened the occupant comfort expectation of other IEQ factors, which resulted in less dissatisfaction with other IEQ factors.

Gauthier et al. [44] investigated the effect of elevated CO_2 levels on occupants' perception of the thermal environment. The authors argued that the performance gap between modelled and perceived thermal comfort sensation could possibly be filled by including measurement of CO_2 levels in thermal comfort modelling. They conducted two experimental studies, one in a controlled chamber setting and the other one in an office setting. Results showed that participants felt warmer on average when CO_2 levels increased. However, the statistical significance of these findings was modest (p=0.08, for the office experiment, and p=0.48 for the controlled chamber study). The authors suggested that this may be caused by 1) the small sample size of 18 participants in their climate chamber-based experiments, 2) the low range of variation of CO_2 levels in the field study, 3) using the average mean vote of occupants' thermal sensation as a continuous variable. The authors suggested increasing the sample size as well as using logistic regression to model thermal sensation for future studies [44]. Though they found only modest evidence to support their hypothesis on completion of climate chamber-based experiments with 18, the authors attributes this to the relatively small sample size of their study. Similar studies, and to some extent findings, on the correlation between indoor CO_2 levels and indoor air temperature and thermal acceptability are found elsewhere [20], [16].

Chinazzo et al. [22] found that, at elevated indoor air temperature (above 25 °C), changes to measured illuminance levels are correlated with changes to perceived thermal satisfaction in both summer and winter times (p-value=0.027 in summer and 0.001 in winter). Occupants reported lower thermal satisfaction levels when exposed to dim light (lower than 300 lux) compared to people exposed to brighter environments. The evidence was supported by a follow-up study in a

controlled experimental environment [23].

1.2.5 Model Predictive Control in Indoor Thermal Environments

Seeking to reduce the prevailing gap between predictions and actual observations of thermal comfort, several approaches have been adapted in indoor thermal environment control systems, such as MPC [24, 43, 135]. When using model predictive control strategy, a building model is used within the control system to evaluate predictions of thermal comfort levels and control the indoor environmental conditions accordingly [24]. Currently, the most used control strategy for building thermal environment is controlling the indoor air temperature at a static set point, which is usually determined by building managers [81]. However, temperature set point control strategies often fail to provide comfortable thermal environment to occupants [24, 74, 133]. It has been found that about 50% of the controlled temperature set points deviate from the comfortable temperature during cooling seasons, which leads to excessive energy consumption and uncomfortable environmental conditions [133].

To overcome the issues associated with temperature set point-based control, several researchers have adapted MPC into indoor thermal environment control systems [14, 24, 74, 81]. It has been found that MPC can reduce building energy use while maintaining occupant's thermal comfort levels [81]. Although some prior works used the PMV model, developed by Fanger [38], to control the buildings' thermal environment, the PMV model-based control fails to reflect the real-time thermal sensation into the the building's thermal environment control system. Besides, discrepancies are often found between actual thermal comfort and the thermal comfort calculated by the PMV model [56, 123].

The development of machine learning techniques has contributed to the introduction of more measurements of indoor environments into building control systems. For instance, recent buildings have moved away from adapting temperature set point-based control system only and added other environmental factors, such as measurement of CO_2 levels [36, 85]. Measurements of CO_2 concentrations have been commonly included in building control systems as a proxy variable for indoor air quality as well as the presence of occupants [36]. An occupancy-based building control system that uses CO_2 sensors with occupancy-based ventilation control can save up to 24% of building energy- use without compromising occupants' thermal comfort [71]. Further, combining CO_2 sensors with occupancy-based ventilation control might save about 55.8% of outside air ventilation power [113]. In a prior experimental work, Kolokotsa et al. [70] developed a *thermal-indoor air quality* (TIAQ) controller, programmed in MATLAB, which controls the thermal environmental conditions and the indoor air quality of the building. The building control model inputs are: indoor air temperature, mean radiant temperature, indoor relative humidity, indoor air velocity, outdoor temperature (to detect whether it is a cooling or heating season), and CO_2 concentrations. The outputs of the model are: the heating or cooling requirements and the window opening (fully open, slightly open, closed). Using their model, they found that a 20% reduction in building energy use can be achieved compared to conventional building control system [69].

1.3 Research Objectives, Novelty, and Thesis Outline

As has been stated in the previous section, a prevailing research challenge with respect to standard models of thermal comfort is that seminal data-informed models have not always accurately predicted true thermal comfort observations from inde-

pendent field studies. Though some prior evidence, reviewed in 1.2, has suggested that there may be correlations between perceived thermal comfort and non-thermal metrics of IEQ, quantifiable evidence of such relationships being drawn from large field studies is still fleeting.

Further, across the reviewed literature, which include the works reviewed in 1.2 and over 200 works surveyed by Schweiker et al.[110], there are no prior field studies, of a medium or large-scale greater than 50 test subjects or occupants, that have sought to quantify, with uncertainty bounds, the correlations between perceived thermal comfort and non-thermal metrics of IEQ and evaluate whether these correlations are statistically significant. Given the challenges of conducting research including real occupants in realistic settings, most existing studies have been conducted with a limited number of participants, whether in a field setting or climate chamber. Moreover, the participants, often young students, have not been necessarily representative of pertinent populations, for instance, of office workers across all ages and genders [110].

This work is the first known work to use Bayesian inference techniques, state-of-the-art IEQ sensors to quantify, with uncertainty bounds, the correlations between occupants' perceived thermal comfort and non-thermal metrics of IEQ. This thesis presents a novel methodology that applies Bayesian statistical methods on a large multi-domain IEQ dataset and that is able to evaluate the statistical robustness and significance of the relationships between perceived thermal comfort and several metrics of IEQ.

This is the first known work to examine whether the predictions of occupant's thermal comfort can be improved upon adding measurements of non-thermal metrics of IEQ and in a manner that supports future thermal comfort prediction. This

thesis then aims to take advantage of the emerging awareness in research of the inter-dependencies between perceived thermal comfort and overall IEQ, and will evaluate an approach to quantitatively predicting perceived thermal comfort as a function of both thermal and non-thermal metrics of IEQ. With these goals in mind, this PhD thesis objectives are as follows:

1. Investigate whether thermal comfort, as perceived by occupants in a field dataset, can be attributed in a measurable and/or significant manner to one or several non-thermal parameters of IEQ.

2. Test the significance and statistical robustness of the correlations between perceived thermal satisfaction and non-thermal IEQ metrics, if any.

3. Formulate a new predictive model of thermal comfort, derived from the Bayesian logistic regression of 2 large IEQ field datasets, which can be used by building experts to improve the prediction accuracy of thermal comfort in office spaces.

4. Investigate whether the prediction accuracy of thermal satisfaction can be improved upon the addition of measurements of non-thermal metrics of IEQ

5. Integrate the new predictive model of thermal comfort into building control systems to predict thermal comfort in office spaces based on thermal conditions and ventilation rates to increase fresh air amounts while maintaining the same levels of thermal comfort and with minimal energy demand increase.

This PhD thesis consists of six main chapters. The first chapter (this chapter) provides relevant background information, identifies the thesis' research questions,

reviews the literature on several relevant topics, highlights the research novelty, and identifies the overall objectives of this research.

In the second chapter, a new technique for establishing empirical correlations between IEQ survey response data and measurable parameters of IEQ using a Bayesian logistic regression framework is developed. The first case study of this research is presented and a prior large IEQ field study of office spaces in Canada and the United States, known as the Cost-effective Open Plan Environments (COPE) dataset, is utilized for the chapter's core analysis. Posterior results of the Bayesian inference processing are used to deduce correlations between perceived occupants' thermal satisfaction and several non-thermal parameters of IEQ, such as CO_2 concentrations, noise levels, and light levels. Posterior checks and model comparison approaches are performed to the Bayesian models to investigate potential improvements on prediction accuracy of occupants' thermal satisfaction

The third chapter presents an update to the work presented in chapter 2 by, first, reformulating the model as a *hierarchical* Bayesian logistic regression problem, second, by adding approximately 150 new samples of subjective and objective IEQ measurements collected from a new field IEQ study carried out at the University of British Columbia (UBC).

Using the UBC dataset as the second case study of this research, the third chapter seeks to address whether, on the addition of new field data and applying a more comprehensive Bayesian inference technique, the evidence of correlations between thermal satisfaction and non-thermal metrics of IEQ is strengthened - or weakened. The experimental apparatus and full methodology of the new field study is described as well the revised methodology for hierarchical inference of the logistic regression model. Results from model fitness and validation techniques, performed

on the candidate models, are presented. Further, the results from the correlation analysis of the subjective and objective data collected in COPE and UBC studies are presented.

In the fourth chapter, the new predictive model of thermal comfort, derived from the hierarchical Bayesian regression of the combined COPE and UBC datasets is adapted and integrated into a building energy model (BEM). The model predicts occupants' thermal satisfaction and heating energy consumption as a function of setpoint thermal conditions and indoor CO_2 concentrations. An open-plan mechanically-ventilated office space located in Vancouver is simulated in Transient System Simulation Tool (TRNSYS). Different configurations of both the heating setpoint and the indoor CO_2 setpoint are examined. The corresponding heating demand and thermal satisfaction are calculated for each examined scenario in order to examine possible energy-savings scenarios to increase the ventilation rates while not compromising the occupant's thermal comfort.

In the fifth chapter, several occupancy profiles are developed and examined to reflect and compare the current occupancy schedules in office spaces after the COVID-19 pandemic. Different configurations of both the heating setpoint and the indoor CO_2 setpoint are examined. The corresponding heating energy demands and thermal satisfaction are calculated for each investigated occupancy profile for a combination of indoor air temperature and indoor CO_2 levels setpoints. The daily value of heating energy demand, air change rate and ventilation rate for different scenarios of heating and CO_2 setpoints are evaluated.

The sixth, and final, chapter summarizes the conclusions of this research, highlights the key research contributions, provides recommendations for implementation of the current results, and discusses future research work.

Chapter 2

Investigating Relationships Between Thermal Comfort and IEQ: First Case Study

2.1 Introduction

As has been stated in section 1.2, research has shown that, in several instances, it has been possible to find discrepancies between measurements of perceived thermal comfort and predictions of thermal comfort. One prior study argued that this prediction gap may be filled by including new parameters in thermal comfort models [44]. Though some prior evidence, reviewed in 1.2, has suggested that there may be quantifiable dependencies between perceived thermal comfort and non-thermal metrics of IEQ, universally-applicable evidence of such relationship is still fleeting.

It is then the purpose of this chapter to investigate whether, based on a moderately large field study (approx. 800 subjects), the relationship between perceived thermal satisfaction and non-thermal IEQ appears to be significant, independent, and statistically robust. In other words, this chapter aims to inform whether occupants' perceived thermal comfort can be attributed, in a measurable and/or significant manner, to one or several non-thermal IEQ parameters typically associated with indoor 'well-being', such as lighting intensity, acoustic performance, and indoor air quality. More specifically, the purpose of this chapter is to investigate whether: a) the observed relationship between perceived thermal satisfaction and non-thermal metrics of IEQ can be represented quantitatively, and if so, b) whether, based on a moderately large field study, the relationship between perceived thermal satisfaction and non-thermal IEQ appears to be significant and statistically robust.

To undertake this work, a novel technique for establishing empirical correlations between IEQ survey response data and measurable parameters of indoor IEQ using a Bayesian logistic regression framework is developed. A prior large IEQ field study of office spaces in Canada and the United States, known as the Cost-effective Open Plan Environments (COPE) dataset, is utilized for this chapter's core analysis. The field IEQ database consists of physical measurements of IEQ as well as subjective data gathered from IEQ satisfaction questionnaire collected from 779 workstations in large Canadian and US cities. A Bayesian framework is developed and posterior results of the Bayesian inference processing are used to deduce correlations between perceived occupants' thermal satisfaction/dissatisfaction and several non-thermal parameters of IEQ, such as CO_2 concentrations, noise levels, and light levels. Posterior checks and model comparison approaches are performed to the Bayesian models to investigate potential improvements on

prediction accuracy of occupants' thermal satisfaction.

The types of IEQ measurements contained in the COPE dataset were more comprehensive than what is found in more well-known datasets such as RP-884 [33] and the ASHRAE Global Thermal Comfort Database II [42]. The COPE study measured ambient CO_2 concentrations at each sampled workstation, ambient noise levels, speech intelligibility, light levels, and several other metrics. The COPE study aimed to determine the effects of different open-plan office design aspects on the indoor environment and on occupant satisfaction with that environment [95]. The study also examined correlations between occupants' environmental and job satisfaction [129] as well as correlations between physical conditions in the workplace and occupants' environmental satisfaction [87, 95, 96].

The questionnaire issued to office workers in the COPE study also differed from established IEQ questionnaires. For example, many of the IEQ questionnaires issued in facilitation of the ASHRAE/RP-884 thermal comfort databases, including the CBE thermal comfort survey [116], pose several questions to occupants in order to characterize overall thermal comfort. They typically ask occupants to rate their thermal sensation, their thermal preference, and/or their thermal acceptability. The explicit questioning of thermal satisfaction, as ranked on a satisfaction scale (i.e., very unsatisfied, unsatisfied, etc.), is not common. Yet, in the COPE study, the survey asked only the following question of its participants in direct relation to thermal comfort: "What is your degree of satisfaction with the temperature in your workspace, right now?". It was measured on a 7-point satisfaction scale.

2.2 A Bayesian Framework for Thermal Comfort under Thermal and Non-thermal IEQ Criteria

Bayesian methods of inference or regression are not novel. Previous works, presented in section 1.2 of the previous chapter, are not the only applications of Bayesian statistics to building science problems, there are several more in the area of building performance simulation [53]. There are therefore two considerations of this research work that make it a novel contribution to the empirical evaluation of thermal comfort. Firstly, Bayesian logistic regression is undertaken on a large field study. Second, the work applies several model checking and statistical significance checks to identify the independence, and potentially the universality of observing a quantifiable effect of non-thermal IEQ metrics on perceived thermal comfort.

The method proposed in this thesis seeks to quantify the effect of non-thermal IEQ conditions, such as indoor lighting levels, and CO_2 levels, on metrics of perceived thermal comfort using Bayesian methods of statistical inference. Specifically, a Bayesian logistic regression model is configured in order to estimate the probability of an occupant feeling thermally dissatisfied, $p(D)$, and thermally satisfied, $p(S)$, as a function of not only to several psychrometric / thermal IEQ parameters $\mathbf{F}$, but also to several non-thermal well-being IEQ parameters defined by a separate set of terms, $\mathbf{W}$.

The relevant thermal parameters for this study, $\mathbf{F}$, are specifically: T = air temperature (ºC), R = relative humidity (%), M = mean radiant temperature (ºC), and V = air velocity (m/s), such that $\{T, R, M, V\} \in \mathbf{F}$. The measurable non-thermal well-being IEQ metrics used in this study, defined by the set $\{C, N, E, L\} \in \mathbf{W}$, are: C = indoor air CO_2 levels (ppm), N = A-weighted indoor noise levels (dBA), E =

speech intelligibility index, L = measured desktop illuminance (lux). The rationale for choosing these four non-thermal parameters is dictated by the dataset made available for this study. This is further discussed in the next section of this chapter.

A Bayesian network is a graphical representation of uncertain quantities that reveal the probabilistic relationship between a set of variables. The Bayesian network diagram proposed for this work is shown in Figure 2.1, where posterior predictions of thermal dissatisfaction $p(D \mid \mathbf{F}, \mathbf{W})$ can be determined as follows:

The nodes represent the random variables and the connecting lines represent the probabilistic dependence between the nodes. The variable, or node, that causes another node is called a "parent", and the affected node is called its "child". If B is a parent and A is a child of B, the probability of A conditioned by B is noted $P(A \mid B)$. Bayes theorem describes probabilistic dependencies between A and B as follows:

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)} \tag{2.1}$$

$$p(D \mid \mathbf{F}, \mathbf{W}) \propto p(\mathbf{W} \mid D) \cdot \; p(\mathbf{F} \mid D) \cdot \; p(\mathbf{D}) \tag{2.2}$$

such that:

$$p(D \mid \mathbf{F}, \mathbf{W}) = \frac{p(\mathbf{W} \mid D) \cdot \; p(\mathbf{F} \mid D) \cdot \; p(\mathbf{D})}{p(\mathbf{F}, \mathbf{W})} \tag{2.3}$$

For the equation above, $p(D \mid \mathbf{F}, \mathbf{W})$ is the probabilistic model of perceived thermal dissatisfaction given the effect of thermal and non-thermal well-being IEQ parameters. $p(\mathbf{F} \mid D)$ is the likelihood of observing the well-being IEQ param-

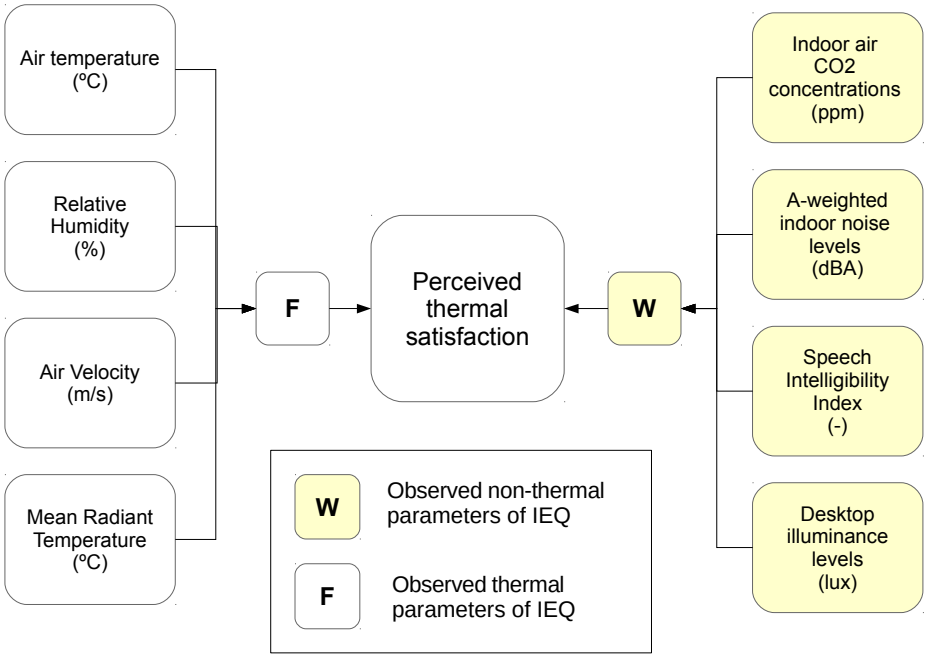

Figure 2.1: A Bayesian Network for the proposed thermal satisfaction modelling framework which incorporates both thermal and non-thermal parameters of indoor environmental quality

eters, $\mathbf{W}$ given the observed thermal dissatisfaction, D. For the equation above, $p(D \mid \mathbf{F}, \mathbf{W})$ represents the probabilistic model of perceived thermal dissatisfaction given thermal and non-thermal IEQ metrics. It is assumed throughout the Bayesian modelling process that the non-thermal parameters, $\mathbf{W}$, are conditionally independent on one another.

The rationale behind choosing Bayesian processing as the modelling choice is related to its inherent advantages: its ability to incorporate different field datasets from heterogeneous sources into one thermal comfort model while updating our

prior knowledge on thermal comfort distributions from past research; it can easily account for unobserved variables [79].

2.3 First Case Study: Field Data from the Cost-effective Open-Plan Environment (COPE) Study

Field IEQ data for this analysis is drawn from the Cost-effective Open-Plan Environment (COPE) field study database made available for this research by the National Research Council of Canada (NRC). The COPE database consists of IEQ data collected from 779 workstations and their occupants in nine buildings between 2000 and 2002 in large Canadian and US cities [96]. The field database contains instantaneous measurements of IEQ conditions at individual workstations coupled with responses from a 'right-here-right-now' IEQ questionnaire of each workstation's occupant. The IEQ parameters covered by the COPE dataset consist of four physical measurements of thermal conditions: temperature, relative humidity, mean radiant temperature, and air velocity, and additionally include 12 measurements of non-thermal conditions, such as noise levels, desktop illuminance, and CO_2 concentrations.

A summary of all the buildings covered by the COPE dataset is laid out in Table 2.1. Data was collected from large cities in Canada and the US belonging to four different Köppen climate classification groups. The first three of the nine buildings surveyed were occupied by federal government organizations in large Canadian cities, with field measurements undertaken in 2000. In 2002, data from four private-sector office buildings, and two provincial government office buildings was added to the dataset. Three of nine buildings were in large Canadian cities, and the remaining buildings were located in two large US cities.

Building	City (Köppen climate)	Sector	Date of Visit	#Floors	# Respondents
1	Ottawa (Dfb)	public	Spring 2000	11 (4 visited)	131
2	Toronto (Dfb)	public	Summer 2000	12 (3 visited)	127
3	Ottawa (Dfb)	public	Spring 2000 + Winter 2000	22 (4 visited)	160
4	Ottawa (Dfb)	private	Winter 2002	15 (1 visited)	52
5	San Rafael (BSk)	private	Spring 2002	3 (3 visited)	85
6	San Rafael (BSk)	private	Spring 2002	5 (1 visited)	48
7	San Francisco (Csb)	private	Spring 2002	8 (1 visited)	72
8	Montreal (Dfa)	public	Spring 2002	4 (2 visited)	47
9	Quebec City (Dfb)	public	Spring2002	3 (3 visited)	56

Table 2.1: Summary of the buildings studied in the 'COPE' database

The following non-thermal measurements, drawn from the COPE dataset, are used in this study to characterize conditions of air quality, lighting, acoustics, and interior design. These are defined by the set $\mathbf{W} = \{C, N, E, L\}$. The non-thermal parameters under study, defined by the set $\mathbf{W}$, are selected based on their availability in the COPE field dataset and on their definition lying under the umbrella of typical 'well-being'-type non-thermal IEQ parameters [119]. Perceived thermal satisfaction and dissatisfaction was also measured by the COPE field study. The dataset captured occupant responses to the question of thermal satisfaction, e.g., "How satisfied are you with the temperature at your workspace right now?". The answer to this question was measured on a Likert scale from 1, 'very dissatisfied', to 7 ' very satisfied'. Proportions of thermal satisfaction responses received, denoted by TS, for each building in the COPE database are shown in Figure 2.2.

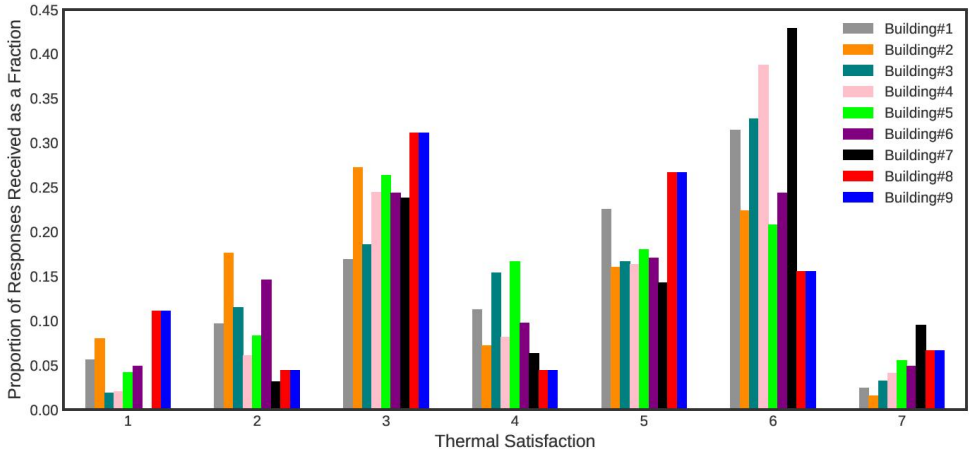

Figure 2.2: Proportion of thermal satisfaction responses received as a fraction across all buildings

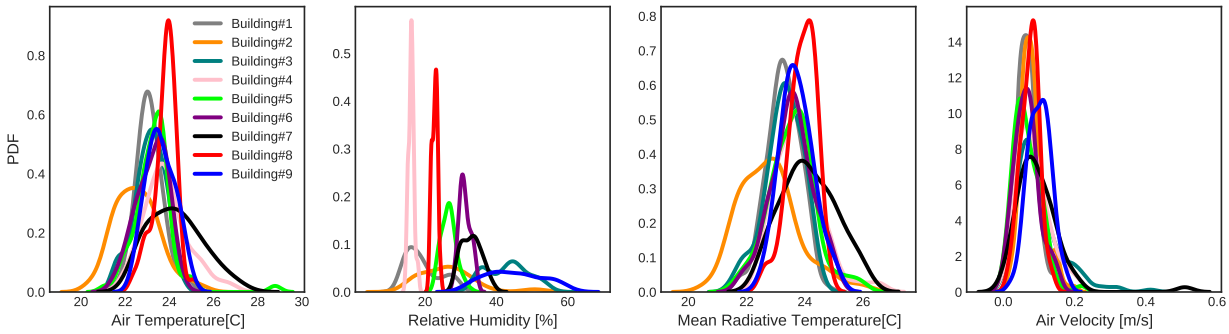

Figure 2.3: Probability distributions of IEQ thermal parameters (**F**) across all buildings

Figure 2.3 and Figure 2.4 illustrate data for all the thermal, **F**, and non-thermal, **W**, IEQ parameters, as derived from the COPE dataset. Probability Density Function (PDF) for each metric per building in the COPE dataset are evaluated and presented in Figure 2.3 and Figure 2.4. The probability density functions of each

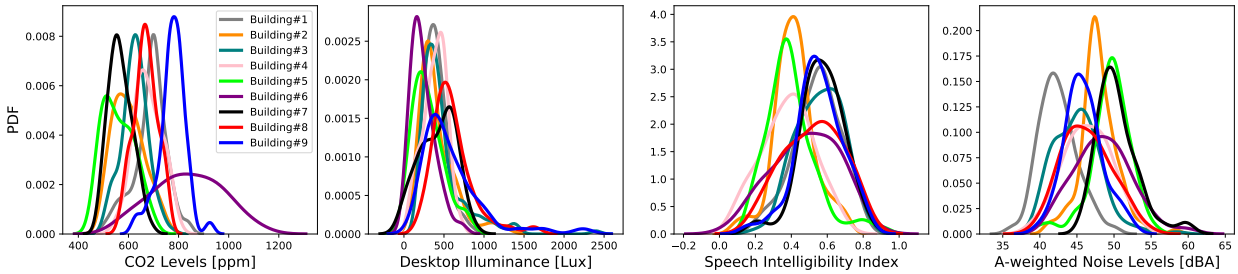

Figure 2.4: Probability distributions of IEQ non-thermal well-being (**W**) parameters across all buildings

parameter are only generated for the purposes of comparison in Figure 2.3 and Figure 2.4 and are not used further in this study.

2.3.1 Description of Bayesian Logistic Regression Model

Bayesian regression is an approach to regression analysis that treats a regression model's correlation coefficients as uncertain and probabilistic. It's viewed that regression models established by Bayesian inference can more appropriately express the inherent uncertainty of a regression model's fit, particularly in contexts where observable data to establish a model is sparse [98]. For this work, Bayesian logistic regression model is developed to represent eq.2.3 against observations found in the COPE database.

The model predicts the logistic relationship between the likelihood of an occupant feeling thermal dissatisfied or satisfied (D or S) and both measured thermal conditions (**F**) and measured non-thermal, well-being, IEQ parameters (**W**) drawn from the dataset. Establishing β as a set of regression model coefficients, $p(D|\mathbf{F},\mathbf{W})$ can be re-written as:

$$p(D \mid \mathbf{F},\ \mathbf{W}) = \frac{1}{1 + e^{-}[\Sigma_{\mathbf{F}}\ (\beta_{\mathbf{F},D}\ .\ \mathbf{F}) + \Sigma_{\mathbf{W}}\ (\beta_{\mathbf{W},D}\ .\mathbf{W}) + \beta]} \tag{2.4}$$

The probability of thermal dissatisfaction, $p(D)$, is inferred from the Bayesian model by modelling the thermal dissatisfaction as a dichotomous dependent variable and the thermal **F** and non thermal **W** parameters are modelled as continuous independent variables. Observed data for $p(D)$ is inferred from the COPE dataset by assuming that for each survey response, i:

$$D_i = \begin{cases} 1\ (\text{or dissatisfied})\ if\ TS < 4 \\ 0\ (\text{or satisfied})\ if\ TS > 4 \end{cases} \tag{2.5}$$

Equations (2.4) can also solve for $p(D \mid \mathbf{F},\ \mathbf{W})$ by replacing D for S. Similarly, the probability of satisfaction, $p(S)$, is modelled as a dichotomous dependent variable and observed data for $p(S)$ is inferred from the COPE dataset by assuming that for each survey response, i:

$$S_i = 1 - D_i \tag{2.6}$$

Neutral responses made by occupants with regards to thermal satisfaction (i.e., $TS = 4$) are ultimately excluded as they are neither clear indicators of thermal satisfaction or dissatisfaction.

2.3.2 Description of Candidate Models

As this thesis seeks to determine whether, and to what extent, measurements of non-thermal parameters are correlated with predictions of indoor thermal comfort, there is no single proposed model that should be explored, but several. In the

Bayesian framework, it is effective to evaluate whether the addition of new parameters to a model, or a change in parameters of a model, improves the prediction accuracy of the model against another model, and also against an original Null hypothesis. The methodology of selecting the Null hypothesis for this study is explained in detail in the following section.

An ensemble of models are created, each represented by a different combination of thermal, **F**, and non-thermal, **W**, variables of IEQ. They are then evaluated and ranked with respect to their relative performance and performance against the Null hypothesis. Table 2.2 lays out a total of 8 different cases of the candidate models. Each case is used to define a logistic regression model that seeks to relate the probability of an occupant feeling thermally satisfied or dissatisfied to **F** and **W** parameters. For example, case 1 establishes the probability of an occupant feeling thermally satisfied or dissatisfied as a function of indoor air temperature, and mean radiant temperature. The respective long-form notation of the models attributed to this case would be respectively $p(S \mid T, M)$ and $p(D \mid T, M)$.

In all, a total of 16 different models are established out of the 8 cases of Table 2.2, as each case refers to a model of $p(S)$ and $p(D)$. It should be noted that the cases chosen for this work reflect the outcome of some trial and error analysis that was undertaken to determine candidate cases of reasonable interest. It is by no means an exhaustive list of all possible models of thermal satisfaction and dissatisfaction that can be created out of the COPE dataset.

Based on initial trial and error undertaken, the following model coefficients are modelled as having a first order linear relationship with $p(D)$ and $p(S)$: C, N, V, and R. The following parameters are modelled as having a quadratic relationship with $p(D)$ and $p(S)$: T, M, E and L. The choice of these relationships were made

Case	Parameters of **F**	Parameters of **W**	Models
1	T,M	-	$p(S \mid T,M)$ and $p(D \mid T,M)$
2	T,M,R	-	$p(S \mid T,M,R)$ and $p(D \mid T,M,R)$
3	T,M,R,V	-	$p(S \mid T,M,R,V)$ and $p(D \mid T,M,R,V)$
4	T,M	C	$p(S \mid T,M,C)$ and $p(D \mid T,M,C)$
5	T,M	E	$p(S \mid T,M,E)$ and $p(D \mid T,M,E)$
6	T,M	N	$p(S \mid T,M,N)$ and $p(D \mid T,M,N)$
7	T,M	L	$p(S \mid T,M,L)$ and $p(D \mid T,M,L)$
8	T,M	C,E,N,L	$p(S \mid T,M,C,E,N,L)$ and $p(D \mid T,M,C,E,N,L)$

Table 2.2: List of cases evaluated for generating models of $p(S)$ and $p(D)$

on a trial-and-error basis and evaluation of aforementioned model selection criteria.

2.3.3 Model Comparison and Evaluation Criteria

When determining the significance of conclusions drawn from statistical analyses, evaluation criteria for Bayesian inference differs from *Frequentist* inference in both terminology and approach. In Bayesian inference, and particularly for Bayesian regression models, metrics that evaluate the predictive accuracy of a proposed model are used to compare individual models against one other in the process of determining the models of best fit to observable data [6, 125].

Previous studies in the field of building science and thermal comfort made use of several different model comparison approaches to evaluate their results. For instance, Karava et al. [64] developed a Bayesian model that is used to identify different building groups with similar thermal characteristics. They used the Watanabe-Akaike Information Criteria (WAIC) scores and Leave-One-Out Cross-Validation (LOO-CV) techniques to compare between the prediction accuracy of different models under study. In a previous study by Kirstensen et al. [73], the authors developed a Bayesian model that predicted annual heating energy consumption

for Danish buildings using building characteristics data. The authors used WAIC scores to compare the expected predictive accuracy of their candidate Bayesian models and select the best performing model according to the WAIC score results. These two methods, along with the Odds Ratio analysis, are described below.

Watanabe-Akaike Information Criteria

The widely applicable, or Watanabe-Akaike, information criterion (WAIC) method [134] is a Bayesian validation approach that evaluates the likely fitness of a model to known and unknown data. The WAIC method does this by estimating the out-of sample prediction error of a candidate mathematical model - in this case the proposed logistic regression model. The WAIC method, which resolves itself to calculating a 'score' of a model, takes into account in-sample accuracy (i.e., an evaluation of how well a model fits observed data) and out-of-sample prediction (i.e., how well a model can predict unobserved / future data) [47, 127].

The WAIC score of an individual model is a probabilistic metric and so there is nominally two parameters of interest when using a WAIC score to compare one model against another: the mean score, and the score's error. A model is viewed to be a better fit to data than another if its mean WAIC score is lower. However, such a claim should only be considered robust if the error of calculated WAIC scores is not larger than the difference between the mean scores themselves. Further explanation is provided in section 2.3.3.

Leave-one-out Cross Validation

Leave-One-Out Cross-Validation (LOO-CV) is another established approach for comparison and selection of Bayesian models. It is usually performed for com-

paring the predictive accuracy of the models. LOO-CV estimates the out of sample predictive fit by repeatedly partitioning the data into training and hold-out sets, iteratively fitting the model with the training dataset and evaluating the fit with the held-out data [52, 104, 127]. The out-of-sample prediction accuracy from a fitted Bayesian model using the log-likelihood evaluated at the posterior simulations of the parameter values.

While WAIC and LOO rank models by evaluating the out-of-sample predictive accuracy, their computation methodology are slightly different [127]. It has been demonstrated that LOO-CV often outperforms WAIC. Although WAIC is asymptotically equal to LOO, it has been found that LOO is more robust in the finite case with weak priors [126, 127]. To compare the models under an arbitrary study, LOO should be evaluated for each model and we should prefer the model(s) with lower LOO.

Judging Model fitness Based on WAIC and LOO Scores

The current state-of-practice for estimating WAIC and LOO scores of arbitrary models generates two sets of values: mean predicted values of WAIC and LOO, respectively *WAIC* and *LOO*, and the standard error of each prediction, respectively $WAIC_{error}$ and LOO_{error}. Applying WAIC and LOO scores to judge models remains a quasi-quantitative or relative process; there is no hard and fast rule to determining model fitness based on specific values of these scores. However, the following approach is accepted [86]. If one model has a lower *WAIC* (or *LOO*) than another, such that $WAIC_1 < WAIC_2$, one can suggest there is some evidence, either weak or moderate, in support of model 1 over model 2. If the WAIC score of this preferred model is so low such that $(WAIC_1 + WAIC_{error,1}) \leq$

$(WAIC_2 - WAIC_{error,2})$, then the evidence in support of model 1 is strong. The same applies for *LOO*.

The relationship between WAIC/LOO mean scores, errors, and strength of a model's evidence base are deemed important and unique considerations of this work's methodology. The aforementioned building science studies of Kirstensen et al. [73] and Karava et al. [64] drew conclusions on model fitness on only the mean estimated WAIC/LOO scores, not their error.

Other Methods of Model Comparison

Visualization and analysis of odds ratios is another model validation approach used particularly in Bayesian logistic regression in order to measure the effect of predictor variables on binary responses. It is used to validate the correlations between each independent variable and the posterior outcome of the model [25]. The odds ratio of $A \mid B$ (i.e., A given B) is the ratio between the odds of dependent variable A in the presence of independent variable B and the odds of A in the absence of B. Therefore, the odds ratio is used to test the sensitivity of a model to its independent variables, by evaluating the ratio between the odds of the dependent variable with and without the independent parameter under test. In logistic regression, the odds ratio represents the constant effect of an independent model parameter on the likelihood that an outcome will occur [19]. Since the logistic regression model uses a logistic function to estimate the relationship between a binary dependent variable and a group of predictor variables, the log of the odds ratio is often an easier approach to interpret. If the odds ratio of an independent model variable is greater 1, it signifies that, all other variables in being equal, an increase in the model variable's value results in an increase in the model's outcome value.

Another way of evaluating fitted models is performing Predictive Posterior Checks (PPC), which is a useful direct way to verify whether, given observed data, the predictions made by a chosen model are ultimately sensible [48]. Posterior predictive checks are performed by generating replicated data using draws from the model's posteriors and then comparing these to the observed data. The model is then validated by checking for alignment between real and simulated data, i.e between observed data and what we expect to be observed under the posterior distribution [49].

2.4 Results

The investigation of results seeks to inform whether, insofar as the COPE dataset is concerned, non-thermal parameters of IEQ influence perceived thermal satisfaction and whether including some of these non-thermal IEQ parameters is useful when generating an empirical model of predicted thermal satisfaction.

The Bayesian statistics Python library, PyMC3, is used to infer posterior distributions of logistic regression model coefficients for all of the 16 investigated models. For each model, 5000 samples are drawn from the posteriors using the Sequential Monte Carlo (SMC) method, a type of Markov Chain Monte Carlo (MCMC) sampling method. The Google Collaboratory (COLAB) interactive Python environment is used to run all simulations in this work. Weakly informative priors for each of the models' regression parameters β are used, as recommended by Gelman et al.[46]. It's to be noted noted that since building 6 is observed to be outlier in regards to the CO_2 distribution, as shown in Figure2.4, it has been removed from the analysis. Therefore, the following results reflect a regression of 731 out of the 779 samples contained in the COPE dataset.

2.4.1 Initial results

The Bayesian logistic regression process seeks to determine the extent to which the likelihood of a binary variable, such as an occupant feeling thermally satisfied or not, is a function of one or several continuous variables. An illustration of the overall scope of results is found in Figure 2.5. It represents a visualization of the regression result of case 8 of Table 2.2. The Figure displays predicted correlations between thermal dissatisfaction, $p(D)$, and thermal satisfaction, $p(S)$, against individual non-thermal parameters of **W** assuming a fixed set of values for all other parameters. Unless otherwise specified, the fixed set of parameters are: T=23.5 °C, R=30%, M=23.5 °C, V=0.08 m/s, C = 600 ppm, N = 46 dBA, E = 0.5, and L = 450 lux. These are the observed mean values for each parameter, taken at the moment of the survey, out of the COPE data set. As the regression coefficients of the model, β, are probabilistic, there is a range of fit of the models. In Figure 2.5, the mean predictive value of $p(D)$ and $p(S)$ are denoted by a solid red line. The standard error of predictions is denoted by the upper and lower dotted red lines around the mean.

These initial results might suggest a potential relationship between surveyed thermal dissatisfaction, surveyed thermal satisfaction, and several non-thermal **W** IEQ parameters. For instance, it is illustrated that indoor CO_2 levels and thermal dissatisfaction, $p(D)$, may be positively correlated. In support of this, the results also support a view that $p(S)$ is negatively correlated to CO_2 concentrations. Furthermore, It is shown in Figure 2.5 that speech intelligibility and thermal dissatisfaction, $p(D)$, are negatively correlated.

The areas of highest fit of the regression model results correspond with the

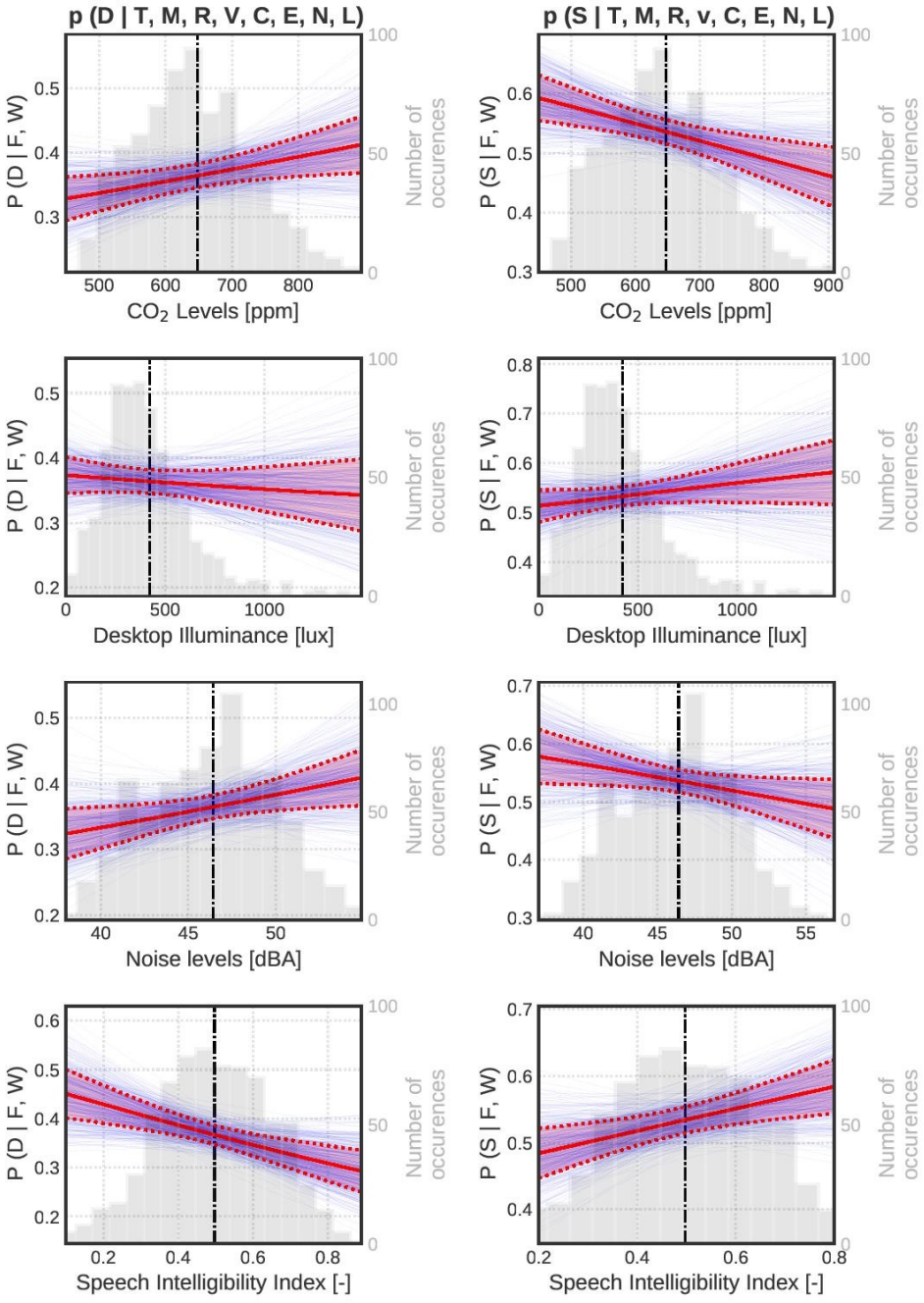

Figure 2.5: Probability $p(D|\mathbf{F},\mathbf{W})$ and $p(S|\mathbf{F},\mathbf{W})$, with thin blue lines indicating individual sample traces, solid red lines indicate mean predicted value from all traces, dashed red bands indicate the standard error of traces, grey bars indicate the probability distribution of each independent parameter as observed in the COPE dataset, and black dashed centre lines are the mean values of observations

number of occurrences of each non-thermal **W** parameter within the COPE dataset, denoted by histogram in grey in Figure 2.5. Mean observed values of each **W** parameter are denoted by vertical, dotted black lines. Model checks, comparison, and selection are undertaken to identify models of best-fit to the COPE dataset observations, and from these best-fit models one can make informed observations from the data.

2.4.2 Model Checks, Comparison, and Selection

Visualization of Odds Ratios

The log odds ratio for each of the non-thermal IEQ parameters, **W**, regression parameters for all $p(D)$ and $p(S)$ models are produced and illustrated in Figure 2.6. The plotted odds ratios suggest that at least two out of the four surveyed non-thermal IEQ parameters may have an attributable independent relationship with predicted thermal dissatisfaction and satisfaction. It is perceived that variables which are more likely to affect a logistic regression model in a statistically significant manner are those with odds ratios which more credibly deviate from 1; anecdotally, this occurs where $\left(p(\beta) < / > 1\right) > 0.95$ if not higher. The odds ratio of the posterior traces of the β_E comply to this for $p(S)$ and a lesser extent for $p(D)$. This is also the case for β_C with respect to $p(S)$. This must not be interpreted as a judgment on parameter acceptability at this stage but more an initial indicator of the likely sensitivity of model results to certain variables. The difference between plots of $p(S)$ and $p(D)$ might be caused by the availability of more 'S' data (i.e. '$TS > 4$' data) as shown in Figure 2.2. Further interpretation of the odds ratio results is discussed further in section 2.5.2.

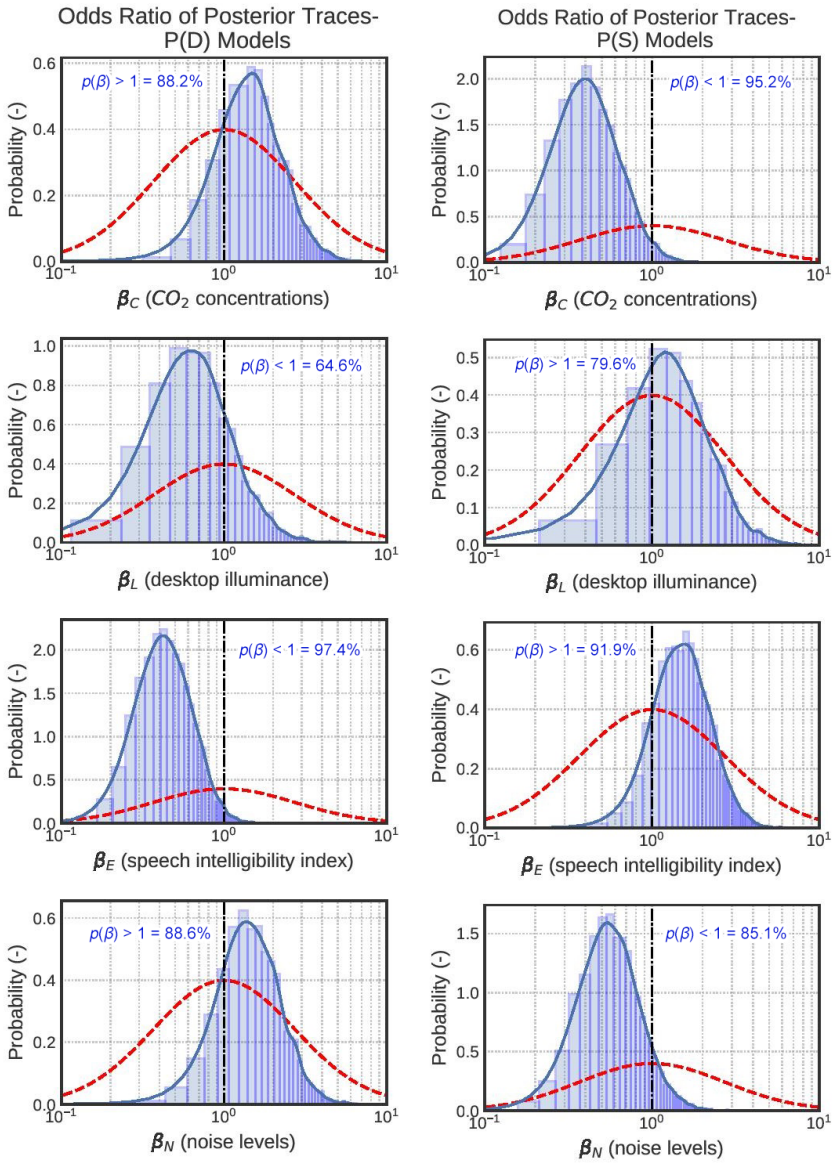

Figure 2.6: Odds Ratio of posterior traces of non-thermal (**W**) IEQ parameters for $p(D)$ (on the left) and $p(S)$ (on the right) Bayesian Models with prior distributions displayed in red

Quantitative Model Comparison and Selection

Two different model comparison approaches are performed to evaluate each of the 16 regression models developed: the Watanabe-Akaike Information Criteria (WAIC) and Leave-one-out Cross Validation (LOO-CV). A Null hypothesis is selected for this process to establish the comparison. The candidate models for $p(D)$ and $p(S)$ that most closely fit the observed data but only include thermal IEQ parameters (i.e. observed parameters of the Fanger PMV-PPD model) are chosen. It should be noted that the choice of the thermal IEQ parameters included in this work is dependent on their availability in the COPE dataset.

This first step reveals that the best representation of $p(S)$ and $p(D)$ as a function of only thermal parameters were the models that including only indoor air temperature, T, and mean radiant temperature, M i.e. $p(S|T,M)$ and $p(D|T,M)$. Tables 2.3 and 2.4 substantiate this to the extent that WAIC and LOO scores for these models is lower than alternatives which include parameters of V, and R. The implications that relative humidity and airflow velocity measurements potentially reduce predictive accuracy of $p(D)$ and $p(S)$ is discussed further in section 2.5.1.

Scores of WAIC and LOO

The WAIC and LOO scores for each of the developed Bayesian models for both $p(S)$ and $p(D)$ models are calculated and the results are summarized in Tables 2.3 and 2.4. We recall that, when comparing the fit and appropriateness of two different models to a dataset, models with lower WAIC and LOO scores are suggested to be a better fit to data. However, the estimated standard errors of the WAIC and LOO

Bayesian Model	WAIC Score	WAIC Standard error	LOO Score	LOO Standard error
$p(S \mid T,M,C)$	999.65	4.26	999.65	4.26
$p(S \mid T,M,E)$	1000.26	4.42	1000.26	4.42
$p(S \mid T,M,C,E,L,N)$	1000.36	5.7	1000.36	5.7
$p(S \mid T,M)$	1001.47	3.06	1001.48	3.06
$p(S \mid T,M,L)$	1001.54	3.53	1001.54	3.53
$p(S \mid T,M,N)$	1002.25	3.66	1002.25	3.66
$p(S \mid T,M,R,V)$	1002.53	3.61	1002.53	3.62
$p(S \mid T,M,R)$	1003.48	2.99	1003.48	3.00

Table 2.3: Scores of WAIC and LOO-CV for the $p(S)$ models, with Null hypothesis shown in red

scores must also be considered in this assessment, and ultimately one model can be considered a robust improvement over another if its WAIC and LOO scores are lower than the Null hypothesis inclusive of their estimated errors.

Bayesian Model	WAIC Score	WAIC Standard error	LOO Score	LOO Standard error
$p(D \mid T,M,E)$	951.65	14.2	951.65	14.19
$p(D \mid T,M)$	954.89	13.66	954.89	13.66
$p(D \mid T,M,C)$	954.95	13.89	954.96	13.89
$p(D \mid T,M,N)$	955.01	13.93	955.02	13.93
$p(D \mid T,M,C,N,L,E)$	955.33	14.43	955.34	14.43
$p(D \mid T,M,L)$	956.54	13.61	956.55	13.61
$p(D \mid T,M,R,V)$	956.91	13.83	956.93	13.75
$p(D \mid T,M,R)$	956.93	13.75	956.92	13.83

Table 2.4: Scores of WAIC and LOO-CV for the p(D) models, with Null hypothesis shown in red

The results indicate that there is evidence in support of some models that predict thermal comfort as a function of both thermal and non-thermal IEQ parameters. With respect to WAIC and LOO, $p(S|T,M,C)$, $p(S|T,M,E)$, and $p(S|T,M,C,E,L,N)$ provide improved predictive accuracy over the Null hypothesis for thermal satisfaction. Likewise, $p(D|T,M,E)$ provides improved predictive accuracy over the

Null hypothesis for thermal dissatisfaction. Though this observation is consistent with the odds ratio analysis visualized in Figure 2.6, the calculated WAIC and LOO scores indicate that preferential selection of any one of these models is not supported by *strong* evidence. The difference in WAIC / LOO scores is too small and/or the difference in the score errors are too large to make such a claim.

Predictive Posterior Checks

Predictive posterior checks are used to illustrate whether a selected model's predictions are more representative of observed data than the Null hypothesis. Figures 2.7 and 2.8 shows the posterior predictive distributions drawn from 1000 samples from different $p(D)$ and $p(S)$ models respectively. The Figure examines specifically how well individual models fit to the upper and lower 50% quantiles of observed conditions. These observations are indicated with a red circle, corresponding to the proportion of occupants, out of the field dataset, who claimed to be thermally dissatisfied (or satisfied) in situations where measured non-thermal conditions of IEQ were in the quantile ranges shown. The Figure appears to substantiate some of the results of the model comparison process. Adding parameters E, C, and N to the Null hypothesis establishes a better fit of $p(S)$ and $p(D)$ to observations than the Null hypothesis itself. Specifically, the Maximum a posteriori estimates (MAPE) drawn from the posterior distributions of models including E, C, N are closer in fit to observed data than the MAPE drawn from posterior distributions of the Null hypothesis.

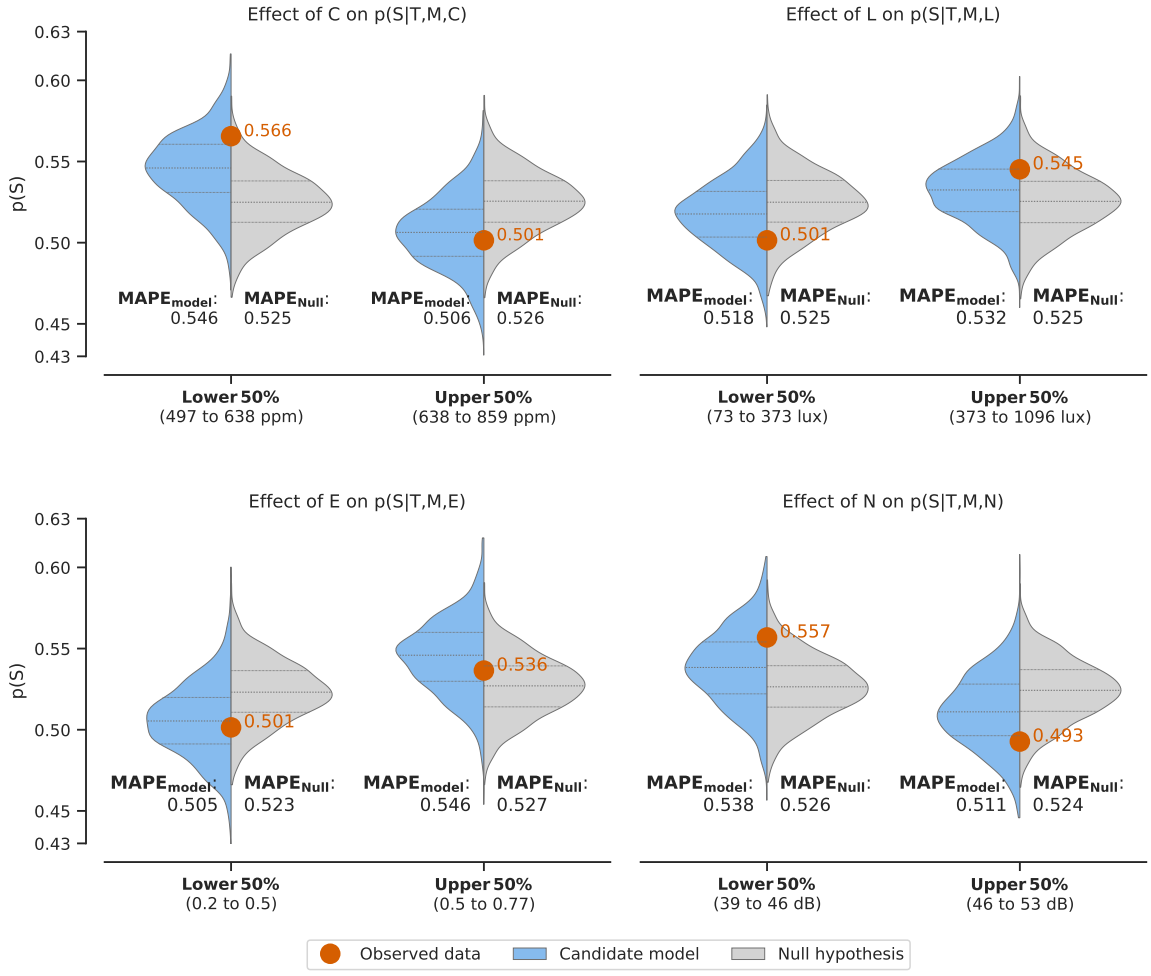

Figure 2.7: Posterior predictive distributions of $p(S)$ for different quantiles of field observations.

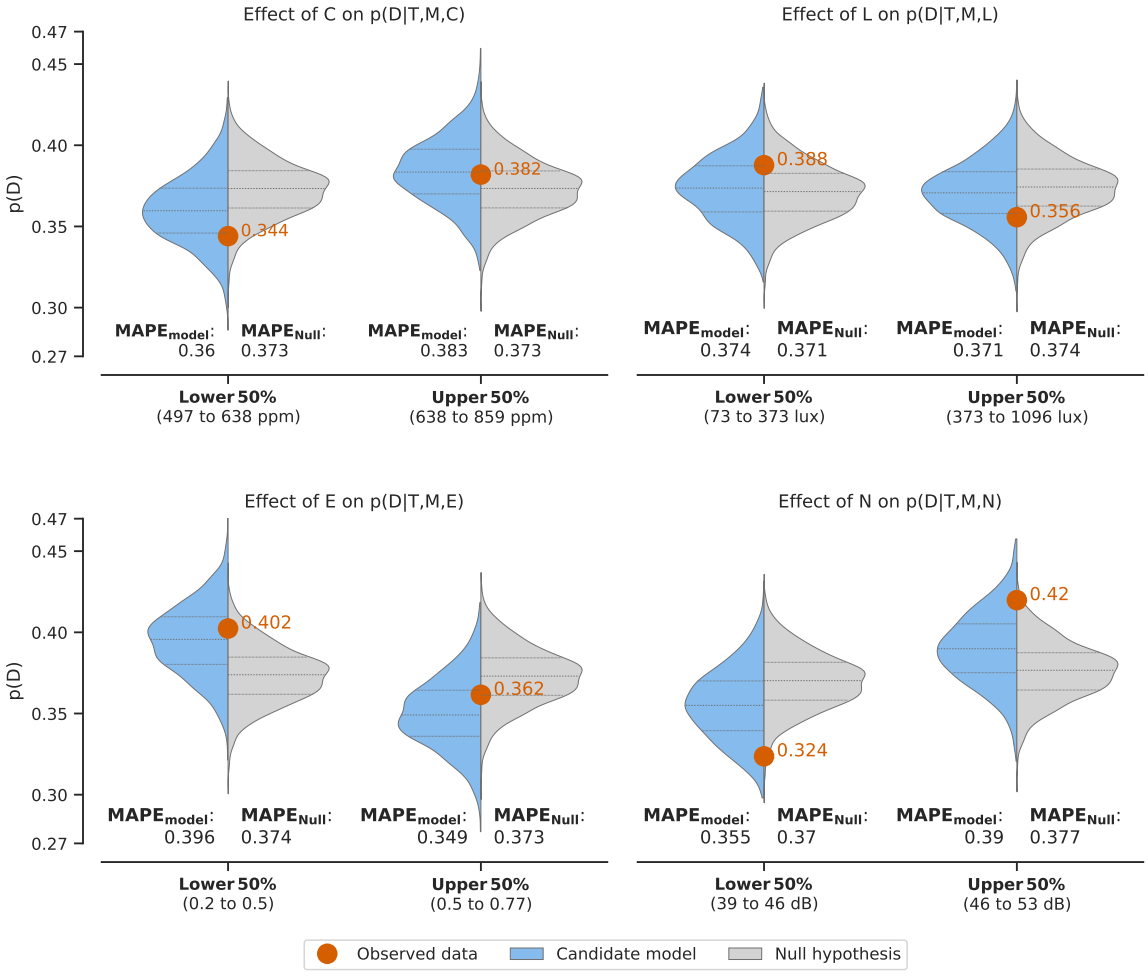

Figure 2.8: Posterior predictive distributions of $p(D)$for different quantiles of field observations.

2.4.3 Drawing Posterior Predictions from the Models

A common illustrative representation of thermal comfort analysis is the plotted relationship between perceived thermal comfort and indoor air temperature and/or operative temperature. In Figure 2.9, a similar illustration is presented in visualizing posterior predictions of thermal satisfaction against both operative temperature and individual non-thermal parameters of **W**. Operative temperature is assumed to be the average of indoor air temperature and mean radiant temperature. The standard deviation of posterior predictions is denoted by the translucent band around the median.

Figure 2.9 reveals that, between an indoor air CO_2 concentration of 500 ppm and 900 ppm, and at an operative temperature of 23 °C, the median likelihood of an occupant in the COPE study feeling thermally satisfied decreases from 0.57 to 0.45. Recognizing the predictive uncertainty of these findings, we can alternatively state that from 500 ppm to 900 ppm, $p(S)$ is predicted to decrease by $30 \pm 8\%$. A similar difference in thermal satisfaction responses is observed for speech intelligibility. Between measured values of the speech intelligibility index from 0.3 and 0.8, $p(S)$ increases by $16 \pm 3\%$.

These are not trivial ranges for the independent variables and predicted effects on thermal satisfaction; for instance, 900 ppm is still well within historically-prescribed CO_2 limits for good indoor air quality in buildings [97]. Speech levels in indoor workplaces also vary widely [59, 115]. If a future study reinforces these findings with more evidence, this would imply that a stronger-than-previously-known predictive relationship may exist between building ventilation rates, interior design, perceived thermal comfort, and ultimately building energy use.

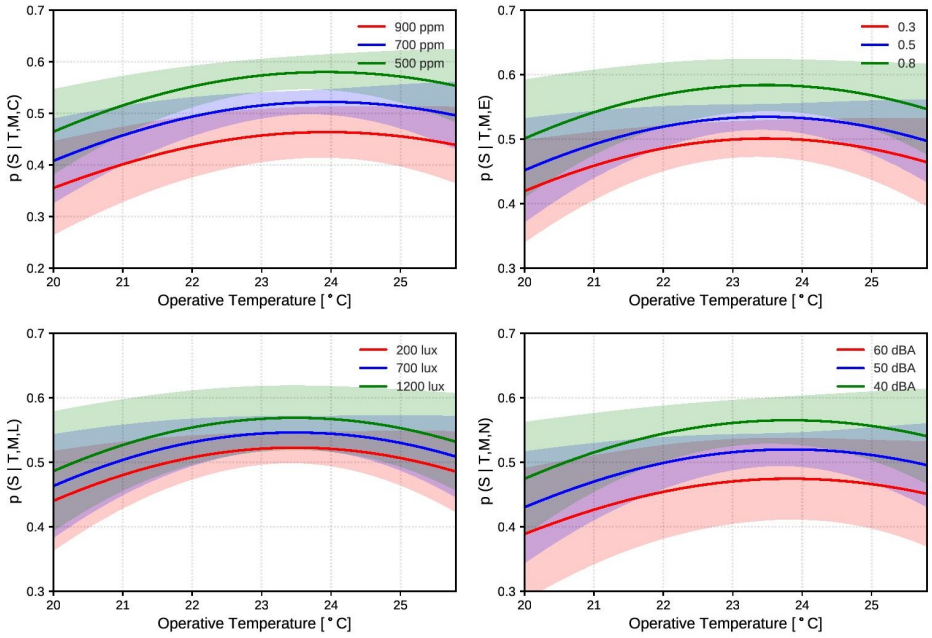

Figure 2.9: $p(S)$ models posterior predictive showing the effect of each non-thermal parameter on the relationship between operative temperature and thermal satisfaction; mean and standard deviation of predictions shown (Unless otherwise specified, R=30%, V=0.08 m/s, T=M=Operative temperature)

2.5 Discussion

2.5.1 On Establishing the Null Hypothesis

The Null hypothesis was established as the best performing predictive models of $p(S)$ and $p(D)$ that included only thermal parameters as independent variables. The models $p(S|T,M)$ and $p(D|T,M)$ were identified as the Null hypotheses respectively for S and D, shown in Tables 2.3 and 2.4. The establishment of these parameters as the Null hypothesis may go against convention; all buildings in the COPE field study are mechanically ventilated and therefore would adhere, by code

if not by physics, to the PMV-PPD thermal comfort regime. The PMV-PPD model would suggest indoor air humidity and air velocity should also be considered influential parameters. Their exclusion from the Null hypothesis is explained in two parts. First, it was noted that most measurements in the COPE dataset were undertaken in offices in cold climate settings, where indoor air humidity, in absolute terms, was relatively low (95% of measurements between 1.9 and 9.5 g/kg, with mean of 5.4 g/kg). Indoor air velocity was also consistently measured to be low, with over 60% of measured values below 0.1 m/s. Within these ranges, the PMV-PPD models informs us that the effect of air velocity and humidity on $p(D)$ and $p(S)$ is likely low or negligible [116]. Though the Null hypothesis can't be claimed to globally fit the regression of thermal comfort field study data better than a model more representative of the PMV-PPD model, $p(S)$ and $p(D)$ can be suitably explained by air temperature and mean radiant temperature, as far as thermal parameters of IEQ are considered and in the COPE dataset specifically.

2.5.2 On Model Comparison and Selection

Four different approaches to comparing predictive Bayesian statistical models to observed data were applied in this work: analysis and visualization of Odds Ratios, calculations of WAIC and LOO scores and visualization of predictive posterior values of $p(S)$ and $p(D)$ against data. The results of all model assessment methods are broadly consistent. For example, the odds ratio for speech intelligibility E, suggests with high credibility that a reduction in E will reduce $p(D)$, all other parameters being equal. It also appears credible, though slightly less so, to state that an increase in E leads to an increase in $p(S)$. Likewise, a decrease in C appears likely to increase $p(S)$ independent of all other parameters.

The WAIC and LOO scores are consistent with the odds ratios findings, but provide a more nuanced assessment of the robustness of these observations. The WAIC and LOO scores substantiate a view that including E in the estimation of $p(D)$ and $p(S)$ increases predictive accuracy, and to an extent also with C. However, the errors in the WAIC and LOO scores are sufficiently high that one cannot state there is strong evidence in support of the models that include E and C. Some evidence, or weak evidence, yes; strong evidence, no.

This provides context to interpreting the predictive posterior checks. In Figures 2.7 and 2.8, it's shown that the MAPE of all models reflected improved predictions of $p(D)$ and $p(S)$ when compared to the Null hypothesis. Beyond this visualization, the WAIC and LOO scores are used to understand which of these model results are likely more robust to future out-of-sample data. Though evidence in support of $p(S|T,M,E)$ and $p(S|T,M,C)$ is not *strong*, the results suggest that the performance of these models with respect to WAIC and LOO is sufficient to warrant more data collection on speech intelligibility and CO_2 concentrations.

If the WAIC and LOO scores for these models further improve after combining the COPE dataset with another field study dataset, it may be possible to establish a more universal statement about the magnitude of correlation between perceived thermal satisfaction, measured CO_2 concentrations, and speech intelligibility. This is discussed in detail in the next chapter (3 and in section in 6.2.3 of chapter 6.

2.6 Summary of Findings

This chapter presented a Bayesian methodology for empirical regression of predicted thermal satisfaction against not only measured thermal parameters of IEQ but also measured non-thermal indoor environmental conditions. Bayesian logis-

tic regression was applied to a field dataset in order to evaluate the relationships between surveyed occupant thermal dissatisfaction, thermal satisfaction and eight thermal and non-thermal indoor environmental quality parameters. Posterior results suggested that there is some evidence in support of speech intelligibility and indoor CO_2 concentrations being correlated with occupant responses to the question of perceived thermal dissatisfaction / satisfaction. It has been found that indoor CO_2 levels and thermal dissatisfaction are positively correlated. In support of this, the results also support a view that perceived thermal satisfaction is negatively correlated to CO_2 concentrations. In addition, It has been found that speech intelligibility and thermal dissatisfaction are negatively correlated.

Analysis and visualization of posterior predictive checks and odds ratios suggested, with high credibility that occupants' thermal satisfaction appears to increase when indoor CO_2 levels decrease. The evidence base for all of the investigated models was tested using the Widely-Applicable Information Criterion (WAIC) and Leave-one-out (LOO) cross-validation techniques. The results suggested that including measurements of CO_2 levels in the estimation of thermal comfort might improve the prediction accuracy of thermal satisfaction. The WAIC and LOO scores, however, show relatively high errors that the models including speech intelligibility index and CO_2 cannot prove strong evidence.

In the following chapter, a new IEQ field study at the University of British Columbia is used to determine whether the evidence base for the findings drawn in this chapter is improved upon the addition of the new data and whether the correlations between thermal comfort and non-thermal metrics of IEQ can be significant and statistically robust. Compared to the COPE field dataset of the early 2000s, the new IEQ field study uses modern equipment and building systems.

Chapter 3

Predicting thermal satisfaction as a Function of CO_2 levels: Second Case Study

3.1 Introduction

In chapter 2, a Bayesian logistic regression model for evaluating the correlations between perceived thermal comfort and several non-thermal IEQ conditions, such as indoor lighting levels and CO_2 levels, was developed. The model was trained on a prior IEQ field dataset of open-plan offices, produced by the National Research Council of Canada (NRC) in the early 2000s [18, 129].

The Bayesian framework was configured so that it could estimate the probability of an occupant stating they feel thermally satisfied, $p(S)$, as a function of not only psychrometric/thermal IEQ parameters, $\mathbf{F}$, but also of one or several

non-thermal IEQ parameters, $\mathbf{W}$, or $p(S \mid \mathbf{F}, \mathbf{W})$ [27]. Initial results from that first analysis revealed some evidence suggesting that adding measurements of indoor CO_2 concentrations and speech intelligibility to predictive models of thermal satisfaction could provide better predictive accuracy than models that would not include these parameters.

In this chapter, this work is updated by, first, reformulating the model as a *hierarchical* (or multi-level) Bayesian logistic regression problem. Second, by adding new samples of subjective and objective IEQ measurements collected from a new field IEQ study carried out at the University of British Columbia in 2019 and 2020. The newly collected field IEQ database, and simulation code for the hierarchical model is publicly accessible as part of this thesis. This chapter then seeks to address whether, on the addition of new field data and applying a more comprehensive Bayesian inference technique, the evidence of correlations between thermal satisfaction and non-thermal metrics of IEQ is strengthened - or weakened.

In section 3.2.1, the experimental apparatus and full methodology of the new field study is described. In section 3.2.2, the revised methodology for hierarchical inference of the logistic regression model is presented. In section 3.2.3, different methods for evaluating model fitness and accuracy are reviewed. In section 3.3.1, the results of the field study are presented. In section 3.3.2, the results from the correlation analysis of the subjective and objective data collected in COPE and UBC studies are presented. In section 3.3.3, the results of inferring the hierarchical model are laid out. In section 3.3.4, model comparison, selection and validation checks performed on the candidate models are presented. All results are then discussed in section 3.4.

3.2 Methods

3.2.1 Design of UBC Field Study: Second Case Study

Overview

Bayesian statistical methods rely on the concept of improving one's understanding of a problem or phenomena through the addition of new data. The COPE dataset that under-laid the previous analysis (in chapter 2 is today over 20 years old. No publicly-available data sources for similar field studies that contain the same granularity of data on IEQ parameters as the COPE study have been found. Hence, a new field study was developed for this research work at the University of British Columbia (UBC) to directly complement the COPE dataset with new observations. The new IEQ study utilizes modernized instrumentation under the auspices of more modern indoor building environments and building systems compared to the COPE field study of the early 2000s. Yet the UBC study also attempts to mirror the experimental procedure of the original COPE study as best as possible.

The result of the new field study is the creation of the 'UBC dataset', consisting of instantaneous physical measurements of IEQ, spatial, and geometric parameters coupled with responses from an IEQ questionnaire. The field study, carried out at the Vancouver campus of the University of British Columbia (UBC), collected responses from approximately 150 office workers in four buildings between 2019 and 2020. Most buildings were visited twice, in summer (between July and August 2019) and winter (between January and February 2020). The goal of visiting the same buildings twice was to capture diverse climatic weather conditions for a more generalized findings. However, during the second visit, the participants were a

mix of new participants and the same participants of the first visit. This might cause some limitations in assuming independence of the collected subjective data. Further responses were expected after February 2020 but were curtailed due to the onset of the COVID-19 pandemic and the mandatory vacating of office spaces at UBC and throughout Vancouver.

Building ID	**Dataset**	**No. Floors**	**No. Samples**	**Date of Visit**	**Mechanical Ventilation**	**Year built**	**Windows**
B1-B9	COPE	Refer to Charles et al.[18] for a summary of the COPE building characteristics					
B10	UBC	4	54	July 2019 & January 2020	ducted-air VAV heating; natural ventilation cooling	2011	operable windows
B11	UBC	2	27	July 2019 & February 2020	ducted-air VAV heating; natural ventilation cooling	1977	operable windows
B12	UBC	5	19	July 2019 & February 2020	ducted-air CAV heating; natural ventilation cooling	1961	operable windows
B13	UBC	6	41	August 2019	ducted-air VAV heating & cooling	2003	non-operable windows

Table 3.1: Summary of the COPE and UBC field study buildings; all UBC building are located in Vancouver, BC, Canada

IEQ Questionnaire

An updated version of the 'right-here-right-now' occupant satisfaction survey developed originally for the COPE study was used for the field study at UBC. Though there exist more state-of-the art survey questionnaires than what was issued by the COPE field study in 2000, retaining similarity between both field studies was critical. Effectively, the UBC IEQ questionnaire was mirrored against the COPE questionnaire so that both datasets can be easily combined and compared against each other. Nevertheless, a few more questions were also added to the UBC ques-

tionnaire in line with current practise vis-a-vis IEQ field evaluation. Overall, the questionnaire presented itself as an online, anonymous survey of perceived 'right-here-right-now' and long-term IEQ conditions.

The survey was designed to be completed online at the workstation of the respondent shortly before physical IEQ measurements would be taken. Questions about IEQ satisfaction included three different perception categories for each IEQ parameter: perceived levels ("How would you rate the temperature at your workspace right now?"), satisfaction with perceived levels ("How would you rate your satisfaction with the air temperature at your workspace right now?"), and preference to perceived levels ("How would rate your preference with respect to the air temperature at your workspace right now?").

The first two questions were answered on a 7-point scale (1 = Cold and 7= Hot)and (1=Very unsatisfied and 7=Very satisfied) respectively. While the third question was answered on a 3-point scale (1=prefer warmer; 2=remain the same; and 3=prefer cooler).

The survey covered satisfaction with the thermal environment, job and workplace satisfaction, long term satisfaction with the air temperature, satisfaction with background noise levels, lighting levels, daylight availability, glare, view to the outside, quality and quantity of artificial lighting, air quality, air movement, humidity, IEQ controllability, and other individual features of the workspace. The survey also covered some long-term satisfaction questions, for example: ("Over the PAST YEAR, how satisfied have you been with the overall air temperature at your workspace?"). These questions were clearly worded, as best as possible, to avoid confusing the participants, however, switching between right-here-right-now and long-term questions might still cause some confusion. A list of all the survey

questions is laid out in Table C.1,C.2 and C.3 in Appendix C.

Mobile IEQ Measurement Station

An experimental IEQ sensor cart was designed and built for this study, illustrated in Figure 3.1. The cart carried all the IEQ sensors required for the local microclimate measurement conducted at each workstation. The measurements taken by the cart align with the measurement station of the original COPE study. The station measured indoor air temperature, relative humidity, mean radiant temperature, CO_2 concentrations, CO concentrations, concentrations of total volatile organic compounds (TVOC)s, air velocity, A-weighted noise levels, and desktop illuminance levels. Table 3.2 describes all the sensors/instruments mounted on the UBC IEQ cart as well as the sensors that were used in the original COPE field study [95].

Experimental Procedure

In the Spring of 2019, administrative representatives of academic departments and central administrative units of UBC were contacted to advertise this study. Ethics approval for the study was sought and received from the UBC Office of Research Ethics (Ref no.: H19-01364). Ethics approval required submission and review of the project's research proposal, the consent form issued to participants, the survey questionnaire, the initial contact e-mail letter to be sent to prospective participants, the overall experimental procedure, and the project information sheet issued to prospective participants. All interested candidates in the study were asked to read an information package given to them and sign a consent form if they would be willing to proceed with the study.

The field study targeted healthy adult participants who were employees of UBC

IEQ metric measured	UBC cart (circa 2019)		COPE cart (circa 2000)	
	Instrument	**Accuracy (Range)**	**Instrument**	**Accuracy (Range)**
Air Temperature	TSI Velocicalc-9565	±0.3°C (-10°C to 60°C)	Omega RTD	±0.1°C (Room temp[i])
MRT	Hearth lab-CUBE	±0.3°C (-40°C to 70°C)	Omega RTD	±0.1°C (Room temp)
Relative humidity	TSI Velocicalc-9565	±3%RH (5 to 95%RH)	General Eastern RH2	±2%RH (20 to 95% RH)
Air Velocity	TSI Velocicalc-9565	±0.015 m/s (0 to 50 m/s)	TSI-8475	±3% (0 to 1 m/s)
Light intensity	Dr. Meter Illuminance sensor	±3% ±10 digits (0.1 to 200,000 Lux)	Minolta T1	±7% (0.01 to 999,00 Lux)
Sound pressure levels	Soundpro- SE	Class 1 (10 to 140 dB)	Rion NA-29	±0.1 dB (27 to 130 dB)
CO_2 concentrations	GrayWolf IQ-610	±3%rdg ±50ppm (0 to 10,000 ppm)	B&K 1302	±0.3 ppm (-)
CO concentrations	GrayWolf IQ-610	±2ppm (0 to 500 ppm)	B&K 1302	±0.3 ppm (-)
VOC concentrations	GrayWolf IQ-610	(5 to 20,000 ppb)	B&K 1302	±0.3 ppm (-)
Biophilia Images	RICOH Theta SC 360-degree camera	-	-	-
Physical measurements	Bosch-GLM 165-40 Laser measure	±1.5 mm (up to 20 m)	-	-

[i] see [128] for context

Table 3.2: IEQ sensors mounted on the UBC and COPE carts

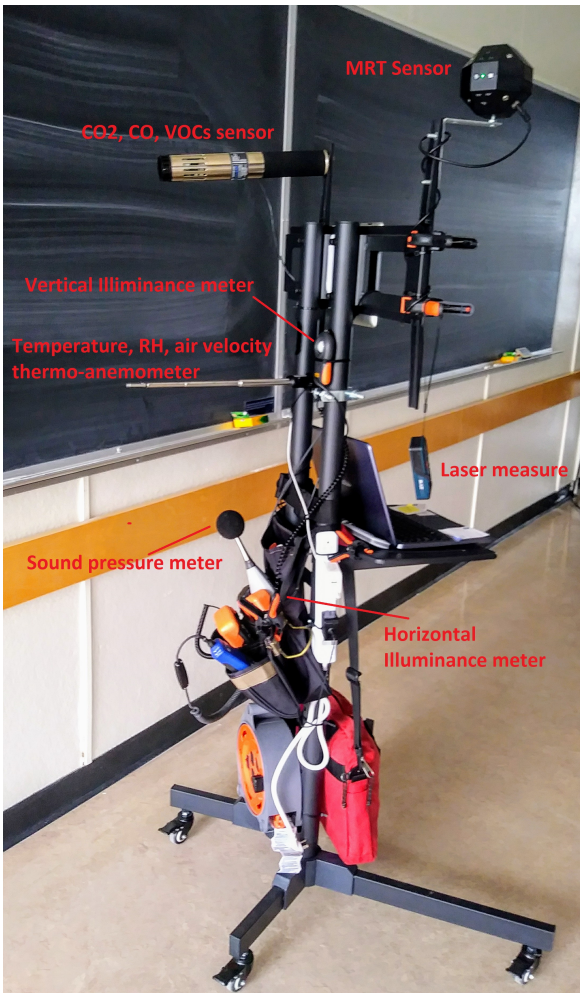

Figure 3.1: IEQ sensor cart developed for the UBC study

and worked day-to-day in office-type work environment. Participants were not discriminated by age, gender, or race. Only occupants of administrative or faculty offices were selected. Interested participants were contacted by e-mail and provided with an information sheet about the nature of the study and its potential implications. They were then asked to opt-in to the experiment via a secure, online sign-up form to schedule their involvement and select a day and time to participate in the

experiment. Despite any possible initial personalized contact with the research team as well as the register of their consent, no personally-identifying information was retained during the experimental phase; all field results were anonymized. All the measurements were made during normal working hours only (8 AM to 4 PM) between Monday and Friday. The specific experimental procedure for each consenting participant was as follows.

On the morning of each experiment day, a link to the online survey was e-mailed to each confirmed participant. Participants were asked to save this link as they would be answering the survey only during a scheduled 15-minute experimentation period later in the day. Occupants were asked to only schedule an experiment time that would assure they would have been in a sedentary position at their office for at least an hour beforehand. This was to ensure that each participant would be, as much as possible, at equilibrium with their surrounding environment prior to their participation.

At the scheduled time, a team of 2-3 researchers visited a participant's workstation. First, they verified whether the participant had consented to the study. Upon confirmation of consent, each participant was given a unique random numeric identifier which served as the participant's only identifier that would later associate their survey responses with the manual and physical data collected. Participants were then asked to open and complete the online IEQ questionnaire at their desk, while the research team left the respondent's workstation. At the time of this visit, the following measures were taken to mitigate the Hawthorne effect, which implies that the participants may alter their behaviour when they are being observed [15]. The research team visit was planned, on an opt-in basis, and completely voluntary. With prior communication, each participant chose their time of

participation at their own preference. All the participants were faculty members or university staff which created a friendly and relaxed environment with the research team. Participants had the full support of their managers and they were encouraged to take a break while the participants finished their data collection. Upon their arrival, the research team started with some introductions with the participants, which is a recommended practice to avoid the Hawthorne effect [100]. Once each participant started filling out the survey, the research team removed themselves from the vicinity of the participants and the participants were not being watched at that time.

Each participant referenced their unique identifier number while answering the survey. On average, participants completed the survey within 10 minutes, upon which they would meet the research team and allow the team to undertake physical measurements. The research team undertook physical measurements by removing the office chair from the participant's workstation and placing the IEQ measurement cart in its place. On installing the cart, the researchers immediately left the workstation, stood at a distance of at least 2 meters away from the sensor cart, and allowed IEQ measurements to be taken. The collecting of IEQ sensor data was semi-automated using a laptop computer installed on the sensor cart. Continuous measurements from the air quality and environmental sensors were collected at intervals of 5-10 seconds over a period of 3-5 minutes. The data was subsequently time-averaged to obtain single values representing the steady-state conditions at a participant's workstation. At the end of the 3-5 minute measurement period, the research team undertook physical/geometric measurements of the workspace and installed light and sound meters on the participant's desk.

The additional manual and spatial measurements taken at various stages of

Observational measurement	Manual measurement
Office Type (open plan; semi private '2-5 desks'; private)	Ceiling Height (m) (from floor)
Is it raining?	Distance to indoor plant (m)
Are there blinds?	Distance to Nearest Window (m)
Are the blinds operable?	Relative Orientation of Window
Are the blinds down?	Distance to Second Window (m)
Clothing (Tops; bottoms)	Relative Orientation of Second Window
Footwear & socks	Desk Width (m)
Desk Material	Desk Depth (m)
Desk Surface Colour	Desk Height (m)
Chair Type (Office Chair, Stool, Armchair)	Partition Height (m) (if applicable)
Chair Material (cloth, mesh, wood)	
Floor Material	
Floor Surface Colour	
Wall Surface Colour	
Ceiling Surface Colour	
Partition Surface Colour (if applicable)	
Is there a desk lamp?	
Are the windows open?	
Desk lamp exists and on?	

Table 3.3: Observational and manual data collected

the 15-minute experimentation window included but not limited to: participant clothing levels, workspace size, outdoor raining conditions, blinds and windows presence and operability, desktop partition height, distance to the nearest indoor plant, distance to the nearest windows, presence of a desk lamp. A 360° image of the workspace was taken using a panoramic camera, mounted on a tripod and placed on each participants' desk.

Clothing levels data was collected by the research team for each participant through observation and then logging the clothing data in the observational measurement spreadsheet on the laptop computer. In summary, Table 3.3 lists all the manual measurement and observational data collected by the research team. Table C.1 and Table C.2 lay out the individual survey questions and the format of

responses of each question. The typical experimental setup followed in the UBC field study is illustrated in Figure 3.2. At the end of each participant's experiment, the sensor cart was removed, the participant's chair replaced, and the participated was thanked for their participation.

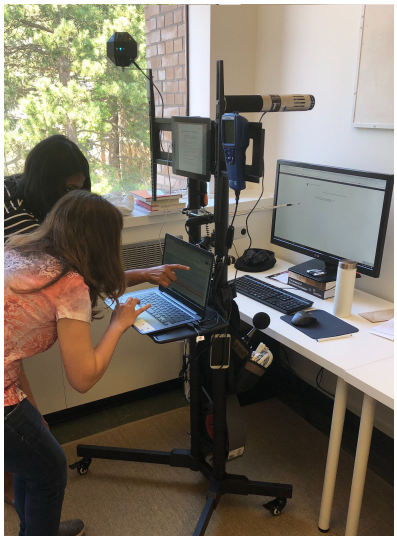

Figure 3.2: Researchers beginning the process of collecting IEQ sensors measurements at a participant's workstation; shortly after beginning the automatic data collection process, the researchers move at least 2m away from the sensor cart

3.2.2 Hierarchical Bayesian logistic Regression Model

Overview

In section 2.3.1, a methodology for inferring the probability of an occupant feeling thermally satisfied, $p(S)$, given a set of thermophysical IEQ parameters, **F**, and any other parameters more typically consistent with non-thermal IEQ conditions, **W**, is described. The method was established as a Bayesian logistic regression problem.

In this chapter, a *hierarchical* (multi-level) Bayesian logistic regression model

for inferring $p(S\,|\mathbf{F},\mathbf{W})$ is developed. Multi-level Bayesian modelling is useful when one faces differently-sized, independent datasets in a Bayesian inference problem. It can allow one to fit a regression model to the datasets while accounting for unexplained variations among the datasets themselves. Hierarchical Bayesian modelling has been used in several prior research works in the field of building science [91, 105, 111]. The process intends to prevent the final model parameters from being influenced by outlying data subsets. Hierarchical modelling is known to often outperform traditional (non-hierarchical) regression techniques with regards to model fitness and identifying true model parameters, particularly when managing multiple independent data sets [45]. It is observed that not only are the COPE and UBC datasets unique, but within these datasets one might assume that the subset field studies of individual buildings B are not necessarily perfect subsets of data. It is plausible that any building-specific study may be affected by some unknown random effects that are possibly, but unknowingly, attributable to the specific building, its location, the time of the survey, etc. On the basis of this assumption, this work assumes there are a total of nine independent field trials within the COPE dataset, and four for the UBC set (for a total of 13).

For context, the generic format of the Bayesian logistic regression model for predicting $p(S)$ as a condition of parameter sets $\mathbf{F}$ and $\mathbf{W}$ is re-stated in this section:

$$p(S\,|\,\mathbf{F},\,\mathbf{W}) = \frac{1}{1+e^{-[\Sigma_{\mathbf{F}}\,(\beta_{\mathbf{F}}\,.\,\mathbf{F})+\Sigma_{\mathbf{W}}\,(\beta_{\mathbf{W}}\,.\mathbf{W})+\beta_0]}} \tag{3.1}$$

where β_0, $\beta_{\mathbf{F}}$, and $\beta_{\mathbf{W}}$ are regression coefficients; $\mathbf{F}$ may include any number of thermo-physical parameters such as indoor air temperature, T or relative humidity, R; $\mathbf{W}$ may include any number of non-thermal IEQ parameters such as indoor CO_2

concentrations, C, ambient noise levels, N, etc.

In the hierarchical model, posterior estimates of β are inferred for each building subset B in Table 3.1. Each β is assigned a prior normal distribution of mean μ and variance σ. A Monte Carlo sampling process (i.e., NUTS algorithm) attempts to converge on identical posterior distributions for each field study specific β, such that one observes all field study-specific β's sharing approximately the same posterior distribution.

Figure 3.3 illustrates a network diagram which details a hierarchical model of $p(S|T,C)$ inferred from the entire COPE + UBC dataset. The subscript t refers to observed data from each individual test subject, of which there are 919 total in the combined datasets.

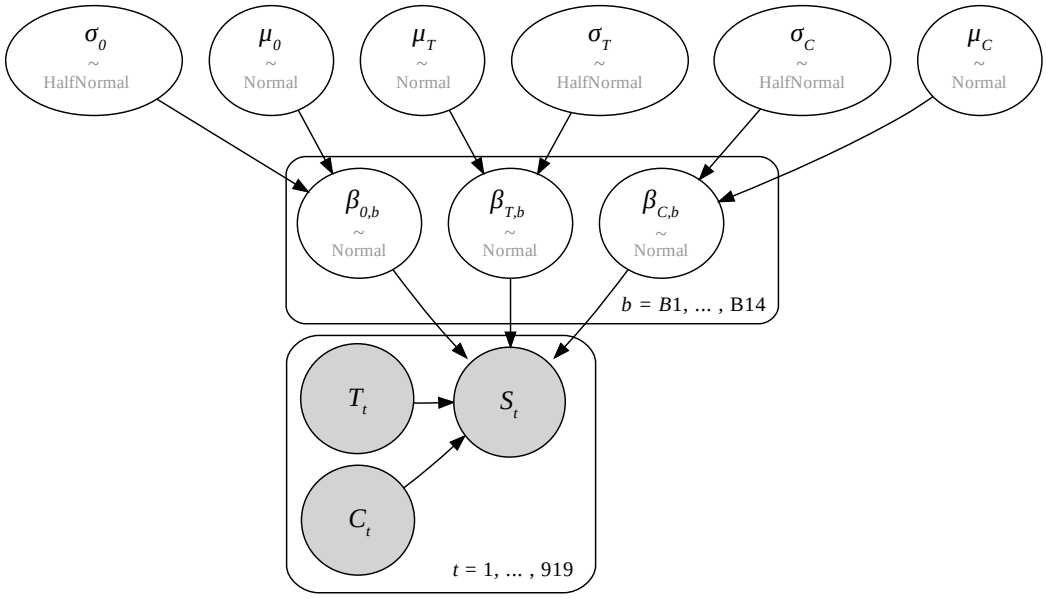

Figure 3.3: Network diagram of hierarchical logistic regression model for $p(S|T,C)$

The approach in Figure 3.3 applies similarly to other models of $p(S|\mathbf{F},\mathbf{W})$. Overall, the methodology is consistent with similar efforts utilizing hierarchical logistic regression models, and more detailed mathematical formulations of these models can be found elsewhere [106, 117, 136].

Candidate Models and Trials

In section 2.3.2, measurements of perceived thermal satisfaction from the COPE dataset were fit to several candidate regression models, each with varying independent parameters of **F** and **W**. In this section, the same model selection and comparison process is undertaken for the same number of candidate models. Table 3.4 lays out a total of 8 different candidate models of $p(S)$. The Null hypothesis assumed in this work is $p(S \mid T)$, the prediction of thermal satisfaction as a function of measurements of indoor air temperature alone.

This may seem a peculiar choice given the contemporary theory of thermal comfort. It is well understood, as per the PMV-PPD model, that predicted thermal comfort is observed to be a condition dependent on many thermal properties, including dry-bulb temperature, mean radiant temperature, relative humidity, and air velocity. This was discussed at length in Chapter 2. It was determined that $p(S \mid T)$ was an appropriate Null hypothesis for the field as all measured values for relative humidity and air velocity had little variability and all fell within the range of typical comfortable values as per the PMV-PPD model. Hence, one should accept $p(S \mid T)$ as an appropriate Null hypothesis for the specific field data evaluated, not necessarily for a universal database of field data.

Like in section 2.3.2, for each candidate model, the following variables are modelled as having a first order linear relationship with $p(S)$: indoor CO_2 concen-

trations, C; ambient noise levels, N; indoor air velocity, V; and relative humidity, R. The following parameters are modelled as having a quadratic relationship with $p(S)$: indoor air temperature, T; partition height, P; and lighting intensity, L. The choice of these relationships was previously discussed in the previous chappter in section 2.3.2.

Each candidate model in Table 3.4 is also trained on the available data in three different trials, first against only the COPE dataset, second against only the UBC dataset, and third against a combined dataset of both the COPE and UBC field studies. Unlike in the first case study performed on the COPE dataset, mean radiant temperature (MRT) is omitted from this phase as a candidate parameter of **F**. In the COPE field studies, a globe thermometer was used to collect estimated measurements of MRT. In the UBC field studies, the CUBE sensor [118] was used to collect measurements of MRT. The CUBE sensor is a relatively novel product that uses a spherical array of radiometers to measure MRT directly in an indoor space. In the intervening period since the UBC studies were carried out, new evidence has come to light which not only indicate significant measurement discrepancies between the CUBE-based MRT estimates and globe-based MRT estimates, but also a more fundamental issue with the accuracy of historical globe-based MRT estimates [118]. In this chapter, measurements of MRT are only reported but not included in the regression model analysis.

3.2.3 Model Comparison and Evaluation of Fitness

In Bayesian inference, a number of model checks and validation techniques are available to determine the evidence base of a proposed model and to evaluate its predictive accuracy [6, 125]. In Chapter 2, two cross-validation and model selec-

Model ID	Parameters of **F**	Parameters of **W**	Candidate model
1 - Null hyp.	T	-	$p(S \mid T)$
2	T, R	-	$p(S \mid T, R)$
3	T, R, V	-	$p(S \mid T, R, V)$
4	T	C	$p(S \mid T, C)$
5	T	P	$p(S \mid T, P)$
6	T	N	$p(S \mid T, N)$
7	T	L	$p(S \mid T, L)$
8	T	C, P, N, L	$p(S \mid T, C, P, N, L)$

Table 3.4: List of candidate models of predicted thermal satisfaction, $p(S)$, as a condition of different thermal ($\mathbf{F}$) and non-thermal ($\mathbf{W}$) parameters , $p(S \mid \mathbf{F}, \mathbf{W})$. T =indoor air temperature; R= indoor relative humidity; V=indoor air velocity; C=indoor CO_2 concentrations; P=partition height; N= ambient noise levels; L=lighting intensity.

tion evaluation criteria were applied: the Watanabe-Akaike Information Criteria (WAIC) scores [134] and Leave-One-Out (LOO-CV) cross validation based on pareto-smoothed importance sampling (PSIS) [104].

WAIC and PSIS-LOO scores are both Bayesian validation approaches that evaluate the likely fitness of a model to data. This is done by estimating the out-of-sample prediction error of a candidate model [47, 127]. A model is viewed to be likely a better fit to data than another if its effected WAIC or PSIS-LOO score is significantly lower (in absolute terms) than the other candidate. The out-of-sample predictive fit (PSIS-LOO score) is estimated by repeatedly partitioning the data into training and hold-out sets, iteratively fitting the model with the training dataset and evaluating the fit with the held-out data.

Although WAIC is asymptotically equal to PSIS-LOO, it has been demonstrated that PSIS-LOO may be more robust in the cases defined by weak priors or influential observations [104]. The former condition may apply to this study

as weak priors are deliberately used in the assignment of the logistic regression model β coefficients. Hence, the expected log predictive density (ELPD) obtained via PSIS-LOO cross-validation is calculated and used in this chapter to evaluate the predictive fit of candidate models.

3.3 Results

3.3.1 Outcomes of UBC Field Study

Figure 3.4 presents kernel density estimates of the measured indoor environmental conditions collected from the UBC field study. These are compared against the distributions of data in the COPE field studies. Figure 3.5 displays the subjective responses to questions in the IEQ questionnaires which overlap between the UBC and COPE field studies. Except for respondents' indication of their age and gender, all other questions regarded occupant's perceived 'right-here-right-now' satisfaction with IEQ, answered on a 7-point scale from 1 = Extremely dissatisfied to 7 = Extremely satisfied.

Figure 3.6 and Figure 3.7 present additional responses to subjective questions in the UBC IEQ survey. These questions were unique to the survey and have no direct overlaps with data collected in the COPE studies and are not used further in this analysis. The raw data collected from the UBC field study is available as part of this thesis in the following GitHub repository: https://github.com/eta-lab/UBC-field_IEQ_study. The codebase for the Hierarchical Bayesian model is also published in the same public Github repository as the UBC data. For a sample of the code, refer to Appendix B (a version of this appendix has been published in [26]. Collected manual, observational, geometric, physical measurement, as well

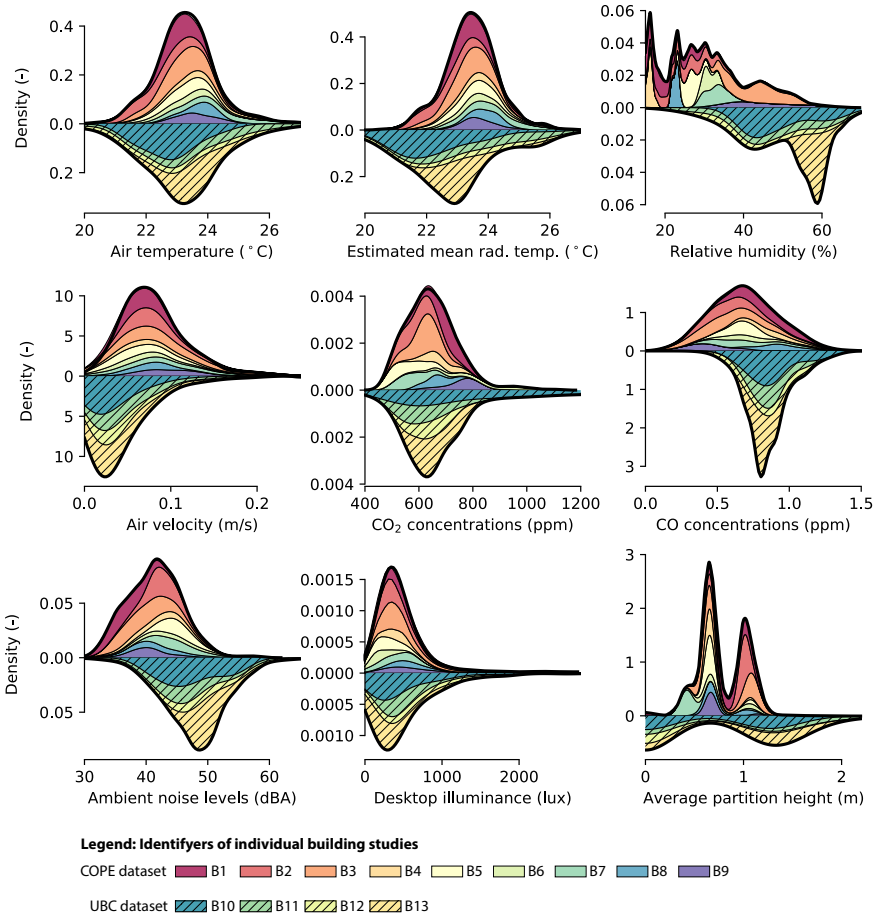

Figure 3.4: UBC field study measured IEQ metrics and comparison with COPE study outcomes

as IEQ objective measurement and satisfaction with thermal conditions data are laid out in Appendix A. For the collected subjective data, refer to Figures 3.6 and 3.7 or refer to the complete dataset in the public Github repository https://github.com/eta-lab/UBC-field_IEQ_study.

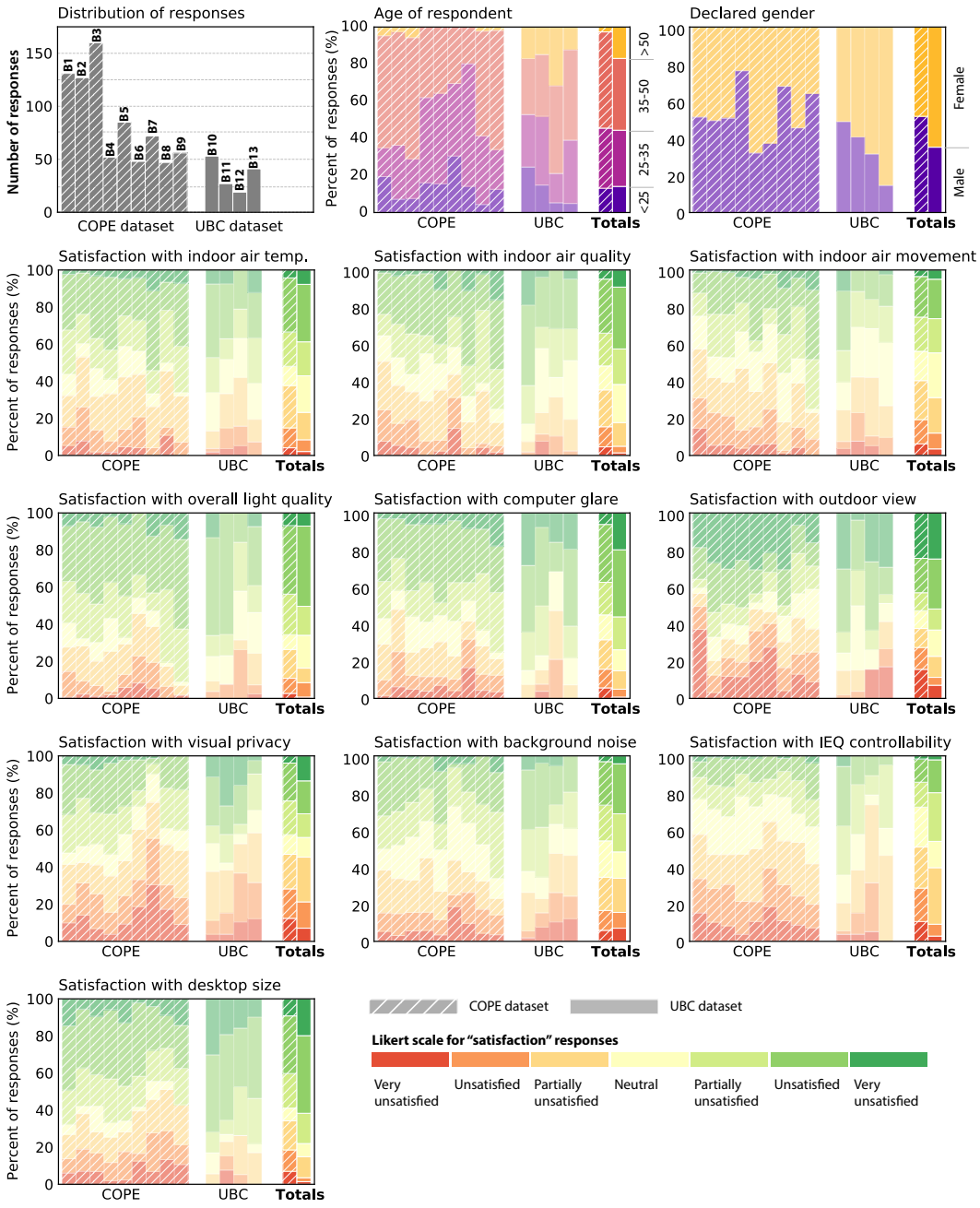

Figure 3.5: UBC field study survey responses and comparison with COPE study outcomes

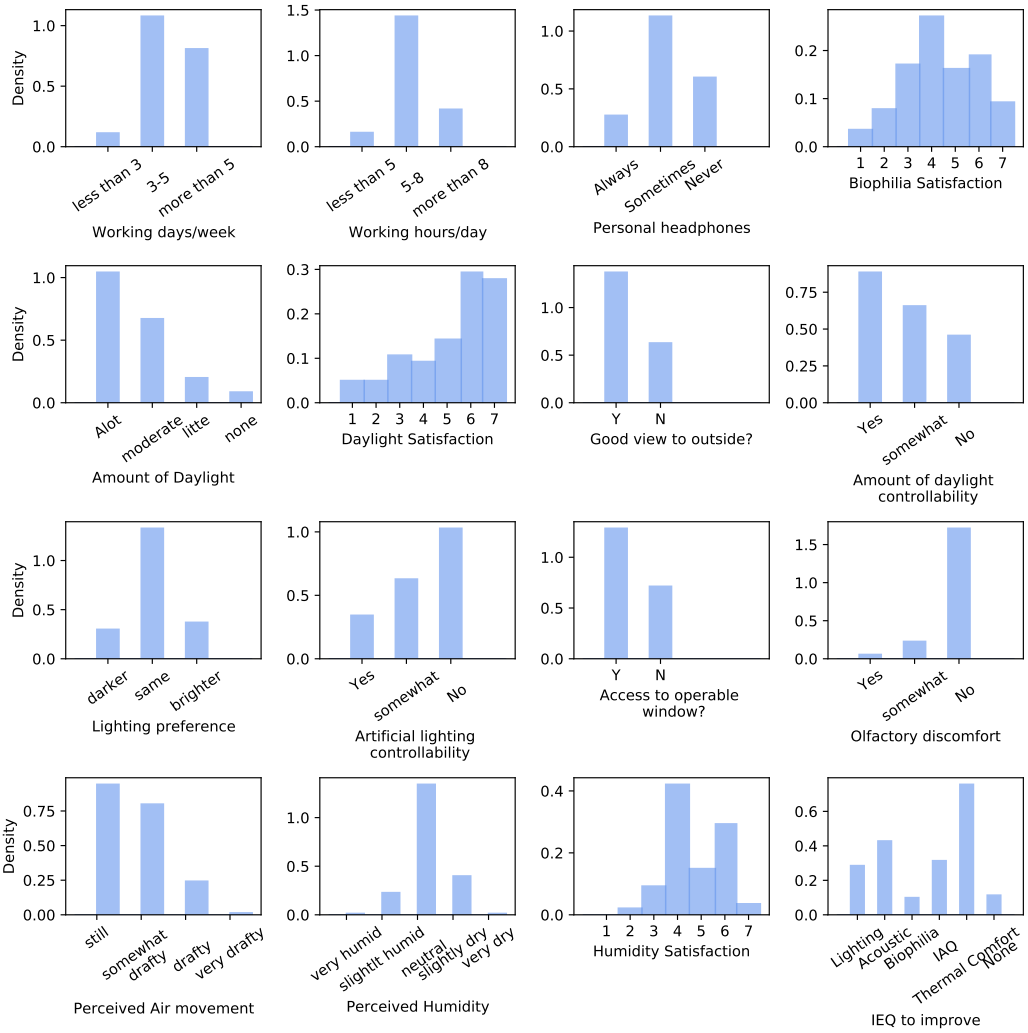

Figure 3.6: UBC field study survey responses-Part I

3.3.2 Correlation Analysis

Figure 3.8 illustrates the results of a T-test analysis comparing the statistical differences between observations between the COPE and UBC datasets. Figure 3.9 and Figure 3.10 demonstrate a Kendall-τ_b correlation analysis of the subjective

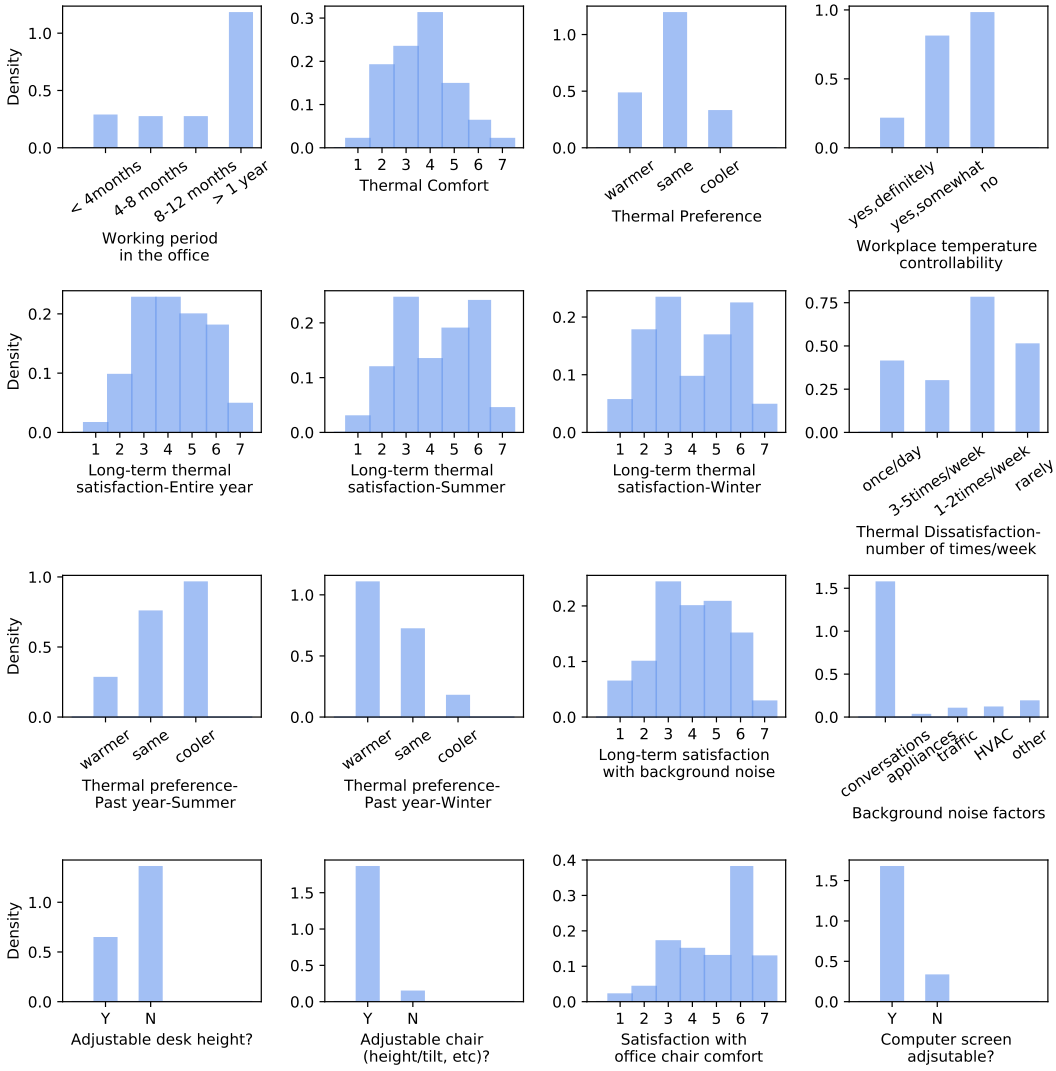

Figure 3.7: UBC field study survey responses-Part II

data collected in COPE and UBC studies, respectively. The colour code ranks the correlation strength from very strong positive correlation (dark red) or very strong negative correlation (dark blue) to no/insignificant correlation (grey). The size of the correlation squares indicates the statistical significance of the finding (p-

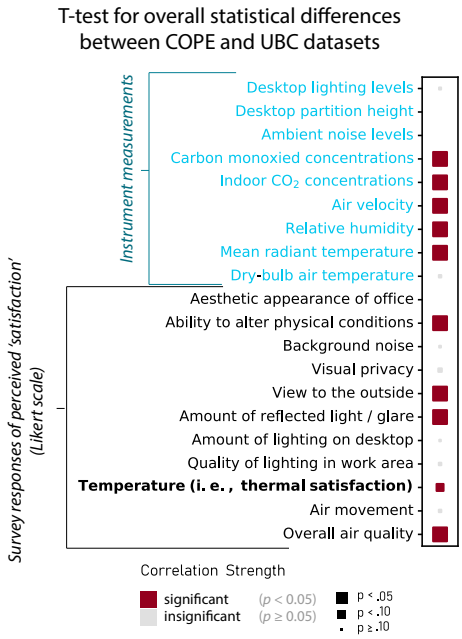

Figure 3.8: T-test evaluating overall statistical differences between COPE and UBC datasets (only parameters shared by both datasets are shown)

value). Figure 3.11 provides Kendall-τ_b values and significance for the correlation between the subjective (ordinal) and instrumentation (continuous) measurements of the COPE and UBC studies independently. Figure 3.12 similarly visualizes the Pearson correlation coefficient and significance between instrumented measurements within the COPE and UBC datasets. Similar to the Kendall-τ_b correlation heatmaps, the correlation's strength levels are displayed in different colours: from very strong positive correlation (dark red) or very strong negative correlation (dark blue) to no correlation (grey). The statistical significance of the relationships is indicated by the size of the squares. Figure 3.12 reveals significant correlations between indoor air temperature and mean radiant temperature in both datasets, and this is an expected observation. Another evident observation from Figure 3.12,

though moderate, is the inverse correlations between measured indoor CO_2 concentrations and measured mean radiant temperature, absolute humidity and relative humidity in the UBC dataset. These observations and other findings are discussed in detail in the next section.

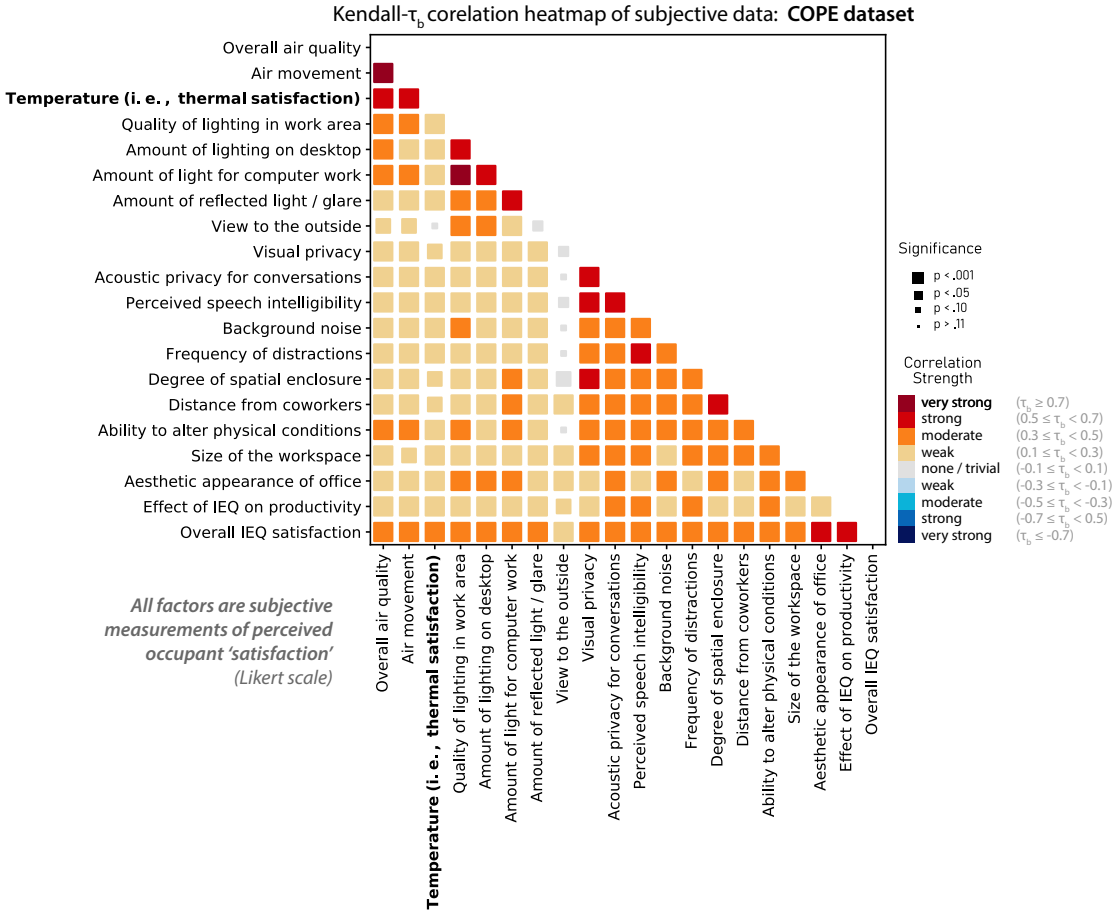

Figure 3.9: Kendall-τ_b correlation analysis of subjective data collected in COPE study

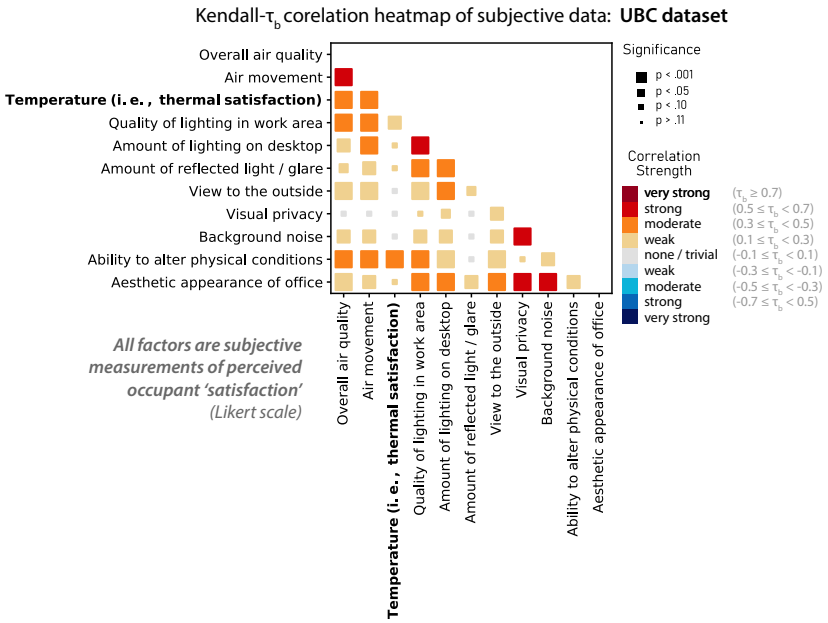

Figure 3.10: Kendall-τ_b correlation analysis of subjective data collected in UBC study

3.3.3 Regression Results

Figure 3.13 presents the posterior predictive results of the Bayesian hierarchical regression process. One thousand samples of $p(\beta \mid \text{field data})$ are drawn to generate one thousand samples of a candidate regression model of $p(S \mid \mathbf{F}, \mathbf{W})$. Posterior predictions of a model inferred from only the COPE dataset are shown in orange. Predictions drawn from a model inferred from the combined COPE & UBC dataset are shown in blue. With the exception of $p(S \mid T)$, in all other cases indoor air temperature is kept fixed at, T=23.5 °C, which is approximately the observed median measured indoor air temperature across both the UBC and COPE datasets. The median posterior predicted probability of thermal satisfaction is denoted by a solid line and the 95% credible interval of thermal satisfaction predictions are shown

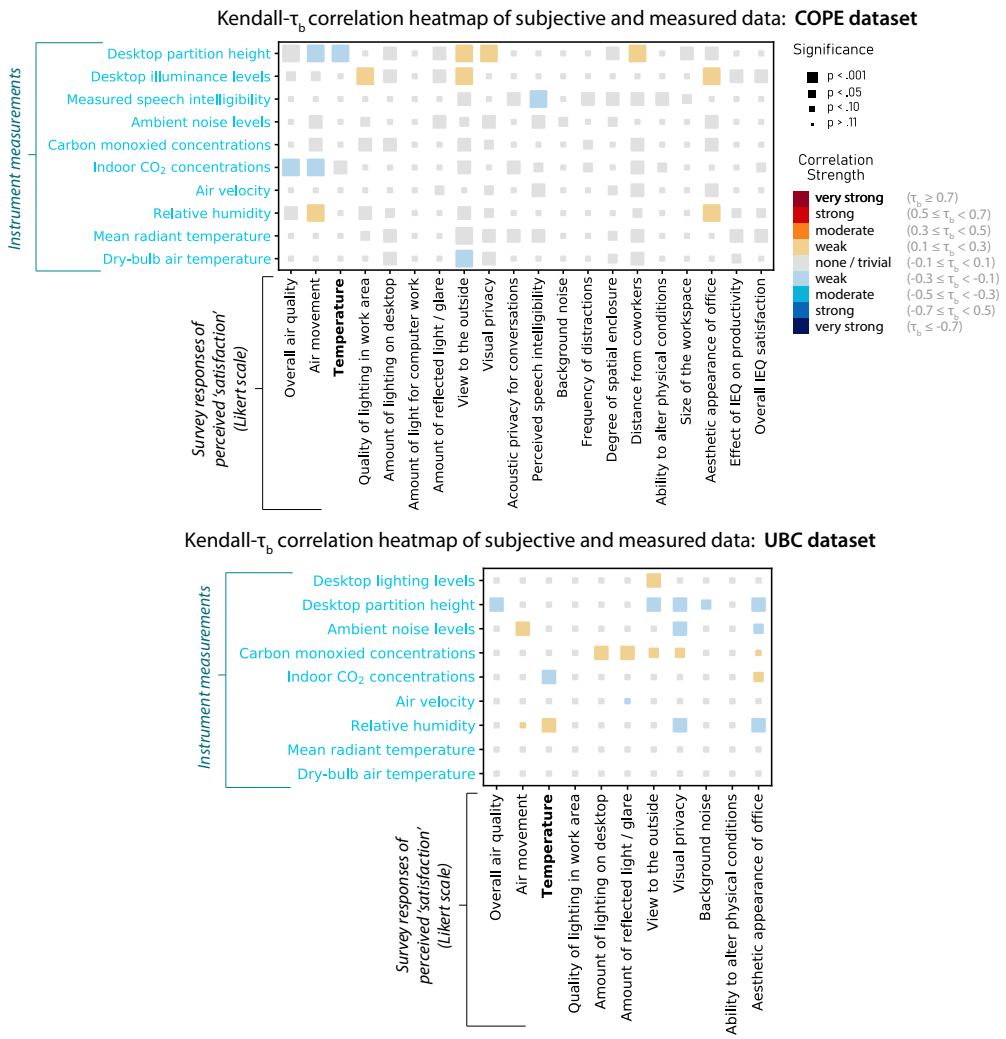

Figure 3.11: Kendall-tau_b correlation heatmaps of subjective and measured data in COPE and UBC studies

as upper and lower dotted lines around the median. Figure 3.13 also visualizes the individual probability density of observed non-thermal parameters, **W**, for both the COPE and COPE+UBC datasets.

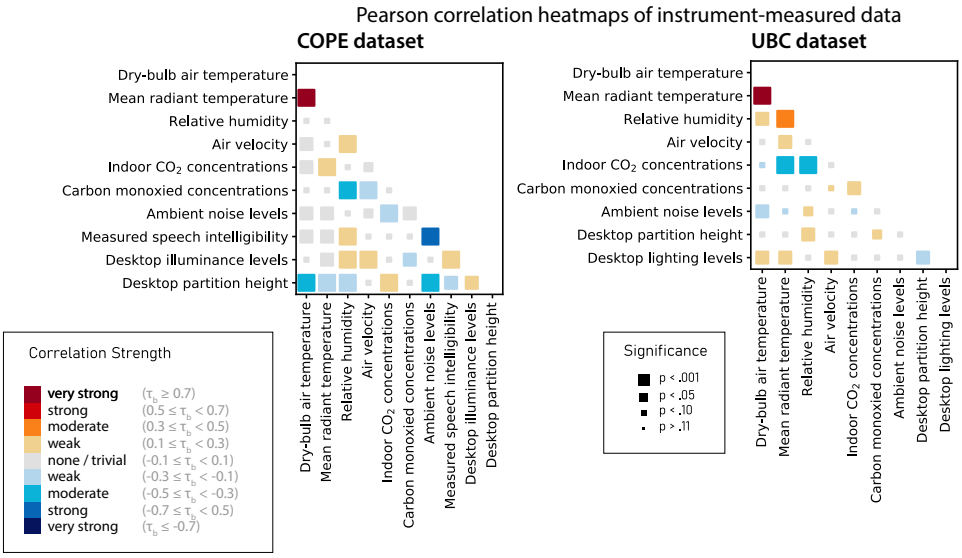

Figure 3.12: Pearson correlation heatmaps of physical measured IEQ data in the COPE and UBC studies

3.3.4 Model Comparison, Selection, and Validation Checks

Calculated ELPD PSIS-LOO scores are used to evaluate the fitness and rank of all of the model cases investigated (see Table 3.4 for original list). The mean scores for each developed model as well as the standard errors of the scores are shown in Table 3.5. The models are ranked in ascending rank (first row is best, last row is worst) according to the absolute mean value of the PSIS-LOO scores for the models trained on the UBC+COPE dataset. Models with lower mean scores than that of the Null hypothesis are displayed in green and the Null hypothesis is shown in red. The ELPD PSIS-LOO scores of all the candidate models trained on the COPE dataset using the prior non-hierarchical modelling are also recalculated from the results presented in Chapter 2 and presented in Table 3.5. The model(s) with the smallest mean scores (in absolute terms) are deemed likely to predict out-of-sample

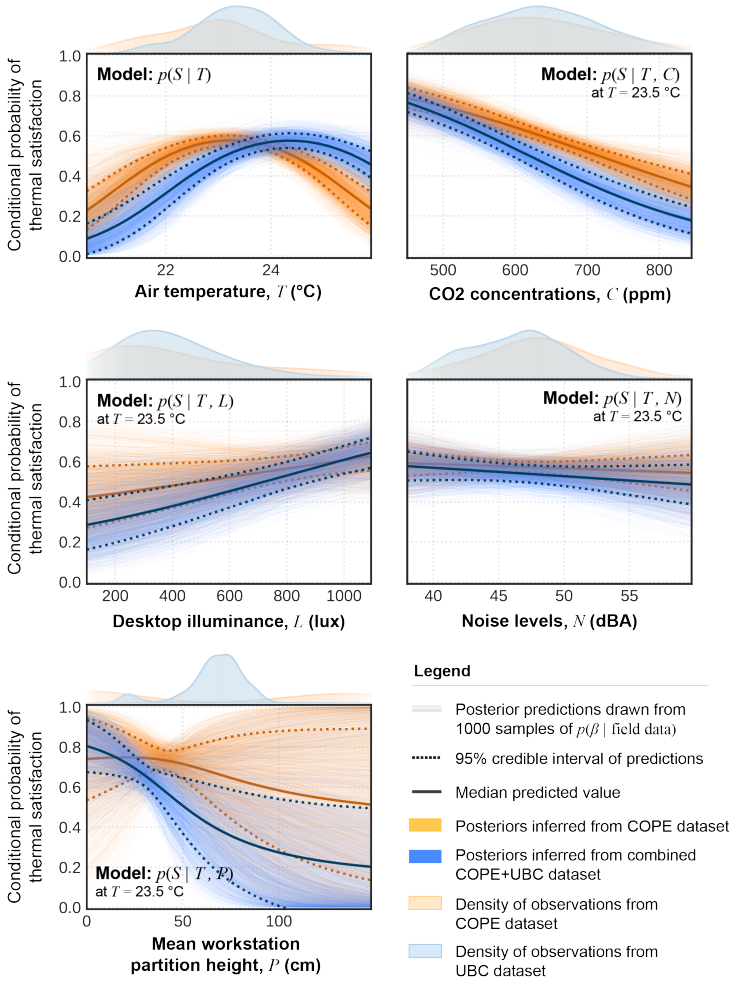

Figure 3.13: Posterior predictions of the probability of thermal satisfaction $p(S|T,\mathbf{W})$

data more accurately than other candidate models with higher values. The scores between the non-hierarchical and hierarchical models of the COPE-only datasets can be compared directly to each other, for comparing models trained on the COPE dataset to models trained on the UBC dataset, Figure 3.14 visualizes a normalized rank of the PSIS-LOO scores. Values of the ELPD PSIS-LOO scores are normal-

Bayesian Model	COPE (non-hierarchical)		COPE (hierarchical)		UBC		COPE+UBC	
	Score	Standard error	Score	Standard error	Score	Standard error	Score	Standard error
$p(S \mid T,C)$	-500.8	2.2	-487.0	7.0	-98.6	4.7	-581.7	8.1
$p(S \mid T,C,P,N)$	-491.9	4.6	-485.2	8.2	-99.3	6.1	-583.3	9.5
$p(S \mid T,C,P,N,L)$	-491.9	4.7	-491.0	8.7	-100.8	6.5	-590.0	9.8
$p(S \mid T,P)$	-495.2	3.9	-492.7	6.9	-98.8	5.1	-592.2	7.7
$p(S \mid T)$	-501.9	1.5	-495.1	5.8	-99.6	3.6	-593.5	6.3
$p(S \mid T,N)$	-502.0	1.8	-497.0	6.0	-101.9	3.9	-595.4	6.5
$p(S \mid T,L)$	-502.4	1.7	-500.2	6.1	-100.6	4.4	-595.7	6.7
$p(S \mid T,R)$	-502.9	1.5	-495.7	6.0	-101.4	4.1	-596.1	6.6
$p(S \mid T,R,V)$	-502.6	1.8	-496.9	6.5	-102.7	4.4	-596.9	6.6

Table 3.5: ELPD PSIS-LOO scores of models trained on the COPE, UBC, and COPE+UBC datasets. The Null hypothesis is shown in red.

ized by dividing each score by the score of the Null hypothesis in each dataset class. For example, the PSIS-LOO score for $p(S \mid T,C)$ as trained on the UBC dataset is divided by -99.6. The standard errors of the scores are also normalized in a similar manner.

On the basis of Table 3.5 and Figure 3.14, one observes that only two candidate models appear to credibly fit better to out-of-sample field observations of thermal satisfaction than the Null hypothesis: $p(S|T,C)$, $p(S|T,C,P,N)$. Of these two, $p(S \mid T,C)$ is preferred on the basis that the mean and standard error of its PSIS-LOO score is unambiguously lower than the mean score of the Null hypothesis. This is at least the case for models inferred on the COPE dataset and the combined COPE & UBC dataset.

3.3.5 Drawing Posterior Predictions from the Models

Figure 3.15 represents a visualization of posterior predictions of perceived thermal satisfaction against both indoor air temperature, T, and indoor CO_2 levels, C, as

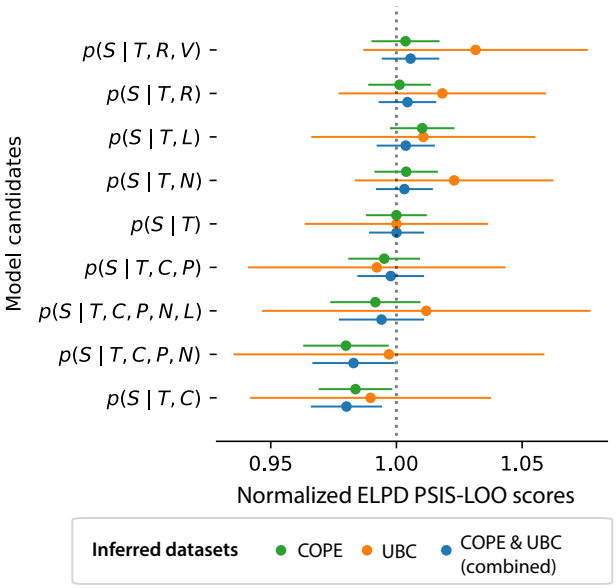

Figure 3.14: ELPD PSIS-LOO scores for each candidate model, normalized around the scores for each dataset's Null hypothesis, $p(S \mid T)$; solid lines depict the standard errors of the mean scores

inferred from the COPE and UBC combined dataset. The predictions are drawn from posteriors of the best performing model, $p(S \mid T, C)$ as identified in the prior section. The visualization is a heatmap of thermal satisfaction predictions, where red colours represent a higher likelihood of perceived thermal dissatisfaction and green and blue colours represent the highest likelihood of perceived thermal satisfaction. The center figure visualizes maximum a posteriori estimates (MAPE), and at either side of the figure is visualized the 95% credible intervals of predictions. As an example, a sample posterior prediction of thermal satisfaction at $T = 23.5\,^{\circ}\mathrm{C}$ and $C = 550$ ppm is evaluated and shown in the figure.

Figure 3.16 represents a visualization of posterior predictions of thermal satisfaction against both operative temperature and indoor CO_2 levels. The standard

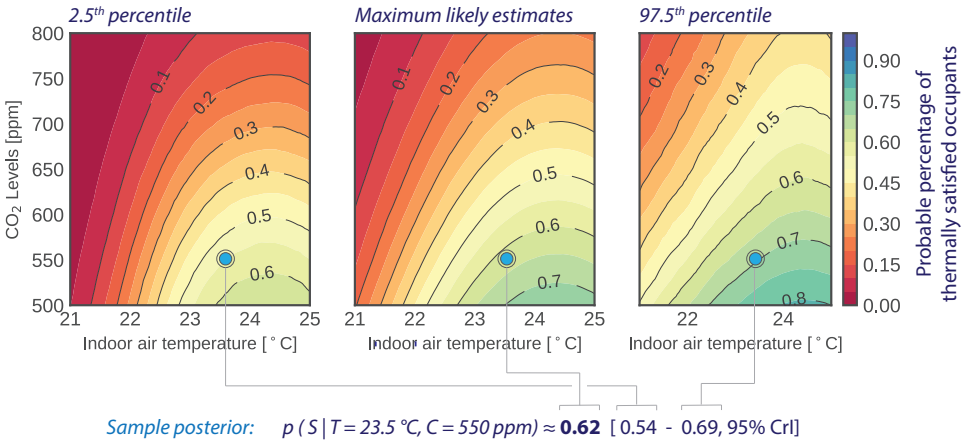

Figure 3.15: Posterior predictions of 'thermally satisfied' occupants as a function of indoor air temperature and CO_2 concentrations (COPE and UBC datasets combined)

deviation of posterior predictions is denoted by the translucent band around the median. Operative temperature is assumed to be the average of indoor air temperature and mean radiant temperature.

As observed from Figure 3.16, the independent negative correlation between CO_2 levels and perceived thermal comfort is notable. Specifically, between an indoor air CO_2 concentration of 500 ppm and 900 ppm, and at an operative temperature of 23 °C, the mean likelihood of an occupant feeling thermally satisfied decreases from 0.58 to 0.46. These are not trivial ranges for the indoor CO_2 levels and predicted effects on thermal satisfaction; as discussed in the previous chapter, 900 ppm is still well within good indoor air quality CO_2 limits for buildings ([97]).

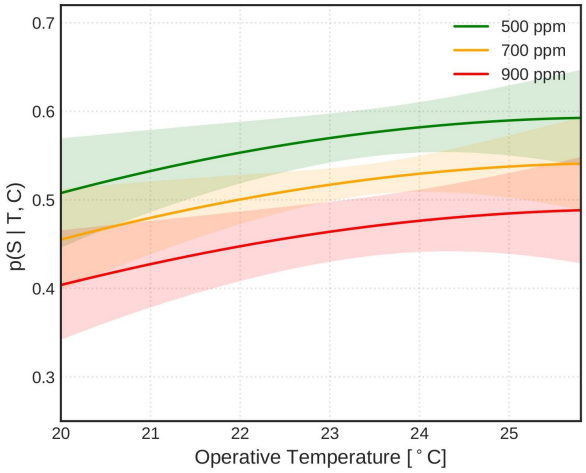

Figure 3.16: Posterior probabilities of the effect of different CO_2 levels on the relationship between thermal satisfaction $p(S)$ and operative temperature

3.4 Discussion

3.4.1 Similarities and Differences between the COPE and UBC datasets

One of the objectives of this second case study was to generate a new dataset, produced via the UBC field campaign, that would serve as an acceptable extension of the original COPE dataset. The immediate use of this dataset would be to undertake Bayesian inference of a model of predicted thermal satisfaction as a function of measured variables in the dataset. It was therefore desired, though not necessarily expected, that the new dataset would possess statistical similarity to the COPE dataset, even if such similarity would be tempered by the comparatively smaller total sample size of the UBC study.

Figure 3.8, which provides an initial glance at comparing statistical differences between the UBC and COPE datasets shows that the datasets are similar with respect to some observed parameters and different with others. With regards to measured air temperature, desktop light levels, ambient noise levels, perceived noise and desktop light satisfaction, the two datasets are similar to one another (t-test $p > 0.05$). This is not the same for measurements of indoor air CO_2 concentrations or perceived air quality, where both appear to have notable differences between the COPE and UBC datasets (t-test $p < 0.001$). Figure 3.4 visualizes some of theses differences directly. For example, though the aggregate distributions for CO_2 measurements in the UBC and COPE datasets are similar, the UBC dataset has a slightly greater proportion of measurements in excess of 800 ppm. These observations are not deemed problematic for this study, as the Bayesian multi-level modelling process is equipped to account for outlier effects. It nevertheless helps to contextualize the subsequent discussion.

On comparing the measured, subjective (survey-derived) variables to one other (see Figures 3.9 and 3.10), the COPE and UBC datasets appear to illustrate similar patterns of relationships. For example, in both the COPE and UBC datasets, responses to questions regarding one's satisfaction with noise, visual privacy, and workspace aesthetics appear to be negligibly or weakly correlated with perceived air quality and satisfaction with air temperature, as shown in the Kendall-τ_b correlation analysis results of subjective data collected in the COPE and UBC datasets (Figures 3.9 and 3.10 respectively). Other perceived factors are strongly correlated with each other, however. In both studies, the response to a question of perceived air quality satisfaction appears to be correlated with perceived thermal satisfaction, satisfaction with light quality, and to a lesser extent, perceived satisfaction

with noise quality. In both studies, perceived thermal satisfaction appears to be correlated mainly to the question of perceived air quality and not other perceived factors. These are not novel findings as prior studies have observed similar conditions [11, 72, 76].

On inspecting potential correlations between subjective and objective (instrumented-measured) variables (Figure 3.11), the similarities between both datasets continue. There are few relevant correlations between any parameter, at least as far as estimates of the Kendall-τ_b coefficient are concerned. It is possible that data sample size is a potential issue in both studies as the significance for the majority of potential correlations is poor.

With regards to comparing measured variables (Figure 3.12), though some similarities exist, differences between both datasets are also more evident than in the above assessments. Strong correlations of MRT and air temperature were expected and are evident, particularly in the COPE dataset and in line with studies discussed previously in section 3.2.2. Measurements of indoor air CO_2 concentrations appear to have a unique relationship to other metrics in each of the datasets. In the COPE dataset, measured concentrations of indoor CO_2 levels appears to have no or trivial relationships with all other measured parameters. In the UBC dataset, even despite the relatively smaller size of the dataset, measured indoor CO_2 concentrations appear to be moderately inversely correlated to measured mean radiant temperature, absolute humidity and relative humidity. This relationship may be attributed to the effect of study season on the data. Approximately 50% of the UBC datasets measurements where taken in the summer period, the other in the winter period. As the majority of surveyed buildings in the UBC study were naturally-ventilated, they experienced higher indoor humidity levels in the summer than in the winter,

and likewise lower indoor CO_2 concentrations due to higher ventilation rates. It is plausible that indoor MRT was also higher in the summer on account of the effect of solar gains along surrounding façades.

Further investigation of statistical correlations across the COPE and UBC datasets may be valuable, though it is not the essence of this thesis. The overarching observation is that, despite some differences, the two datasets are not so dissimilar that they could not be combined into a single expanded field study of open-concept offices. There are no significant anomalies.

3.4.2 Evidence in Support of the Bayesian Regression Models

In comparison to the prior work performed on the COPE dataset and presented in Chapter 2, the hierarchical Bayesian logistic model not only appears to possess a higher degree of fitness to the underlying data but also reveals a higher degree of certainty on the significance of individual model parameters in predicting $p(S)$ even in the COPE dataset. This finding was not unexpected in light of other domain studies that have witnessed the same relative performance improvement of multilevel modelling [117, 136].

In general, the prior findings appear to be reinforced and strengthened in the new UBC study. Out of all candidate models of thermal satisfaction, $p(S)$, that are conditional on only measurements of thermophysical properties (air temperature, relative humidity, indoor air velocity), $p(S \mid T)$ remains the model with the highest fitness to the data. It has the smallest PSIS-LOO score of all applicable candidates, and this is observed in both datasets independently as well as in the combined datasets. It remains an appropriate Null hypothesis for this study. On adding some non-thermal parameters to the prediction of $p(S)$, such as noise lev-

els, and illuminance levels, models $p(S|T,N)$ and $p(S|T,L)$ respectively, there is evidence of some improvement in model fitness, as presented in Table 3.5 and Figure 3.14. However, like in the first case study, the evidence basis in support of this improvement remains inconclusive for all parameters except indoor CO_2 concentrations.

Compared to the prior results, one observes a revised indicator suggesting CO_2 is a credible predictor of perceived thermal satisfaction. This is observed not only through Figures 3.13 and 3.15, but critically through PSIS-LOO scores. As the specific value of the PSIS-LOO score is also affected by dataset size, it may not be easy to assess differences between model performance, chosen variables, and chosen training data in Table 3.5. Yet, the PSIS-LOO score of $p(S \mid T,C)$ inferred from the COPE dataset is sufficiently smaller than the score of the Null hypothesis, even in consideration of the standard error of the score (see Table 3.5 and Figure 3.14).

Though inferring $p(S \mid T,C)$ from only the UBC dataset does not reveal such a strong indicator in support of the candidate model, the mean PSIS-LOO score of this model is still lower than the Null hypothesis. More so, when combining both the UBC and COPE datasets together, the PSIS-LOO score reduces further. Ultimately, in all cases, $p(S \mid T,C)$ is observed to be the best model explaining the underlying field data.

3.4.3 Comparison to Results from First Case Study

It is observed from the results in Table 3.5 that re-applying the assessment of the COPE dataset [27, 31] using a *hierarchical* Bayesian logistic regression model reduced the ambiguity of the regression results and in doing so improved the evi-

dence in support of individual candidate models. By specifying a second-stage distribution that described how building-specific parameters vary within the dataset, the hierarchical models could better account for observed and unobserved sources of heterogeneity. Comparing the PSIS-LOO scores of the hierarchical and non-hierarchical Bayesian models trained on the COPE dataset, apart from models including partition height P, the hierarchical models are consistently preferred over the non-hierarchical models; i.e., the ELPD PSIS-LOO score for hierarchical models is credibly smaller (in absolute terms) than for the non-hierarchical models.

Additional anomalies between the hierarchical and non-hierarchical models are observed that also substantiate the potential advantages of hierarchical modelling. Using the non-hierarchical modelling approach of the prior work, models including partition height, P, would appear to be credibly fit better to the field data than models not including partition height. This is not the case for the hierarchical models. One observes in Figure 3.4 that recorded partition heights in the COPE study were heavily clustered by building type. If one allows the possibility that random (unknown) effects, and not partition height, could account for differences in perceived thermal satisfaction between individual buildings - which is the assumption of the hierarchical model - it becomes less clear that partition height is independently correlated to observed differences in thermal satisfaction. There is simply not enough variation in P within each building's field data to be confident that partition height could be an independent predictor of $p(S)$. This same feature of hierarchical modelling leads to one observing greater certainty on the relationship between measured CO_2 concentrations and predicted thermal satisfaction. Building B9 of the COPE dataset was observed to have uncharacteristically high measurements of indoor CO_2 levels while simultaneously having the greatest

proportion of thermally satisfied occupants than any other building. Assuming that it is possible that B9's observed conditions could be attributable to some unknown random effect, the more-aligned relationship between CO_2 and thermal satisfaction, as observed in the other buildings, is treated as more indicative of the true relationship. Effectively, with respect to $p(S \mid T, C)$, the posterior distributions of the hierarchical model considers B9 to be an outlier, suggesting that the influence of B9's field results on the true value of β_C is relatively small compared to other buildings. It was observed that building B9 featured a slightly reduced number of data points compared to the rest of the buildings in the COPE dataset (56 data points and the average per-building sample size in the dataset is 86.4). This might contribute to the observed anomaly, although this remains speculation.

Chapter 4

Building Energy Model of an Office Space based on Indoor CO_2 and Temperature Controls

4.1 Introduction

The posterior results inferred from the COPE and UBC field studies, presented in Chapter 2 and 3, revealed that indoor CO_2 concentrations levels appear to be strongly correlated with perceived thermal satisfaction. Model checks and validation techniques suggested that including measurements of CO_2 increases the prediction accuracy of thermal satisfaction modelling in open-plan offices [27, 28, 30, 31].

In this chapter, the new predictive model of thermal comfort, derived from the hierarchical Bayesian regression of the combined COPE and UBC datasets is

adapted and integrated into a BEM that simulates an open-concept mechanically-ventilated office space located in Vancouver. The model predicts occupants' thermal satisfaction and heating energy consumption as a function of setpoint thermal conditions and indoor CO_2 concentrations.

The building energy model aims to investigate the potential energy costs resulting from pumping higher amounts of fresh air indoors while lowering the heating setpoint and maintaining the same levels of thermal comfort. A 241.5 m^2 open-plan mechanically-ventilated office space located in Vancouver is simulated in TRNSYS for the month of January. The control system is set so that the indoor air temperature and the indoor CO_2 levels are maintained fixed. Different configurations of both the heating setpoint and the indoor CO_2 setpoint are examined. The corresponding heating demand and thermal satisfaction are calculated for each examined scenario in order to examine possible energy-savings scenarios to increase the air change rates while not compromising the occupant's thermal comfort.

4.2 Proposing a New Predictive Thermal Comfort Model for Building Controls

In this section, a new predictive model of thermal comfort, derived from the hierarchical Bayesian regression of the COPE and UBC datasets, presented in Chapter 3, is formulated. The proposed model predicts the probability of occupants' thermal satisfaction as a function of indoor air temperature, T, and indoor CO_2 concentrations levels, C, $p(S \mid T, C)$, as follows:

Model Parameter(β)	MAPE	95% CrI
β_o	0.15	CrI [-0.22 : 0.5]
β_T	1.3	CrI [0.093 : 2.2]
β_{T^2}	-0.25	CrI [-1.8 : 1]
β_{CO_2}	-0.74	CrI [-1.4 : -0.099]

Table 4.1: Maximum a posteriori estimates (MAPE) of each model parameter for $p(S \mid T,C)$ model with 95% Credible intervals (CrI)

$$p(S \mid T,C) = \frac{1}{1+ e^{-\Gamma}}$$

$$\Gamma = [(\beta_T \,.\, T) + (\beta_{T^2} \,.\, T^2) + (\beta_C \,.\, C) + \beta_o] \quad (4.1)$$

Where T = air temperature (°C), C = indoor CO_2 levels (ppm), β_C, β_T, and β_{T^2} are the model parameters for C, T, and T^2 respectively, and β_0 is the constant coefficient. C is modelled as having a first-order linear relationship with $p(S)$ and T is modelled as having a quadratic relationship with $p(S)$.

The maximum a posteriori estimates (MAPE) and the 95% Credible Interval (CRI) of each model parameter (β) are summarized in Table 4.1. MAPE method is a full Bayesian parameter estimation approach in which a prior is specified in the estimate procedure, unlike the Frequentist's Maximum likelihood Estimation (MLE). The parameter's uncertainty is expressed by a prior, and the posterior is evaluated using the knowledge in the prior and the information in the likelihood [93].

Figure 4.1 displays the model parameters' posterior probability distributions (β) drawn from the Markov Chain Monte Carlo sampling. The maximum a posteriori estimation (MAPE) and the 95% credible interval (highest posterior density,

HDI, interval) of each β are displayed. The significance of the effect of the model parameter on thermal satisfaction increases as the maximum a posteriori estimation (MAPE) deviates from 0 (displayed as a vertical orange line).

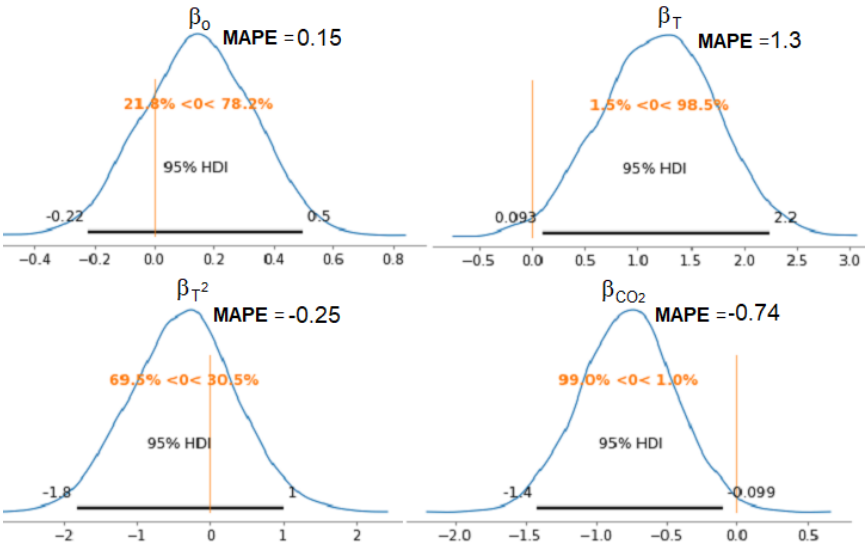

Figure 4.1: Posterior traces of the $p(S \mid T, C)$ model parameters (β). For each model parameter: its maximum a posteriori estimation (MAPE) is displayed along with the corresponding 95 % credible interval.

Figure 4.1 shows that the CO_2 model parameter credibly deviates from 0, which is evident from the $(p(\beta_C) < 0) = 99.0\%$. This observation reinforces the prior findings and suggests that predictions of thermal satisfaction are highly influenced by measurements of indoor CO_2 levels.

It should be noted that the proposed Bayesian thermal comfort model is a probabilistic model which predicts the thermal satisfaction as a probabilistic distribution with model parameters' uncertainty bounds presented above. For future use of the model, it is recommended that the model is used to sample from, using Markov chain Monte Carlo (MCMC) or other sampling method an adequate number of times

until the predicted outcomes are well-converged, instead of using its deterministic form for more accurate predictions. In this thesis, 5000 draws are used to infer the posterior predictions of the logistic regression model parameters. It is recommended, for observational studies that involve logistic regression in the analysis, to have a minimum sample size of 500 to sufficiently derive the statistics that represent the parameters [13]. If using PyMC3, the recommended default number of draws in logistic regression is 1000 draws.

4.2.1 Model Validation

Model validation is an essential part of the model development process. It aims to validate the model's robustness and accuracy, verify that the model is not over-fitting or under-fitting the data, and prove it is not an ungeneralized model (i.e. model's predictions cannot be generalized on other data). In statistics and machine learning, model validation is usually done by splitting the data into two subsets: training dataset and testing dataset. The model is first fit on the training data in order to make predictions on the test data ([120]). To validate the model, it was first trained by fitting it to the COPE dataset to infer predictions for the model parameters. Using the UBC dataset as the testing data, predictions of thermal satisfaction are evaluated and compared against true observations drawn from the test data.

The probability distribution of the predicted thermal satisfaction, $p(S)$, inferred from the training data is shown in Figure 4.2. The maximum likely predicted $p(S)$, displayed as a blue dashed vertical line, is compared against the mean of the true $p(S)$, observed from the test data and displayed as an orange dashed vertical line. The 95% credible interval of the predicted $p(S)$ is shown in Figure 4.2 as a grey shaded area. The model validation process, displayed in Figure 4.2, reveals that

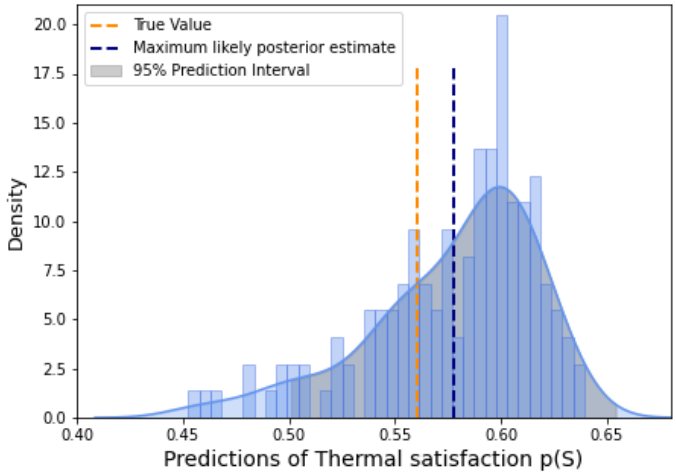

Figure 4.2: Comparison of posterior predictions of $p(S \mid T,C)$ to true value for the test dataset (95% prediction interval shaded in grey , true $p(S)$ displayed as orange dashed line, maximum likely estimate of $p(S)$ displayed as blue dashed line

the true value of $p(S)$ lies within the 95% credible interval of predictions of thermal satisfaction with maximum likely predicted $p(S)$ = 0.578 and mean observed $p(S)$=0.56 (percentage prediction error = 3.2%).

4.3 Developing a New Building Energy Model

The predictive thermal comfort model is integrated into a TRNSYS [10] framework to investigate the potential energy savings that can be achieved by using the model to control both the heating setpoint and the required fresh air without compromising the occupant's thermal satisfaction. TRNSYS is commonly used to model and simulate the building energy consumption allowing for a transient simulation of each individual mechanical system within the building [66]. The building model is first constructed using Google SketchUp [1] to build the geometry of the office

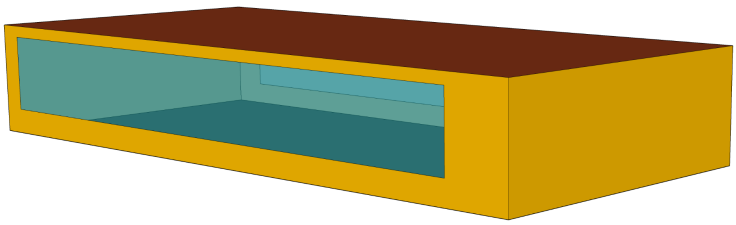

Figure 4.3: The simulated open-plan office geometry

space, specify and separate the air nodes, configure the building's surface types (walls, floor, roof, ceiling) as well as specify the boundary conditions of each zone. TRNBuild [67] is then used to specify and configure the building characteristics such as zones' ventilation types, internal gains, heating and cooling system types, infiltration types, and constructions materials and specifications.

An open-plan office in the Centre for Interactive Research on Sustainability (CIRS), a four-storey building on the campus of the University of British Columbia (UBC), is selected as a case study for this work. The office, illustrated in Figure 4.3, is modelled as a single-zone with a floor area of 241.5 m^2 and a floor-to-ceiling height of 3.7 m. The model has two exterior walls exposed to the outside and two walls with adiabatic zone boundary conditions (i.e. no heat transfer between the zone and its surrounding is assumed).

The office is located on the 3^{rd} floor and has two large windows on the north and south sides of the building, as shown in Figure 4.3. The fenestration area is 50 m^2 per side which is equivalent to a 0.58 window-to-wall ratio (WWR). The windows have a U-value of 1.1 $W/m^2.k$ and a G-value of 0.62. The external walls' construction has a thickness of 0.377 m and a U-value of 0.25 $W/m^2.k$. The floor and ceiling constructions of the zone are modelled as adiabatic boundary conditions

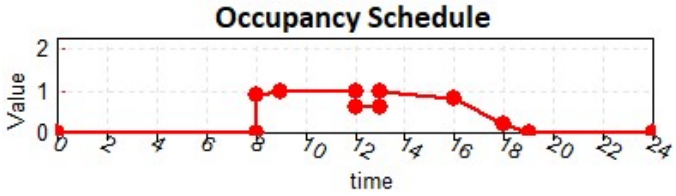

Figure 4.4: The daily schedule used to scale the internal heat gain and indoor CO_2 production rates

with the outside.

The internal heat gains generated by the occupants are modelled in TRNBuild by specifying a maximum thermal gain generated by the occupants and scaling this internal gain throughout the day using a predetermined office occupancy schedule. The occupants' maximum internal heat gain is 75 W convective gain and 75 W radiative gain. These gains are scaled using the occupancy schedule shown in Figure 4.4. It is assumed that 25 people occupy this office, which is equivalent to an occupancy density of 10 m^2/person.

A lighting control system is configured so that the lights are turned on if the occupancy is greater than 0, and turned off otherwise. The radiative heat gain from lighting is assumed to be equal to 25.2 KJ /hr /m^2. The heat gain produced by the electrical equipment heat gain is assumed to be 43.2 KJ / hr / m^2, and is scaled using the occupancy schedule displayed in Figure 4.4.

4.3.1 Simulated HVAC System and Control Strategy

The HVAC system of the centre for interactive research on sustainability building (CIRS) consists of a ducted-air Variable Air Volume (VAV) heating system with operable windows for natural ventilation cooling. The HVAC system also has a Heating Recovery Ventilation (HRV) system for heat recovery in the heating sea-

son. The control system used in this building energy model is a temperature and CO_2-based control system, where both the heating setpoint and CO_2 set point are controlled.

The building energy model investigated in this thesis uses the CIRS building as a case study, yet the model itself has been simplified to focus on the integration of the predictive thermal comfort model into the building control system. Hence the building energy model can be considered a theoretical model that aims to explore the significance of the newly developed model and its impact on building control systems.

A total of 36 simulations are undertaken in the building energy model, with simulation corresponds with a unique indoor air temperature setpoint and indoor CO_2 setpoint. The range of setpoints covered spans an air temperature setpoint of between 21 ^{o}C and 25 ^{o}C and CO_2 levels of between 500 ppm and 1000 ppm. This broadly covers the range of observed indoor temperatures and CO_2 levels in the prior UBC and COPE datasets. Under each simulation, a setback air temperature setpoint of 15 ^{o}C is assigned and the ventilation system is deactivated during unoccupied periods.

The instantaneous indoor CO_2 concentration, $C_{t=n}$, is calculated as a function of the ventilation rate as follows:

$$C_{t=n}[ppm] = \frac{VC_{t=n}}{V_{zone} * 1000} * 10^6 \tag{4.2}$$

Where V_{zone} is the volume of the zone in m^3 (V_{zone} = 893.55 m^3) and $VC_{t=n}$ is the instantaneous volume of indoor CO_2 in litres and calculated as follows:

$$VC_{t=n} = \Delta t[\frac{1}{s}] * ([V^{\cdot} + V_{inf} * (C_{out} - C_{in,t=n-1})] + [VC_{gen} + VC_{t=n-1}]) \quad (4.3)$$

Where $V^{\cdot}$ is the ventilation rate in L/s, V_{inf} is the infiltration rate in L/s, C_{out} is the outdoor CO_2 concentrations level, $C_{in,t=n-1}$ is the indoor CO_2 concentration at time=n-1 (i.e. at time=current time - Δt), $VC_{t=n-1}$ is the volume of indoor CO_2 in litres at the previous time step, and VC_{gen} is the occupants' CO_2 generated rate, in litres per second, and is calculated as follows:

$$VC_{gen}[L/s] = \frac{0.31}{60} * N_p \quad (4.4)$$

N_p is the number of occupants. The assumed CO_2 generated rate per person is adopted from ASHRAE Standard 62.1(2007) for an activity level of 1.2 met units corresponding to sedentary persons.

The heating energy demand for the month of January, the coldest month in the year in Canada, is calculated for each scenario using TRNSYS. The mechanical ventilation system is assumed to supply a uniform indoor air temperature. The occupants' thermal satisfaction is evaluated for each combination of indoor air temperature and CO_2 concentration level using the Bayesian predictive thermal satisfaction model in Equation 4.1.

4.4 Building Simulation Results and Discussion

A total of 36 simulations were completed, each simulating the office space for the month of January. Simulation time-step sizes were 5 min. Table 5.2 summarizes the raw data resulting from the 36 simulations performed for each different sce-

C / T	20 °C	21 °C	22 °C	23 °C	24 °C	25 °C
500 ppm	10.13039	11.88981	13.58193	15.10081	16.55359	17.99521
600 ppm	8.510738	9.742854	11.04089	12.45306	13.77039	15.07612
700 ppm	6.024908	6.444541	7.082668	8.184533	9.264256	10.34188
800 ppm	5.157384	4.619636	5.259302	6.206163	7.167583	8.134288
900 ppm	4.517593	3.677084	4.236762	5.116224	6.007772	6.905814
1000 ppm	4.502112	2.952077	3.66343	4.484804	5.327326	6.180053

Table 4.2: Simulation results: Monthly heating energy demand [KWh/m^2] for 36 scenarios of indoor air temperature setpoint and indoor CO_2 setpoint

nario. The table displays the monthly heating energy demand in [KWh/m^2] for each scenario of indoor air temperature setpoint and indoor CO_2 setpoint. Figure 4.5 displays the aggregate simulation results in the form of a contour plot. The plot, generated via MATLAB, interpolates data produced by the simulations, specifically monthly heating energy demand versus indoor air temperature setpoint and indoor CO_2 setpoint. The heating energy demand values evaluated for different configurations of indoor air temperature and CO_2 levels are displayed as solid black contours.

The predictions of thermal satisfaction, stemming from the developed hierarchical Bayesian model, are visualized as a heatmap where dark red colours represent higher likelihoods of perceived thermal dissatisfaction and green colours represent a higher likelihood of thermal satisfaction. The median observed values for both the indoor air temperature and indoor CO_2 levels for the COPE and UBC datasets are also shown.

The simulation results, shown in Figure 4.5, suggest that there might be multiple pathways for increasing building ventilation rates indoors which affect energy demand and thermal comfort differently. For instance, for building managers and

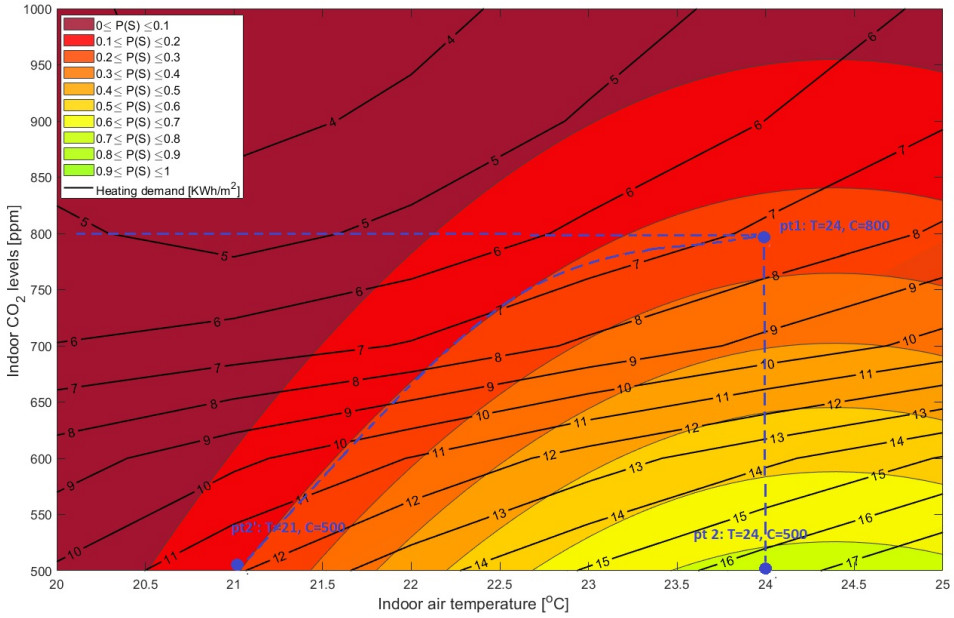

Figure 4.5: Heating energy demand in [KWh/m^2] as a function of indoor air temperature and indoor CO_2 levels (denoted as black solid contours). Predictions of thermal satisfaction, p(S), represented as a heat map: dark red represents higher thermal dissatisfaction and green colour represents higher thermal satisfaction levels.

operators that are under pressure to increase the current air change rates, if the conditioned space is currently maintained at an indoor air temperature of 24 ^oC and a CO_2 level of 800 ppm (point 1), one solution may be to increase ventilation rate such that the indoor air temperature is kept at 24 ^oC and the CO_2 level is lowered to 500 ppm (point 2). In that case, the total energy demand for heating and ventilation would be increased by 131 % but perceived thermal satisfaction would increase as well.

However, if one would seek to only maintain current levels of perceived thermal satisfaction, building managers could choose to decrease indoor heating setpoints to 21 ^oC while lowering the indoor CO_2 concentrations to 500 ppm (point

2'). In this case, the increase in energy demand might only be increased by about 66%. Multi-contextual thermal comfort models, like that presented in Eq. 4.1 and explored in this chapter, may be used by building managers to make changes to the building's control system so that fresh air rates can be increased with minimal energy demand increase and no effect on occupant thermal comfort.

While the findings of this chapter suggested that it is possible to have adequately-ventilated office spaces with minimal increase in heating energy use and with unchanged levels of thermal comfort, these results should reflect the prevailing hybrid work settings as well. Since the COVID-19 pandemic, the number of office workers who work from home in a hybrid mode has increased [57]. The following chapter will examine different occupancy schedules that represent the post-COVID-19 hybrid work environments in office spaces. The heating energy consummation will be evaluated for different configurations of setpoints and for several occupancy profiles to investigate possible energy savings that result from pumping higher amounts of fresh air while lowering the heating setpoint.

Chapter 5

Increasing Ventilation Rates and Energy Efficiency in Post-COVID-19 Buildings

5.1 Introduction

In the wake of the COVID-19 pandemic, the office working setup has greatly changed with more workplaces moving towards hybrid working models and teleworking, where office occupants are no more working 100% of their full-time capacity at the office. Different hybrid working approaches have been recently adopted in workplaces and the current, post-COVID-19, occupancy schedules need to be reflected in building energy models. This is especially of importance when the heating energy demand of office spaces is studied. Another key change in office buildings after the COVID-19 pandemic is the need to increase the amount

of fresh air requirements to mitigate the risk of airborne virus transmission. Various researchers and organizations, including the World Health Organization (WHO) [138] , recommended increasing the percentage of indoor fresh air and the current values of air change per hour (ACH) to help reduce the spread of COVID-19 in indoor environments, which means that the indoor CO_2 levels setpoint needs to be lowered [3, 12, 103]. It has been recommended to increase current ventilation rates to 4 to 6 air changes per hour [3]. The added cost of moving more air as well as heating a larger volume of air poses a considerable challenge to building managers and operators [3].

In the previous chapter, the new Bayesian hierarchical thermal comfort model inferred from the COPE and UBC datasets was adapted and integrated into a building energy model. The control system of the BEM was set so that the indoor air temperature and the indoor CO_2 levels are maintained fixed. In this chapter, several occupancy profiles have been examined to reflect and compare the current occupancy schedules in office spaces after the COVID-19 pandemic. Different configurations of both the heating setpoint and the indoor CO_2 setpoint are examined. The corresponding heating energy demands and thermal satisfaction are calculated for each investigated occupancy profile for a combination of indoor air temperature and indoor CO_2 levels setpoints.

Aiming to provide a solution for building managers, who are under pressure to increase current ventilation rates during the COVID-19 pandemic, this chapter aims to investigate whether buildings' control systems can be changed so that current fresh air amount can be increased with minimal energy increase and without sacrificing occupants' thermal comfort, an important challenge to building modellers in a post-COVID-19 era.

Schedule 0 (Pre-COVID-19)	Schedule 1 (Post-COVID-19)	Schedule 2 (Post-COVID-19)	Schedule 3 (Post-COVID-19)
5 days/ week, 100% full capacity	3 days/ week, 100% full capacity	5 days/ week, 50% full capacity	5 days/week, 60% full capacity

Table 5.1: Daily occupancy schedules used to scale the internal heat gain and indoor CO_2 production rates

5.2 Post-COVID-19 Occupancy Schedules

Four daily schedules are designed to reflect and compare possible occupancy profiles for post-COVID-19 back to work settings. Schedule 0 is designed to simulate a 100% full-time occupancy capacity for all of the 5 working days and is used as the baseline to reflect the pre-pandemic occupancy profiles. Schedules 1, 2, and 3 are possible occupancy profiles for post-COVID-19 work settings that involve hybrid working models where the office is never occupied with either the 100% occupancy capacity or the entire 5 weekdays [29]. Table 5.1 and Figure 5.1 summarize the examined occupancy profiles. The weekend occupancy profiles are assumed to have no occupancy for the entire 2 days.

The occupancy schedules, displayed in Figure 5.1, are used to scale the internal convective and radiative heat gain generated by occupants. The office lighting control system of the building energy model has the same configuration presented in section 4.3. Similarly, the heat gains produced from lighting and electrical equipment have the same configuration used in section 4.3 and are scaled using the occupancy schedules in Figure 5.1.

The building energy model developed in TRNSYS and presented in section 4.3 is used to undertake a total of 144 simulations, 36 simulations for each of the four investigated occupancy schedules. Each of the 36 simulations corresponds with

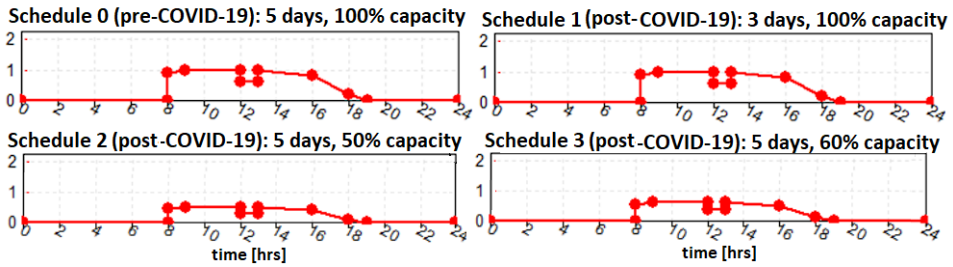

Figure 5.1: Weekdays schedules used to scale the internal heat gain and indoor CO_2 production rates

a unique indoor CO_2 setpoint indoor air temperature setpoint. The range of setpoints covered spans an air temperature setpoint of between 21 and 25 ^{o}C and CO_2 levels of between 500 ppm and 1000 ppm. Under each simulation, a setback air temperature setpoint of 15 ^{o}C is assigned and the ventilation system is deactivated during unoccupied periods. The heating energy demand for the month of January, the coldest month in the year in Canada, is calculated for each scenario and for each of the developed occupancy schedules using TRNSYS. Thermal satisfaction predictions are evaluated for each combination of indoor air temperature and CO_2 concentration level using the Bayesian predictive thermal satisfaction model in equation 4.1.

5.3 Results and Discussion

Monthly heating energy demand, for the month of January, for each of the 36 examined scenario of indoor air temperature and indoor CO_2 setpoints, and for each of the investigated occupancy schedules, are evaluated in [KWh/m^2]. The simulation time-step size is selected to be 10 min for the 144 simulations. For each simulation, daily air changes per hour, indoor CO_2 concentrations levels [ppm], and

C / T	20 °C	21 °C	22 °C	23 °C	24 °C	25 °C
500 ppm	8.355506	9.471294	10.5032	11.46159	12.44644	13.45435
600 ppm	7.251001	8.025247	8.935622	9.822631	10.72242	11.64701
700 ppm	5.56417	5.815475	6.452946	7.193784	7.95787	8.746684
800 ppm	4.831999	4.729806	5.289289	5.959954	6.659504	7.384382
900 ppm	4.40411	4.062286	4.649039	5.27416	5.933471	6.621838
1000 ppm	4.400005	3.732644	4.272521	4.870145	5.50467	6.168761

Table 5.2: Monthly heating energy demand [KWh/m^2] for 36 scenarios of indoor air temperature setpoint and indoor CO_2 setpoint for schedule 1 'post-COVID-19', (3 days/week, 100% full capacity)

C / T	20 °C	21 °C	22 °C	23 °C	24 °C	25 °C
500 ppm	6.163481	9.279357	10.44124	11.62209	12.76425	13.89434
600 ppm	6.243814	6.964488	7.853328	8.744538	9.721083	10.70167
700 ppm	6.340197	5.439579	6.110312	6.854235	7.723479	8.599809
800 ppm	6.340202	4.914611	5.472928	6.211086	7.038243	7.879744
900 ppm	6.340206	4.834376	5.378807	6.105275	6.91999	7.750841
1000 ppm	6.340202	4.834006	5.377166	6.103486	6.918105	7.748857

Table 5.3: Monthly heating energy demand [KWh/m^2] for 36 scenarios of indoor air temperature setpoint and indoor CO_2 setpoint for schedule 2 'post-COVID-19', (5 days/week, 50% full capacity)

ventilation rates [L/s] are calculated for each day in the month of January. Daily values of heating energy demand, in [KWh/m^2], are also evaluated and compared with daily values of outdoor air temperature. Tables 5.2, 5.3, and 5.4 summarize the raw data resulting from the 36 simulations performed for each different scenario of schedules 2, 3, and 4 (post-COVID-19) respectively.

In order to compare the four occupancy schedules (shown in Figure 5.1), the monthly heating energy demands for the three scenarios of increasing the air change rates (scenario 1, 2, 2'), displayed in Figure 4.5, are calculated and displayed in Figure 5.2.

The results, shown in Figure 5.2, reveal that it is possible to increase the indoor

C / T	20 °C	21 °C	22 °C	23 °C	24 °C	25 °C
500 ppm	7.979304	9.759829	10.97843	12.27332	13.48071	8.867627
600 ppm	6.249528	7.51302	8.498482	9.451971	10.49356	8.867995
700 ppm	6.117166	5.602873	6.277244	7.048032	7.955586	8.867995
800 ppm	6.117132	4.768472	5.247532	6.065769	6.910965	8.867995
900 ppm	6.117138	4.407573	4.907643	5.698576	6.513329	8.867995
1000 ppm	6.117137	4.366517	4.858673	5.641803	6.448555	8.867995

Table 5.4: Monthly heating energy demand [KWh/m^2] for 36 scenarios of indoor air temperature setpoint and indoor CO_2 setpoint for schedule 3 'post-COVID-19', (5 days/week, 60% full capacity)

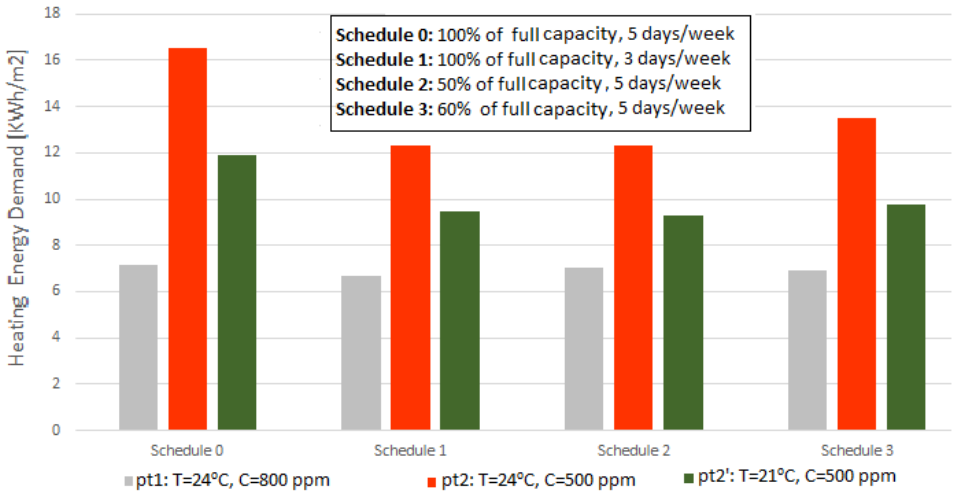

Figure 5.2: Monthly heating energy demands of the investigated occupancy schedules for pt. 1, 2, and 2'.

fresh air amounts while maintaining the same level of occupants' thermal comfort by lowering the heating setpoint and the CO_2 concentration levels. This is applicable to the pre-COVID-19 schedule as well as all the investigated post-COVID-19 occupancy schedules, as shown in Figure 5.2. More specifically, the results show that adopting scenario pt.1 to pt.2 to increase the air change rates, which implies that the indoor CO_2 levels are lowered and the heating set point is maintained,

Schedule	% Increase in heating energy (pt.1 to pt.2.)	% Increase in heating energy (pt.1 to pt.2.'.)
Schedule 0 (5 days, 100%)	130.96%	65.95 %
Schedule 1 (3 days, 100%)	86.897 %	42.22%
Schedule 2 (5 days, 50%)	81.3556%	31.842%
Schedule 3 (5 days, 60%)	95.063 %	41.22%

Table 5.5: Comparison between the percentage increase in monthly heating energy demand for both scenarios of increasing the ventilation rates for the four investigated occupancy schedules

consumes more heating energy than adopting scenario pt.1 to pt.2', in which the heating set point is lowered to 21 ^{o}C and CO_2 concentration levels are lowered to 500 ppm.

Comparing schedules 1 and 3, both schedules correspond to 60% of the pre-COVID-19 working schedule. Schedule 1 represents 100% of the full occupancy capacity with occupants coming only 3 days per week, while in schedule 3, the office workers occupy the office during the five working days but only 60% of the full capacity are present at a time. It is also noteworthy that schedule 1 saves slightly more heating energy than schedule three.

Table 5.5 summarizes the percentage increase in the monthly heating energy demand if building managers and operators can adopt scenario 1-2 vs. scenario 1-2' for increasing the ventilation rates indoors, calculated for the four investigated occupancy profiles. It is revealed from the results that, using the Bayesian predictive model of thermal comfort (presented in Eq.4.1) to increase the amount of fresh air while lowering the heating setpoint and maintaining the same thermal comfort levels, saves more energy than maintaining the indoor temperature setpoints for all the investigated schedules.

The daily values of heating energy demand in [KWh/m^2] for the three investi-

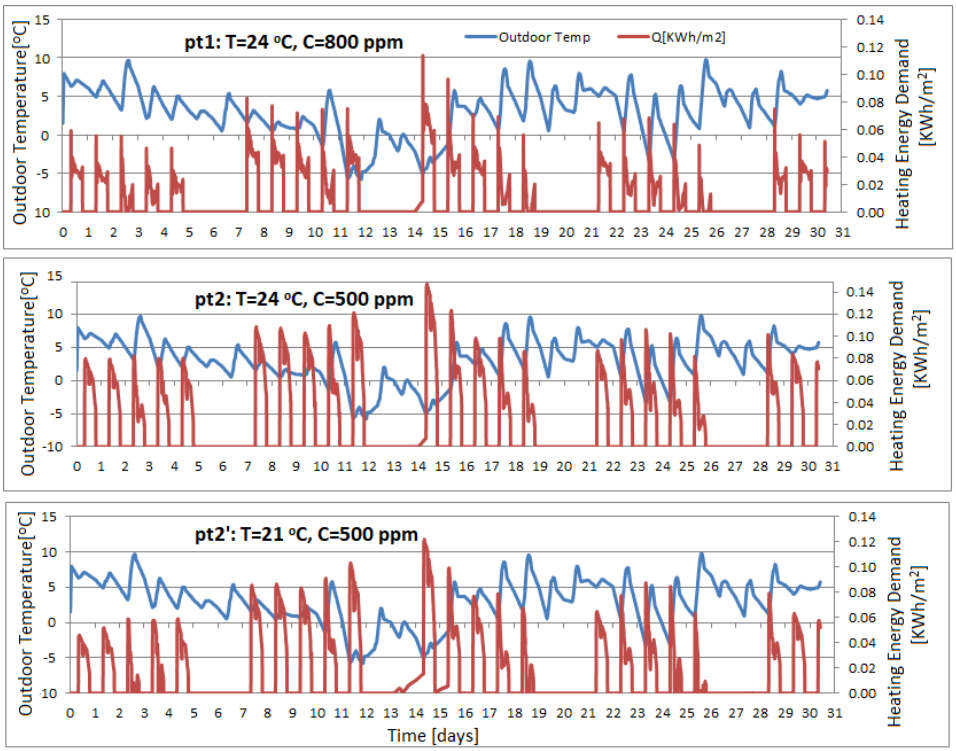

Figure 5.3: Daily heating energy demand [KWh/m2] and outdoor temperature [°C] for the month of January, schedule 0 (pre-COVID-19) for the three investigated points (T=24 °C, C=800 ppm; T=24 °C, C=500 ppm; and T=21 °C, C=500 ppm)

gated scenarios, pt.1 (T = 24 °C and C = 800 ppm), pt.2 (T=24 °C, C=500 ppm), and pt.2' (T=21 °C, C=500 ppm) for the entire month of January are evaluated and displayed along with the outdoor temperature in Figure 5.3. It is noted that the heating energy consumption increases as the outdoor temperature decreases, as seen in Figure 5.3. As the amount of fresh air increases from point 1 to point 2, the heating energy demand increases as well. In scenario 2', however, the heating energy demand is lower than that for point 2 since both points have the same CO_2 concentration setpoint, but point 2' has a lower heating setpoint than point 2, which

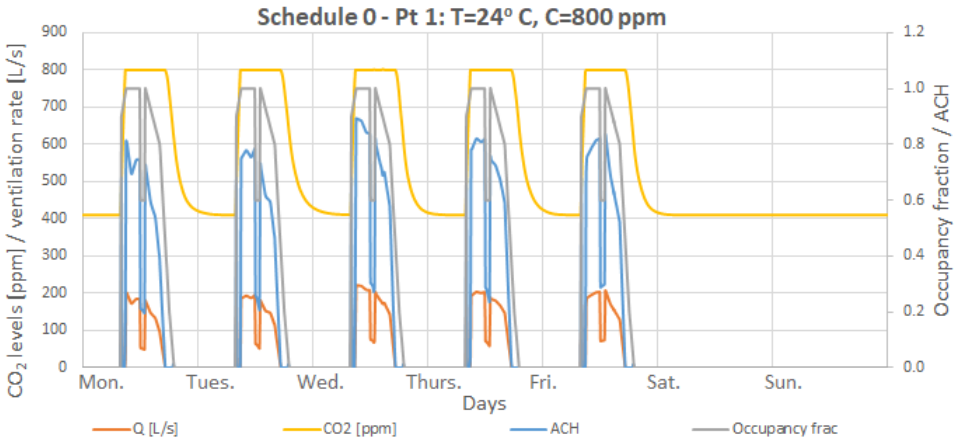

Figure 5.4: Schedule 0 (5 days/week, 100% full-time) CO_2 levels [ppm], ventilation rate [L/s], Occupancy fraction, and Air change rate for pt. 1 (T=24°C, C=800 ppm).

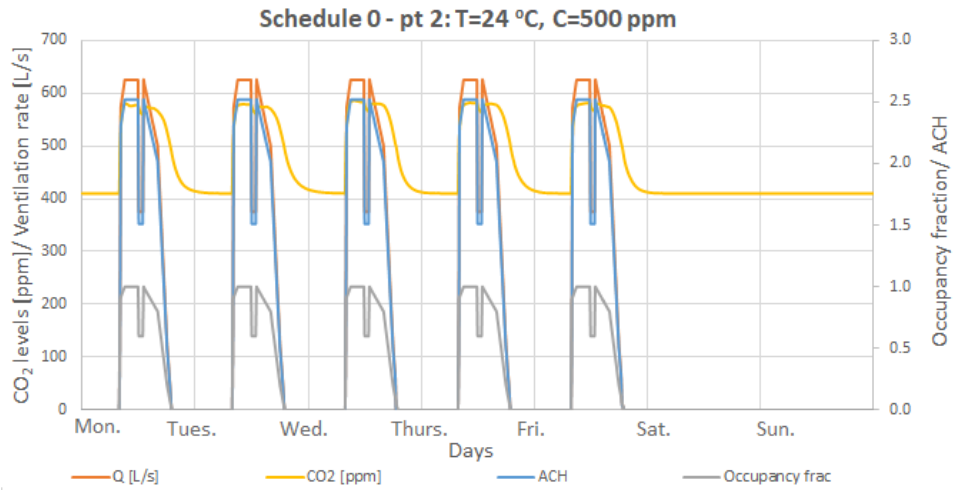

Figure 5.5: Schedule 0 (5 days/week, 100% full-time) CO_2 levels [ppm], ventilation rate [L/s], Occupancy fraction, and Air change rate for pt. 2 (T=24°C, C=500 ppm).

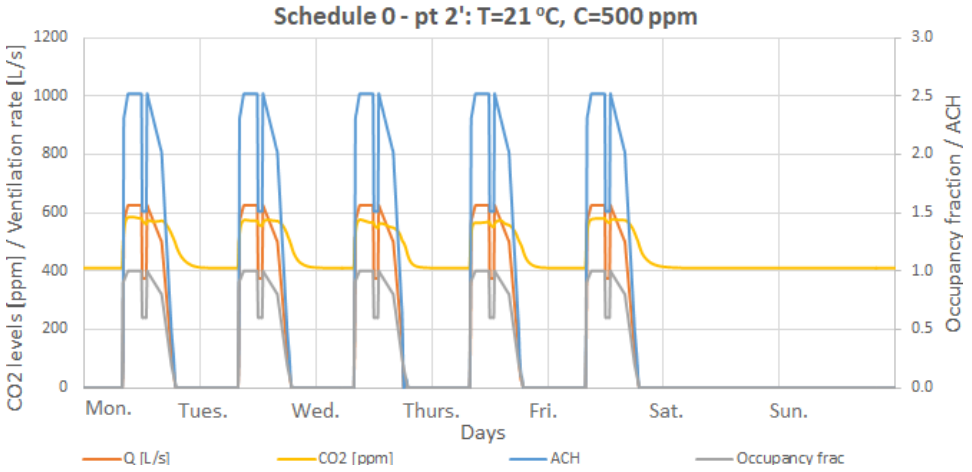

Figure 5.6: Schedule 0 (5 days/week, 100% full-time) CO_2 levels [ppm], ventilation rate [L/s], Occupancy fraction, and Air change rate for pt. 2' (T=21^{o}C, C=500 ppm).

contributed to reducing the heating energy demand.

Figures 5.4,5.5, and 5.6 display the values of daily CO_2 concentrations levels [ppm], ventilation rates [L/s], occupancy fraction, and values of Air changes per hour (ACH) for one week of January for the pre-COVID-19 schedule (schedule 0) for pt.1, p.t2, and p.t2' respectively.

It is noted from the figures that, as the occupancy fraction increases, the ventilation rates and air changes per hour increase as well, for the same CO_2 set point. Moreover, it is shown in Figures 5.4,5.5, and 5.6 that, as the indoor CO_2 sepoint is lowered (from point 1 to point 2 and from point 1 to point 2'), the values of ventilation rates and air change per hour are increased, for the same occupancy fraction.

The value of ventilation rate and air changes per hour is zero at the weekends and on the days when the occupancy fraction is set to zero.

5.4 Summary of Findings

The new data-driven predictive model of thermal comfort, presented in Chapter 4, was integrated into a building energy model (BEM) framework simulating an open-plan mechanically ventilated office in Vancouver.

The building control system was configured so that the heating set point and CO_2 levels are fixed. Results from 144 simulations, performed on four different occupancy schedules (for both pre-COVID-19 and post-COVID-19 hybrid work schedules) using different configurations of heating and indoor CO_2 concentration setpoints, showed that, with minimal building energy demand increases, it might be possible to increase the amount of fresh air while maintaining thermal comfort levels. This finding is valid for all the investigated post-COVID-19 occupancy profiles: the increase in heating energy demand, resulting from pumping higher amounts of fresh air and lowering the temperature setpoint, was always lower when the new thermal comfort model is adapted to keep the levels of occupant's thermal comfort maintained, as far as the building energy model of the CIRS building is concerned.

The simulation results suggested that by implementing the new model in building controls, the building energy use could possibly be reduced without sacrificing occupants' thermal comfort while increasing the amount of fresh air so that the spread of airborne diseases indoors is mitigated, an ongoing challenge faced by building operators, especially now after the COVID-19 pandemic.

In order for these findings to be further validated and generalized beyond the theoretical model used in this chapter, the new predictive model should be integrated and tested in a fully-simulated HVAC control system. More on the implica-

tions and limitations of the findings are discussed in the following section.

5.4.1 Limitations of Approach

The building energy model (BEM) developed in Chapter 4 and 5 aims to investigate and validate the significance of the new predictive thermal comfort model and to explore energy savings scenarios that might result from pumping higher amounts of fresh air while controlling both the heating and CO_2 setpoints. As discussed in Chapter 4, although the building energy model simulates a real office indoor space as a case study, the model has been simplified and is deemed theoretical.

For future use of the developed building energy model, more sophisticated simulations are required which can be more representative of fully-simulated HVAC control systems. Further, it is recommended to validate the model by integrating it into the buildings' HVAC system and including the full inner workings of fans, ducts, VAV boxes, heating elements,..etc.

Furthermore, it should be noted that the simulation results presented in this chapter are limited to building heating energy demand from the month of January. While January is the coldest month of the year in Canada and might be considered a suitable proxy for the rest of the heating months, it might not represent the heating season in the most accurate way. Future adaptations of this model should also include data for the cooling season (summer months), so the cooling energy demand could be evaluated and the annual building energy consumption could be fully modelled.

Vancouver's weather files were used to model the simulated open-plan office. While it might be a good representation of the climatic conditions of the IEQ data used to train the Bayesian model, more climatic conditions should be considered in

the future, so that universal conclusions can be drawn from the model. A detailed discussion of this thesis's conclusions, implications, and limitations is provided in the following chapter.

Chapter 6

Conclusions, Contributions, and Recommendations for Future Work

6.1 Overview

This chapter presents a summary of the conclusions of this PhD thesis, highlights the key research contributions, and discusses the limitations of the current study. The chapter also discusses the research implications and provides recommendations for future adaptation of this work.

6.2 Conclusions, Contributions, and Limitations

Seeking to bridge the gap between observations and predictions of thermal comfort, this thesis presented a novel predictive framework to improve the prediction accuracy of occupants' thermal satisfaction in office spaces. This work presented

a Bayesian framework that estimates the probability of an occupant feeling thermally satisfied as a function of, not only psychrometric IEQ parameters, but also to non-thermal metrics of IEQ. Using Bayesian statistical techniques, the proposed framework investigated and quantified the correlations between perceived thermal comfort and non-thermal metrics of IEQ. Posterior checks and model comparison approaches were performed to the Bayesian models to investigate potential improvements on prediction accuracy of occupants' thermal satisfaction and to test the statistical robustness and significance of the relationships between perceived thermal comfort and several metrics of IEQ. The study is the first to do so with respect to a large field study. Posterior results revealed that higher CO_2 concentrations are found to be independently correlated with lower incidences of occupant's thermal satisfaction in open-plan offices. Further, this is the first work to demonstrate that predictions of occupant's thermal comfort can be improved upon adding measurements of indoor CO_2 concentrations levels. While many recent studies have identified the multi-domain and mutli-contextual nature of thermal comfort - that thermal comfort may be related to other indices of IEQ, this work is one of few studies to evaluate these relationships quantitatively and in a manner that can support future thermal comfort prediction.

In the next subsections, the key research contributions are summarized and answers to the research questions are discussed.

6.2.1 On the Correlations Between Thermal Comfort and Non-thermal Metrics of IEQ

Posterior results drawn from the first case study, as well as model checks and validation performed on the Bayesian models trained on the COPE dataset showed

that perceived thermal satisfaction is correlated with speech intelligibility and measurements of CO_2 concentration levels. More specifically, the results of all model assessment methods performed on the COPE dataset suggested with high credibility that a reduction in speech intelligibility levels or in CO_2 concentration levels will increase the levels of thermal satisfaction, independent of all other parameters. A regression analysis of the relationship between speech intelligibility and perceived acoustic comfort was not undertaken. Moreover, correlations between perceived acoustic comfort and perceived thermal comfort are out of the scope of this research.

In the second case study, the UBC field IEQ dataset was developed and used to test the significance and robustness of the results inferred from the COPE dataset. Measurements of speech intelligibility levels were not collected during the second field study campaign. The posterior results drawn from the Bayesian logistic regression model trained on the COPE and UBC datasets showed that measurements of indoor CO_2 concentrations are correlated with occupants' perceived thermal satisfaction. Further, it was observed that including indoor CO_2 concentrations as an independent variable when predicting occupant thermal satisfaction credibly improved a model's fitness to observed data.

Independent of all other conditions, and only with regards to the workspaces surveyed in this work, only approximately 0.25 [0.17 - 0.38, 95% CrI] of surveyed occupants would be predicted to feel thermally satisfied at conditions of 23.5°C and 800 ppm of CO_2. This probability would appear to increase to 0.70 [0.60-0.77, 95% CrI] at conditions of 23.5°C and 500 ppm. Though it appears implausible that this is a universally-applicable finding, the fact that the prior observed relationship between CO_2 and predicted thermal satisfaction was strengthened with new field

data calls is notable. Investigation of the underlying relationships between measurements of indoor CO_2 concentrations and thermal comfort appears warranted.

In the literature review section of this thesis (1.2), several field studies were reviewed that provide some precedent of observed correlations between thermal comfort and non-thermal metrics of IEQ. These reviewed works point to potential second-order effects that would suggest a more intuitive relationship exists between non-thermal conditions and the perceived thermal comfort of building occupants. The argument is as follows, with the example of indoor noise levels: elevated noise levels may make an occupant frustrated with noise levels; frustration with noise levels may make an occupant more generally frustrated with other conditions of their workspace; therefore, the personal frustration caused by elevated noise levels may lead one to answering a question regarding perceived thermal comfort more negatively than they would in more pleasant acoustic settings. Although measured CO_2 concentrations levels appear to be independently correlated with thermal satisfaction, they may be only proxy or latent metrics of underlying phenomena that have a more intuitive relationship with an occupant's perception of the environment.

Nevertheless, the significance of the results of this research is sufficient to recommend more data collection on CO_2 concentrations in future thermal comfort field studies.

6.2.2 On the Root Causes and Significance of Observed Correlations Between CO_2 Concentrations and Thermal Comfort

There are two questions for this discussion that must be decoupled: what do the presented results show, and what do the presented results mean?

The results show that if one would wish to retroactively predict the response of occupants in the COPE and UBC field studies to the question of perceived thermal satisfaction, one might improve the predictive accuracy of a model by including CO_2 concentrations as a predictor. It is revealed that adding measurements of indoor CO_2 concentrations to predictive models of thermal satisfaction provide better predictive accuracy than models that would not include these parameters. It is notable that upon adding a new field study (UBC) to observations recorded 20 years earlier (COPE), the evidence in support of this statement increased. However, this is not the same as stating this work has identified a general causal relationship between CO_2 concentrations and thermal comfort.

It should be noted that observing a correlation between the *measurements* of two parameters is not the same as stating that one parameter is more likely than not dependent on the other or vice versa. This work therefore only finds some evidence to support a statement that *measurements* of carbon dioxide are *correlated* to reported thermal satisfaction. There is no evidence base to state that the physical concentration of carbon dioxide in indoor air is correlated directly with perceived thermal satisfaction.

Contextualizing the effect of CO_2 concentrations on perceived or actual occupant health and well-being is a contentious topic [89, 122]. It is therefore important to state that this thesis does not find evidence suggesting that ambient air CO_2 concentrations have a direct and/or general effect on one's body or mind such that one's perceived thermal satisfaction also changes. There are no study that would suggest the human brain and body is so sensitive to CO_2 conditions that a 300 ppm increase in indoor CO_2 levels would cause a $>50\%$ decrease in the likelihood one feels thermally satisfied, as is predicted by the generated model of $p(S \mid T, C)$.

It appears more reasonable to suggest that measured CO_2 in the COPE and UBC datasets might be a *latent* variable of some phenomena that can affect the reported thermal satisfaction of occupants. One hypothesis follows. It was observed in Figures 3.9 and 3.10 that responses to perceived thermal satisfaction and air quality satisfaction were correlated to each other. In Figure 3.11 it was also observed that CO_2 levels were inversely correlated, though weakly, with perceived air quality satisfaction in the COPE study. It may be more consistent with the guiding principles of indoor air quality that, at periods of relatively high indoor CO_2 levels, occupants are less likely to feel satisfied with air quality. They may subsequently feel thermally unsatisfied as a second-order effect, as has been discussed previously by Jamrozik et al. [60].

The subjective element of comfort consists of physiological, psychological and behavioural factors. As discussed earlier in Chapter 1, occupants' perception of thermal comfort is holistic and maybe affected by the psychological effect of many physical conditions that occupants encounter in the built environment. This thesis only finds some evidence to support a statement that measurements of CO_2 are correlated to reported thermal satisfaction in the field IEQ dataset. There is no evidence base to state that the physical concentration of CO_2 in indoor air is correlated directly with perceived thermal satisfaction. Therefore, a potentially actionable interpretation of the results could be to consider that measured indoor CO_2 levels may be a credible contextual variable of some underlying design or operating condition of the indoor environment, including air quality, which has some effect on how an occupant perceives the thermal environment.

For example, it has been found that occupant's thermal perception is affected by the number of people in a shared space [88, 108]. Then the argument is that in-

creased levels of indoor CO_2 can be caused by the increased number of occupants in the space, which in turn might affect the occupants perception of the thermal environment. The use of contextual variables for predicting thermal comfort quantitatively is not out of line with existing thermal comfort models. For example, the Adaptive model of thermal comfort relies on the mean running outdoor air temperature as a contextual variable, using it as a proxy for explaining the various adaptive measures taken by humans in naturally-ventilated buildings to improve their comfort. Its incorporation in the Adaptive model allows the model to be mathematically simple in explaining the complex physics of thermal comfort, yet still an accurate predictor of thermal acceptability in naturally-ventilated buildings.

Compared to prior studies that have quantified the relationship between thermal satisfaction and indoor CO_2 levels, discussed in Chapter 1, there is alignment between this study and literature. For example, Gauthier et al. observed an inverse relationship between indoor CO_2 levels and thermal sensation [44]. In this thesis, the inverse correlation is similarly observed for thermal satisfaction. More detailed quantitative comparisons to this prior work and others are challenging, however. In Gauthier et al., the authors found some statistical basis for the strength of their observed correlations ($p=0.08$, for the office experiment, and $p=0.48$ for the controlled chamber study), but they did not frame their analysis as a predictive model of thermal comfort. This work is the first known study to infer a predicted level of perceived thermal satisfaction as a condition of indoor CO_2 levels specifically.

6.2.3 On the Universality of the Findings

The observations drawn in this thesis stem from a regression analysis of a field dataset of over 900 occupant surveys. This is the first study to observe a quan-

tifiable relationship between metrics like indoor CO_2 concentrations and perceived thermal comfort at the scale of a large field study. However, this dataset is not large enough to be used to establish the universality of these observations.

Furthermore, the field dataset was not evaluated against a standard thermal comfort model, such as the PMV-PPD model, due to a lack of measured values for 'clo' and 'met'. Extending this analysis with a more comprehensive dataset, as well including other subjective metrics in the regression is an important future study. However, such analysis is not necessarily vital for assessing all of the implications of this work on future building design and controls. Vis-á-vis the PMV-PPD model, in mechanically-ventilated buildings there are limited avenues to measure 'clo' and 'met' during day-to-day operations. Recently, in Li et al. [80], the authors determined appropriate indices for continuous evaluation of thermal comfort compliance using typically-available building sensor data - though only sensors of psychometric conditions.

The PMV-PPD or Adaptive models of thermal comfort were derived from data of thousands of test subjects is telling, and more data and analysis is needed to draw any universal conclusion on the independent effect of non-thermal metrics of IEQ on perceived thermal comfort. However, the Bayesian approach to comparing and selecting models makes it possible to determine whether or not one is approaching universality of findings on the addition of new data. Hence, one interpretation of this study is not only that some evidence of a relationship between indoor CO_2 concentrations and thermal comfort is presented, but also that a method is grounded for assessing the incremental change in the evidence base upon the addition of future data.

6.2.4 On the Implementation of the New Predictive Model in Building Control Systems

In this thesis, a new data-driven predictive model of thermal comfort that predicts occupants' thermal satisfaction as a function of indoor air temperature and CO_2 concentrations was developed. The model was integrated into a building energy model framework to simulate a mechanically-ventilated open-plan office. The control system was set up so that the indoor air temperature and CO_2 levels are maintained fixed. Different configurations of indoor air temperature and indoor CO_2 levels setpoints were examined. The corresponding heating energy demand and occupants' thermal satisfaction were calculated for each examined scenario in order to investigate potential energy savings associated with pumping more fresh air while lowering the heating setpoint and while not compromising the occupant's thermal comfort. Four daily occupancy schedules were developed to reflect and compare different occupancy profiles for post-COVID-19 back-to-work hybrid working models.

The simulation results showed that, by using the new predictive thermal comfort model, it is possible to increase the ventilation rates with minimal building energy demand increase while not compromising occupants' thermal comfort. For all the studied post-COVID-19 occupancy profiles, the increase in heating energy, resulting from pumping higher amounts of fresh air, was always lower when the Bayesian predictive thermal comfort model was used so that the heating setpoint is lowered while maintaining fixed thermal comfort levels.

The buildings sector is facing several conflated challenges, particularly now in a post-COVID-19 world. Energy use should be minimized to support climate change objectives, but indoor air quality and well-being cannot be sacrificed - if

anything, it should be improved as well. This model presented a solution for building managers and operators who have been under pressure to increase the current amounts of fresh air to lower the risk of spreading the COVID-19 virus and other diseases indoors. The new thermal comfort model can then be used, to control for both heating set point and ventilation rate, so that the air change rate can be increased and thermal satisfaction levels are not compromised. This research has suggested that an open plan office with high amounts of fresh air can provide the same level of thermal comfort at lower temperatures than an office with 'typical' fresh air ventilation rates. Establishing these relationships in a manner in which building designers can account for these effects in building simulation is important.

6.3 Limitations, Implications and Recommendations for Future Work

There are several potential implications of this work on future research. Though the Bayesian analysis, conducted in this thesis, revealed that indoor CO_2 concentrations appear to be an effective predictor of occupants' thermal satisfaction, this was not the case with other factors such as ambient noise levels and illuminance levels. As discussed earlier in the thesis, some studies in the literature have found correlations between perceived thermal satisfaction and both noise and lighting levels. Yet, studies showing that thermal comfort predictions can be improved upon the addition of these variables in a statistically significant manner are still very limited.

It is recognized that this thesis did not investigate or evaluate the sensitivity of the presented results to the dependent error of measurement devices used in

the field study. Measurement error is an acknowledged issue in multi-perceptual studies within the built environment [78]. Though the observations inferred regarding correlations between thermal comfort, temperature, and CO_2 concentrations existed in ranges beyond reported measurement error, it is viable and recommended that measurement error should be considered in any future adaptation of this work. It would be possible to include measurement error within the context of the Bayesian model developed for this study.

It should be noted that, for some indoor environmental conditions, occupants' perceived satisfaction may be influenced by their long-term satisfaction; occupants may have opinions on thermal discomfort that are not only attributed to a 'right-here-right-now' condition but long-term trends with respect to a workspace's IEQ performance. In both the UBC and COPE datasets, and in an attempt to mitigate this limitation, physical measurements of IEQ were collected at each workstation over a period of 10 minutes. Additionally, occupants were continuously reminded to answer the questionnaire by considering only their 'right-here-right-now' perception of the built environment [96], [129], [30]. However, such limitation should be taken into consideration in any future adaptation of this work.

Overall, this thesis, and in particular the results illustrated in Figure 3.15, make it possible to speculate how measurements of CO_2 concentrations could potentially improve the accuracy of thermal comfort prediction models *if* the observations in this thesis continue to be found in extended studies. Yet, the core recommendation for future work is for expanded measurements of non-thermal IEQ conditions in future thermal comfort field studies. These should include, at least, measurements of indoor CO_2 levels. A greater field dataset is needed to determine whether the findings of this study are possibly universally applicable, and/or whether knowledge of

indoor CO_2 concentrations may improve personalized models of thermal comfort that are building- and context-specific. Further, collecting other objective indoor environmental metrics, for example, physiological parameters may contribute to improving the predictions of occupants' thermal comfort and the development of personalized thermal comfort models.

Recently, emerging studies have looked at the continuous evaluation of thermal comfort compliance using commercially-available building IEQ sensors [80]. In addition, portable and cost-effective off-the-shelf CO_2 sensors are becoming more common to see deployed in commercial and residential buildings [114, 131]. One possible implication of the current work could be that additional monitoring of metrics such as indoor CO_2 concentrations could further improve thermal comfort compliance estimates. Commercially-available RESET-accredited off-the-shelf IEQ sensors (see: https://www.reset.build) can be used for that purpose. A number of these accredited sensors are able to measure indoor CO_2 sensors within a similar range of error as the instrumentation sensor used in this field study.

Further, the IEQ data collected in this thesis were all collected in North America, within different climatic conditions, yet it is still hard to generalize the results beyond that scope. It is then recommended to collect data in a wider range of climatic conditions. This will contribute to more generalized and universal findings to be applied beyond the North-American scope.

Another key recommendation to be considered in the future adaption of this work is to include more seasonal variation while collecting measurements of indoor environmental conditions, when possible.

Further, including measurements of occupants' clothing insulation and metabolic rate in future IEQ data collected, whenever possible, is essential to evaluate the

findings against a standard thermal comfort model, such as the PMV-PPD model.

Regarding the statistical analysis of field IEQ data, Bayesian inference may continue to be an appropriate statistical framework to revisit this assessment in the future. As for the development of the IEQ survey questionnaire, and particularly in regard to the satisfaction questions, equidistance between individual responses in the satisfaction scale has been assumed in this work, yet this was not communicated to participants. Future works should consider this while designing IEQ survey questionnaires.

Further work could extend the current analysis by including other subjective IEQ metrics in the regression analysis. Drawing correlations between occupants' perceived satisfaction, subjective IEQ metrics, and other aspects of the indoor environment is a key to improving our understanding of the complex multi-perceptual nature of IEQ and comfort, especially in office spaces.

Further research is required to understand the physical reasons behind the observed correlations between high levels of CO_2 concentrations and increased thermal dissatisfaction. Evaluating whether indoor CO_2 is a proxy or latent variable of underlying phenomena that is affecting the occupants' perception of the thermal environment is a potential area of investigation. For instance, examining whether there is a relationship between higher occupancy density in office spaces and perceived thermal comfort, may contribute to understanding whether CO_2 was only a proxy of highly occupied office spaces or it was directly affecting perceived thermal satisfaction. The more correlations are drawn between different IEQ metrics, the more it will be possible (or not) to draw causal relationships between indoor CO_2 concentrations levels and perceived thermal comfort.

Bibliography

[1] Google sketchup. http://sketchup.com/. Accessed: 2022-09-22.

[2] Y. Al Horr, M. Arif, A. Kaushik, A. Mazroei, E. Elsarrag, and S. Mishra. Occupant productivity and indoor environment quality: A case of GSAS. *International Journal of Sustainable Built Environment*, 6(2):476–490, 2017. ISSN 22126104. doi:10.1016/j.ijsbe.2017.11.001.

[3] J. G. Allen, P. MacNaughton, U. Satish, S. Santanam, and J. D. Vallarino, J.and Spengler. Associations of cognitive function scores with carbon dioxide, ventilation, and volatile organic compound exposures in office workers: a controlled exposure study of green and conventional office environments. environmental health perspectives. *Environmental health perspectives*, 124(6):805–812, 2016.

[4] O. Alm, T. Witterseh, G. Clausen, J. Toftum, and P. O. Fanger. The impact of human perception of simultaneous exposure to thermal load, low-frequency ventilation noise and indoor air pollution. *In 8th International Conference on Indoor Air Quality and Climate*, pages 270–275, 1999.

[5] F. Alsaleem, M. K. Tesfay, M. Rafaie, K. Sinkar, D. Besarla, and P. Arunasalam. An iot framework for modeling and controlling thermal comfort in buildings. *Frontiers in Built Environment*, 6, 2020. doi:10.3389/fbuil.2020.00087.

[6] T. Ando and R. Tsay. Predictive likelihood for bayesian model selection and averaging. *International Journal of Forecasting*, 26(4):744–763, 2010.

[7] ASHRAE. Standard 55-2013 - Thermal Environmental Conditions for Human Occupancy. *Ashrae*, 8400:58, 2013. ISSN 1098-6596. doi:ISSN1041-2336.

[8] F. Auffenberg, S. Stein, and A. Rogers. A personalized thermal comfort model using a bayesian network. *In Proceedings of the 24th International Conference on Artificial Intelligence,*, page 25–31, 2015.

[9] I. Balazova, G. Clausen, and D. P. Wyon. The influence of exposure to multiple indoor environmental parameters on human perception, performance and motivation. *In Proceedings of CLIMA.*, 2007.

[10] W. A. Beckman, L. Broman, A. Fiksel, S. A. Klein, M. S. Eva Lindberg, and J. Thornton. Trnsys the most complete solar energy system modeling and simulation software. *Renewable energy*, 5:1–4, 1994.

[11] L. G. Berglund and W. S. Cain. Perceived air quality and the thermal environment. *In Proceedings of IAQ*, 89:93–99, 1989.

[12] J. R. Biden Jr. National strategy for the covid-19 response and pandemic preparedness. *Simon and Schuster*, 2021.

[13] M. Bujang, N. Sa'at, and J. L. Sidik, T. ample size guidelines for logistic regression from observational studies with large population: Emphasis on the accuracy between statistics and parameters based on real life clinical data. *Malays J Med Sci.*, 25(4):122–130, 2018.

[14] A. Buonomano, U. Montanaro, A. Palombo, and S. Santini. Dynamic building energy performance analysis: A new adaptive control strategy for stringent thermohygrometric indoor air requirements. *Applied Energy*, 163: 361–386, 2016. doi:10.1016/j.apenergy.2015.10.182.

[15] J. P. Campbell, V. A. Maxey, and W. A. Watson. Hawthorne effect: Implications for prehospital research. *Annals of Emergency Medicine*, 26 (5):590–594, 1995. doi:10.1016/s0196-0644(95)70009-9.

[16] S. J. Cao and H. Y. Deng. Investigation of temperature regulation effects on indoor thermal comfort, air quality, and energy savings toward green residential buildings. *Science and Technology for the Built Environment*, 25 (3):309–321, 2019.

[17] S. Carlucci, R. Bai, L.and De Dear, and L. Yang. Review of adaptive thermal comfort models in built environmental. *Building and Environment*, 2018.

[18] K. E. Charles, J. A. Veitch, K. M. Farley, and G. R. Newsham. Environmental satisfaction in open-plan environments: 3. further scale

validation. *National Research Council of Canada, Internal Report IRC-RR-152*, 2003.

[19] H. Chen, C. P., and S. Chen. How big is a big odds ratio? interpreting the magnitudes of odds ratios in epidemiological studies. *Communications in Statistics—Simulation and Computation*, 39:860–864, 2010.

[20] S. Cheung, T. C. abd Schiavon, E. T. Gall, M. Jin, and W. W. Nazaroff. Longitudinal assessment of thermal and perceived air quality acceptability in relation to temperature, humidity, and co2 exposure in singapore. *Building and Environment*, 115:80–90, 2017.

[21] T. Cheung, S. Schiavon, T. Parkinson, P. Li, and G. Brager. Analysis of the accuracy on pmv–ppd model using the ashrae global thermal comfort database ii. *Building and Environment*, 153:205–217, 2019.

[22] G. Chinazzo, L. Pastore, J. Wienold, and M. Andersen. A field study investigation on the influence of light level on subjective thermal perception in different seasons. *In Proceedings of 10th Windsor Conference: Rethinking Comfort.*, pages 12–15, 2018.

[23] G. Chinazzo, J. Wienold, and M. Andersen. Daylight affects human thermal perception. *Scientific reports 9.1*, pages 1–15, 2019.

[24] B. Coffey, F. Haghighat, E. Morofsky, and E. Kutrowski. A software framework for model predictive control with genopt. *Energy and Buildings*, 42(7):1084–1092, 2010.

[25] H. Cooper, L. V. Hedges, and J. C. Valentine. *The Handbook of Research Synthesis and Meta-Analysis*. Russell Sage Foundation, 2009.

[26] S. Crosby and A. Rysanek. Extending the Fanger PMV model to include the effect of non-thermal conditions on thermal comfort. *In Proceedings of eSIM 2020*, 2021.

[27] S. Crosby and A. Rysanek. Correlations between thermal satisfaction and non-thermal conditions of indoor environmental quality: Bayesian inference of a field study of offices. *Journal of Building Engineering*, 35: 102051, 2021.

[28] S. Crosby and A. Rysanek. A novel multi-domain model for thermal comfort which includes building indoor co2 concentrations. In *Proceedings of Building Simulation 2021: 17th Conference of IBPSA*, volume 17 of

Building Simulation, pages 2687–2694, Bruges, Belgium, September 2021. IBPSA. ISBN 978-1-7750520-2-9. doi:https://doi.org/10.26868/25222708.2021.30760. URL https://publications.ibpsa.org/conference/paper/?id=bs2021_30760.

[29] S. Crosby and A. Rysanek. On higher ventilation rates and energy efficiency in post-covid-19 buildings: A new thermal comfort model based on indoor co2 levels and temperatures. *In proceedings of Comfort at the Extremes (CATE) 2022, Edinburgh, UK*, 2022.

[30] S. Crosby and A. Rysanek. Predicting thermal satisfaction as a function of indoor co2 levels: Bayesian modelling of new field data. *Building and Environment*, 209:108569, 2022. doi:10.1016/j.buildenv.2021.108569.

[31] S. Crosby, G. Newsham, J. Veitch, and A. Rysanek. Bayesian inference of thermal comfort: evaluating the effect of "well-being" on perceived thermal comfort in open plan offices. *IOP Conference Series: Materials Science and Engineering*, 609(4):20–28, 2019.

[32] R. de Dear and G. Brager. Indoor Environmental Quality (IEQ) Title Developing an adaptive model of thermal comfort and preference. *ASHRAE Transactions*, 104:1–18, 1998. ISSN 00012505.

[33] R. de Dear and J. Richard. A Global Database of Thermal Comfort Field Experiments. *ASHRAE transactions*, (December):1141, 1998.

[34] R. de Dear, T. Akimoto, and E. Arens. Progress in thermal comfort research over the last twenty years. *International Journal of Indoor Environment and Health*, 23:442–461, 2013.

[35] N. Djongyang, R. Tchinda, and D. Njomo. Thermal comfort: A review paper. *Renewable and Sustainable Energy Reviews*, 14(9):2626–2640, 2010.

[36] A. I. Dounis, M. Bruant, G. Guarracino, P. Michel, and M. Santamouris. Indoor air-quality control by a fuzzy-reasoning machine in naturally ventilated buildings. *Applied Energy*, 54(1):11–28, 1996.

[37] D. Enescu. A review of thermal comfort models and indicators for indoor environments. *Renewable and Sustainable Energy Reviews*, 79:1353–1379, 2017. doi:10.1016/j.rser.2017.05.175.

[38] P. O. Fanger. *Thermal comfort, analysis and application in environmental engineering*. McGraw Hill, New York, 1972.

[39] Y. Feng, S. Liu, J. Wang, J. Yang, Y.-L. Jao, and N. Wang. Data-driven personal thermal comfort prediction: A literature review. *Renewable and Sustainable Energy Reviews*, 161:112357. doi:https://doi.org/10.1016/j.rser.2022.112357.

[40] P. M. Ferreira, A. E. Ruano, S. Silva, and E. Z. Conceição. Neural networks based predictive control for thermal comfort and energy savings in public buildings. *Energy and Buildings*, 55:238–251, 2012. ISSN 03787788. doi:10.1016/j.enbuild.2012.08.002.

[41] P. Frontczak, M.and Wargocki. Literature survey on how different factors influence human comfort in indoor environments. *Building and Environment*, 64(4):922–937, 2011.

[42] V. Földváry Ličina, T. Cheung, H. Zhang, T. de Dear, R.and Parkinson, E. . Arens, and S. Kaam. Ashrae global thermal comfort database ii. *Methods*, pages 06–24, 2021.

[43] G. Gao, J. Li, and Y. Wen. Deepcomfort: Energy-efficient thermal comfort control in buildings via reinforcement learning. *IEEE Internet of Things Journal*, 7(9):8472–8484, 2020.

[44] S. Gauthier, B. Liu, G. Huebner, and D. Shipworth. Investigating the effect of co2 concentration on reported thermal comfort. *in: Proceedings of CISBAT 2015 International Conference on Future Buildings and Districts*, 2015.

[45] A. Gelman. Multilevel (hierarchical) modeling: what it can and cannot do. *Technometrics*, 48(3):432–435, 2006.

[46] A. Gelman, A. Jakulin, M. G. Pittau, and Y. S. Su. A weakly informative default prior distribution for logistic and other regression models. *Annals of Applied Statistics*, 2(4):1360–1383, 2008. ISSN 19326157. doi:10.1214/08-AOAS191.

[47] A. Gelman, A. Vehtari, and J. Hwang. Understanding predictive information criteria for bayesian models. *Statistics and computing*, 24(6): 997–1016, 2014.

[48] G. Gelman and J. Hill. *Data Analysis using Regression and Multilevel/Hierarchical Models*. Cambridge University Press, 2007.

[49] G. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin. *Bayesian Data Analysis*. Chapman and Hall/CRC, 2020.

[50] Y. Geng, W. Ji, B. Lin, and Y. Zhu. The impact of thermal environment on occupant ieq perception and productivity. *Building and Environment*, 121: 158–167, 2017.

[51] A. Ghahramani, J. Pantelic, M. Vannucci, L. Pistore, S. Liu, B. Gilligan, S. Alyasin, E. Arens, K. Kampshire, and E. Sternberg. Personal co2 bubble: Context-dependent variations and wearable sensors usability. *Journal of Building Engineering*, 22:295–304, 2019.

[52] L. Held, B. Schrödle, and H. Rue. Posterior and cross-validatory predictive checks: a comparison of mcmc and inla. *Statistical modelling and regression structures*, pages 91–110, 2010.

[53] Y. Heo, R. Choudhary, and G. A. Augenbroe. Calibration of building energy models for retrofit analysis under uncertainty. *Energy and Buildings*, 47:550–560, 2012.

[54] L. Huang, Y. Zhu, Q. Ouyang, and B. Cao. A study on the effects of thermal, luminous, and acoustic environments on indoor environmental comfort in offices. *Building and Environment*, 49(1):304–309, 2012. ISSN 03601323. doi:10.1016/j.buildenv.2011.07.022.

[55] M. A. Humphreys and J. Fergus. The validity of iso-pmv for predicting comfort votes in every-day thermal environments. *Energy and Buildings*, 34(6):667–684, 2002.

[56] R.-L. Hwang, T.-P. Lin, and N.-J. Kuo. Field experiments on thermal comfort in campus classrooms in taiwan. *Energy and Buildings*, 38(1): 53–62, 2006. doi:10.1016/j.enbuild.2005.05.001.

[57] ILO. Working from home: Estimating the worldwide potential". *International Labour Organization -ILO policy brief*, 2020.

[58] D. Int-Hout. Comfort vs. energy use. *ASHRAE Journal*, 55(7):143, July 2013.

[59] N. Jacobs, B. Roberts, H. Reamer, C. Mathis, S. Gaffney, and R. Neitzel. Noise exposures in different community settings measured by traditional dosimeter and smartphone app. *Applied Acoustics*, 167:107408, 2020.

[60] A. Jamrozik, C. Ramos, J. Zhao, J. Bernau, N. Clements, T. Vetting Wolf, and B. Bauer. A novel methodology to realistically monitor office occupant reactions and environmental conditions using a living lab. *Building and*

Environment, 130(December 2017):190–199, 2018. ISSN 03601323. doi:10.1016/j.buildenv.2017.12.024.

[61] K. L. Jensen, J. Toftum, and P. Friis-Hansen. A Bayesian Network approach to the evaluation of building design and its consequences for employee performance and operational costs. *Building and Environment*, 44(3):456–462, 2009. ISSN 03601323. doi:10.1016/j.buildenv.2008.04.008.

[62] M. V. Jokl and K. Kabele. Optimal (comfortable) operative temperature estimation based on physiological responses of the human organism. *Acta Polytechnica*, 46(6):3–13, 2006.

[63] S. Kamaruzzaman, C. Egbu, E. Zawawi, A. Ali, and A. Che-Ani. *Energy and Buildings*, 43:407–413, 2011.

[64] P. Karavae and S. Ham. Identifying peer groups in a multifamily residential building for eco-feedback design. *5th International High Performance Buildings Conference, Purdue University*, 2018.

[65] J. Kim, S. Schiavon, and G. Brager. Personal comfort models–a new para-digm in thermal comfort for occupant-centric environmental control. *Building and Environment*, 132:114–124, 2018.

[66] S. A. Klein, W. A. Beckman, J. W. Mitchell, and J. A. Duffie. Trnsys-reference manual. *Solar Energy Laboratory, University of Wisconsin-Madison, Madison, WI*, 2000.

[67] S. A. Klein, W. A. Beckman, J. W. Mitchell, J. A. Duffie, N. A. Duffie, T. L. Freeman, J. C. Mitchell, J. E. Braun, B. L. Evans, and J. P. Kummer. Trnsys 16—volume 6 multizone building modeling with type 56 and trnbuild. *Solar Energy Laboratory, University of Wisconsin: Madison, WI, USA*, page 199, 2004.

[68] E. Kocaman, M. Kuru, and G. Calis. Do thermal comfort standards ensure occupant satisfaction? learning from occupants' thermal complaints. *In Creative Construction Conference 2019*, pages 682–687, 2019.

[69] D. Kolokotsa. Comparison of the performance of fuzzy controllers for the management of the indoor environment. *Building and Environment*, 38 (12):1439–1450, 2003.

[70] D. Kolokotsa, G. Saridakis, A. Pouliezos, and G. Stavrakakis. Design and installation of an advanced eib™ fuzzy indoor comfort controller using matlab™. *Energy and Buildings*, 38(9):1084–1092, 2006. doi:10.1016/j.enbuild.2005.12.007.

[71] M. Kong, B. Dong, R. Zhang, and Z. O'Neill. Hvac energy savings, thermal comfort and air quality for occupant-centric control through a side-by-side experimental study. *Applied Energy*, 306:117987, 2022.

[72] T. Kostiainen, I. Welling, M. Lahtinen, K. Salmi, E. Kähkönen, and J. Lampinen. Modeling of subjective responses to indoor air quality and thermal conditions in office buildings. *HvacR Research*, 14(6):905–923, 2008.

[73] M. H. Kristensen, A. Brun, and S. Petersen. Predicting danish residential heating energy use from publicly available building characteristics. *Energy and Buildings*, 173:28–37, 2018.

[74] Y. Kwak, J. H. Huh, and C. Jang. evelopment of a model predictive control framework through real-time building energy management system data. *Applied Energy*, 155:1–13, 2015.

[75] E. Laftchiev and D. Nikovski. An iot system to estimate personal thermal comfort. *2016 IEEE 3rd World Forum on Internet of Things (WF-IoT)*, 2016. doi:10.1109/wf-iot.2016.7845401.

[76] L. Lan, P. Wargocki, D. P. Wyon, and Z. Lian. Effects of thermal discomfort in an office on perceived air quality, SBS symptoms, physiological responses, and human performance. *Indoor air*, 21(5): 376–390, 2011.

[77] J. Langevin, J.and Wen and P. Gurian. Modeling thermal comfort holistically: Bayesian estimation of thermal sensation, acceptability, and preference distributions for office building occupants. *Building and Environment*, 69:206–226, 2013.

[78] S. Lechner, M. Schweiker, A. Wagner, and C. Moosmann. Review for "does thermal control improve visual satisfaction? interactions between occupants' self-perceived control, visual, thermal, and overall satisfaction". *Indoor Air*, 2021. doi:10.1111/ina.12851/v2/review1.

[79] S. Lee, I. Bilionis, P. Karava, and A. Tzempelikos. A bayesian approach for probabilistic classification and inference of occupant thermal preferences in office buildings. *Building and Environment*, 118:323–343, 2017.

[80] P. Li, T. Parkinson, S. Schiavon, T. M. Froese, de Dear, A. R., Rysanek, and S. Staub-French. Improved long-term thermal comfort indices for continuous monitoring. *Energy and Buildings*, 224:110270, 2020.

[81] W. Li, J. Zhang, and T. Zhao. Indoor thermal environment optimal control for thermal comfort and energy saving based on online monitoring of thermal sensation. *Energy and Buildings*, 197:57–67, 2019.

[82] Z. Lin and S. Deng. A study on the thermal comfort in sleeping environments in the subtropics—developing a thermal comfort model for sleeping environments. *Building and Environment*, 43:70–80, 2008.

[83] S. Liu, S. Schiavon, H. P. Das, M. Jin, and C. J. Spanos. Personal thermal comfort models with wearable sensors. *Building and Environment*, 162: 106281, 2019.

[84] S. Liu, S. Schiavon, H. P. Das, M. Jin, and C. J. Spanos. Personal thermal comfort models with wearable sensors. *Building and Environment*, 162: 106281, 2019. doi:10.1016/j.buildenv.2019.106281.

[85] N. Ma, D. Aviv, H. Guo, and W. W. Braham. Measuring the right factors: A review of variables and models for thermal comfort and indoor air quality. *Renewable and Sustainable Energy Reviews*, 135:110436, 2021. doi:https://doi.org/10.1016/j.rser.2020.110436.

[86] Z. H. Marion, J. A. Fordyce, and B. M. Fitzpatrick. A hierarchical bayesian model to incorporate uncertainty into methods for diversity partitioning. *Ecology*, 99(4):947–956, 2018.

[87] C. J. Marquardt, J. A. Veitch, and K. E. Charles. Environmental satisfaction with open-plan office furniture design and layout. *Institute for Research in Construction.Ottawa*, 2002.

[88] Marín-Restrepo, Laura, Trebilcock, Maureen, Porras-Salazar, and J. Ali. Adaptation by coexistence: Contrasting thermal comfort perception among individual and shared office spaces. *Architectural Science Review*, 2020.

[89] A. K. Mishra, S. Schiavon, P. Wargocki, K. W. Tham, and K. Lyngby. Carbon dioxide and its effect on occupant cognitive performance: A literature review. *Windosr 2020*, 2020.

[90] M. Moezzi and K. B. Janda. From "if only" to "social potential" in schemes to reduce building energy use.". *Energy Research Social Science*.

[91] K. Motoki, T. Saito, R. Nouchi, R. Kawashima, and M. Sugiura. Light colors and comfortable warmth: Crossmodal correspondences between thermal sensations and color lightness influence consumer behavior. *Food Quality and Preference*, 72:45–55, 2019. doi:10.1016/j.foodqual.2018.09.004.

[92] K. W. Mui, T. W. Tsang, and L. T. Wong. Bayesian updates for indoor thermal comfort models. *Journal of Building Engineering*, 29:101–117, 2020.

[93] W. NA, X. Zhang, P. Liu, X. Zhao, and Y. Cao. An empirical bayesian approach for calibrating long-term district-scale building energy consumption for heating. doi:http://dx.doi.org/10.2139/ssrn.4180073.

[94] K. Nagano and T. Horikoshi. ew comfort index during combined conditions of moderate low ambient temperature and traffic noise. *Energy and Buildings*, 37(3):287–294, 2005.

[95] G. R. Newsham, J. Veitch, J. G. Charles, K. E.and Clinton, J. G. Marquardt, J. S. Bradley, ..., and J. Readon. Environmental satisfaction in open plan environments: Relationships between physical variables. *Institute for Research in Construction, National Research Council Canada, Ottawa.*, 2004.

[96] G. R. Newsham, J. A. Veitch, and K. E. Charles. Risk factors for dissatisfaction with the indoor environment in open-plan offices: An analysis of COPE field study data. *Indoor Air*, 18(4):271–282, 2008. ISSN 09056947.

[97] M. O. Ng, M. Qu, P. Zheng, Z. Li, and Y. Hang. CO_2-based demand controlled ventilation under new ashrae standard 62.1-2010: a case study for a gymnasium of an elementary school at west lafayette, indiana. *Energy and Buildings*, 43(11):3216–3325, 2011.

[98] J. E. Oakley and A. O'Hagan. Probabilistic sensitivity analysis of complex models: a bayesian approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 66(3):751–769, 2004.

[99] N. A. Oseland and M. A. Humphreys. Thermal comfort: Past, present and future. *Proceedings of a Conference Held at the Building Research Establishment*, 1994.

[100] D. Oswald, F. Sherratt, and S. Smith. Handling the hawthorne effect: The challenges surrounding a participant observer. *Review of Social Studies*, 1 (1):53–74, 2014. doi:10.21586/ross0000004.

[101] N. Pellerin and V. Candas. Combined effects of temperature and noise on human discomfort. *Physiology and Behavior*, 78(1):99–106, 2003.

[102] N. Pellerin and V. Candas. Effects of steady-state noise and temperature conditions on environmental perception and acceptability. *Indoor air*, 14 (2):129–136, 2004.

[103] T. M. Peters, D. Rabidoux, C. O. Stanier, and T. R. Anthony. Assessment of university classroom ventilation during the covid-19 pandemic. *Journal of Occupational and Environmental Hygiene*, 19(5):295–301, 2022. doi:https://doi.org/10.1080/15459624.2022.2053142.

[104] J. Piironen and A. Vehtari. Comparison of bayesian predictive methods for model selection. *Statistics and Computing*, 27(3):711–735, 2017.

[105] C. Roda, P. M. Bluyssen, C. Mandin, S. Fossati, P. Carrer, Y. de Kluizenaar, V. Mihucz, E. de Oliveira Fernandes, and J. Bartzis. Psychosocial work environment and building related symptoms. *Healthy Buildings Europe 2015, 18-20 May 2015, Eindhoven, The Netherlands*, page 1, 2015.

[106] J. N. Rouder and J. Lu. An introduction to Bayesian hierarchical models with an application in the theory of signal detection. *Psychonomic bulletin review*, 12(4):573–604, 2005.

[107] N. Rupp, R.and Vásquez and R. Lamberts. *Energy and Buildings*, 105: 178–205, 2015.

[108] M. Schweiker and A. Wagner. The effect of occupancy on perceived control, neutral temperature, and behavioral patterns. *Energy and Buildings*, 117:246–259, 2016. doi:10.1016/j.enbuild.2015.10.051.

[109] M. Schweiker, G. Huebner, B. Kingma, R. Kramer, and H. Pallubinsky. Drivers of diversity in human thermal perception–a review for holistic comfort models. *Temperature*, 5(4):308–342, 2018.

[110] M. Schweiker, E. Ampatzi, M. S. Andargie, R. K. Andersen, E. Azar, V. M. Barthelmes, C. Berger, ..., and L. P. Edappilly. Review of multi-domain approaches to indoor environmental perception and behaviour. *Building and Environment*, 2020.

[111] M. Simoni, I. Annesi-Maesano, T. Sigsgaard, D. Norback, G. Wieslander, W. Nystad, M. Canciani, P. Sestini, and G. Viegi. School air quality related to dry cough, rhinitis and nasal patency in children. *European Respiratory Journal*, 35(4):742–749, 2010. doi:10.1183/09031936.00016309.

[112] M. Sourbron and L. Helsen. Evaluation of adaptive thermal comfort models in moderate climates and their impact on energy use in office buildings. *Energy and Buildings*, 43(2-3):423–432, 2014.

[113] Z. Sun, S. Wang, and Z. Ma. In-situ implementation and validation of a co2-based adaptive demand-controlled ventilation strategy in a multi-zone office building. *Building and Environment*, 46(1):124–133, 2011.

[114] V. Tanasiev, G. C. Pătru, D. Rosner, N. H. Sava, G., and A. Badea. Enhancing environmental and energy monitoring of residential buildings through iot. *Automation in Construction*, 126:103662, 2021.

[115] S. K. Tang and C. T. Wong. Performance of noise indices in office environment dominated by noise from human speech. *Applied Acoustics*, 55(4):293–305, 1998.

[116] F. Tartarini, S. Schiavon, T. Cheung, and T. Hoyt. CBE thermal comfort tool: online tool for thermal comfort calculations and visualizations. *SoftwareX 12*, 100563, 2020. doi:https://doi.org/10.1016/j.softx.2020.100563.

[117] S. Tasaka. Bayesian hierarchical regression models for QoE estimation and prediction in audiovisual communications. *IEEE Transactions on Multimedia*, 19(6):1195–1208, 2017.

[118] E. Teitelbaum, K. Chen, F. Meggers, and et al. Globe thermometer free convection error potentials. *Sci Rep*, (10):2652, 2020.

[119] The International WELL Building Institute. *The WELL Building Standard*, 2014.

[120] M. E. Tipping. Bayesian inference: An introduction to principles and practice in machine learning. *Summer School on Machine Learning*, Springer, Berlin, Heidelberg.:41–62, 2003.

[121] J. Toftum, R. V. Andersen, and K. L. Jensen. Occupant performance and building energy consumption with different philosophies of determining acceptable thermal conditions. *Building and Environment*, 44(10), 2009.

[122] Z. Tu and et al. Human responses to high levels of carbon dioxide and air temperature. *Indoor air*, 31:872–886, 2021.

[123] J. Van Hoof. Forty years of fanger's model of thermal comfort: comfort for all? *Indoor air*, 18(3):182–201, 2008.

[124] J. Van Hoof, M. Mazej, and J. L. Hensen. Thermal comfort: research and practice. *Frontiers in Bioscience*, 15(2):765–788, 2010.

[125] A. Vehtari and J. Ojanen. A survey of bayesian predictive methods for model assessment, selection and comparison. *Statistics Surveys*, 6: 142–228, 2012.

[126] A. Vehtari, A. Gelman, and J. Gabry. Efficient implementation of leave-one-out cross-validation and waic for evaluating fitted bayesian models. *arXiv preprint arXiv:1507.04544.*, 2015.

[127] A. Vehtari, A. Gelman, and J. Gabry. Practical bayesian model evaluation using leave-one-out cross-validation and waic. *Statistics and computing*, 27 (5):1413–1432, 2017.

[128] J. A. Veitch, K. M. Farley, and G. R. Newsham. Environmental satisfaction in open plan environments: 1. scale validation and methods. *Institute for Research in Construction, National Research Council Canada, Ottawa.*, 2002.

[129] J. A. Veitch, K. E. Charles, K. M. Farley, and G. R. Newsham. A model of satisfaction with open-plan office conditions: Cope field findings. *Journal of Environmental Psychology*, 27(3):177–189, 2007.

[130] J. Vischer. Towards an environmental psychology of workspace: How people are affected by environments for work. *Architecture Science Review*, 51:97–108.

[131] A. Vishwanath, Y. H. Hong, and C. Blake. Experimental evaluation of a data driven cooling optimization framework for hvac control in commercial buildings. *In Proceedings of the Tenth ACM International Conference on Future Energy Systems*, pages 78–88, 2019.

[132] E. Wagner, A.and Gossauer, C. Moosmann, T. Gropp, and R. Leonhart. Thermal comfort and workplace occupant satisfaction—results of field studies in german low energy office buildings. *Energy and Buildings*, 39: 758–769, 2007.

[133] F. Wang, Z. Chen, Y. Jiang, Q. C. Zhao, and Y. Zhao. Preliminary study on perception-based indoor thermal environment control. *In the 13th International Conference on Indoor Air Quality and Climate.*, 2014.

[134] S. Watanabe. Equations of states in singular statistical estimation. *Neural Networks*, 23(1):20–34, 2010.

[135] S. R. West, J. K. Ward, and J. Wall. Trial results from a model predictive control and optimisation system for commercial building hvac. *Energy and Buildings*, 27:271–279, 2014.

[136] J. S. Witte, S. Greenland, R. W. Haile, and C. L. Bird. Hierarchical regression analysis applied to a study of multiple dietary exposures and breast cancer. *Epidemiology*, pages 612–621, 1994.

[137] L. T. Wong, K. W. Mui, and C. T. Cheung. Bayesian thermal comfort model. *Building and environment*, 82:171–179, 2014.

[138] World Health Organization. Roadmap to improve and ensure good indoor ventilation in the context of COVID-19. *https://www.who.int/publications-detail-redirect/9789240021280*, 2021.

[139] W. Yang and H. J. Moon. Cross-modal effects of noise and thermal conditions on indoor environmental perception and speech recognition. *Applied Acoustics*, 141:1–8, 2018.

[140] B. T. Yau, Y. H.and Chew. A review on predicted mean vote and adaptive thermal comfort models. *Building Services Engineering Research and Technology*, 35(1):23–35, 2014.

[141] X. Zhang, P. Wargocki, Z. Lian, and C. Thyregod. Effects of exposure to carbon dioxide and bioeffluents on perceived air quality, self-assessed acute health symptoms, and cognitive performance. *Indoor Air*, 27(1):47–64, 2017.

[142] Q. Zhao, Z. Lian, and D. Lai. Thermal comfort models and their developments: A review. *Energy and Built Environment*, 2(1):21–33, 2021. ISSN 2666-1233. doi:https://doi.org/10.1016/j.enbenv.2020.05.007.

Appendix A

UBC Field IEQ Database

Date-Time	WSID	building#	Temp	RH	CO2	CO	VOC	CLO	MRT	V
7/15/2019 10:39	1	10	24.4	57.2	582	0.9	46	0.41	25.21	0.03
7/15/2019 10:59	2	10	23.5	59.9	606	0.8	40	0.59	23.45	0.07
7/17/2019 10:03	3	10	21.6	64.2	641	0.9	73	0.72	21.31	0.01
7/17/2019 10:14	4	10	21.6	64.8	646	0.9	68	0.51	21.64	0.01
7/17/2019 10:28	5	10	21.3	66.5	672	0.9	62	0.59	21.7	0.01
7/17/2019 11:08	6	10	22.7	60.5	584	0.9	55	0.54	23.18	0.01
7/17/2019 11:17	7	10	23	59.1	579	0.9	51	0.51	22.52	0.01
7/17/2019 11:45	8	10	22.2	61.7	556	0.8	46	0.9	22.39	0.03
7/17/2019 11:54	9	10	21.9	62	510	0.8	46	0.54	21.47	0
7/17/2019 15:54	10	10	23.8	58.8	560	0.8	43	0.81	23.79	0.02
7/17/2019 16:02	11	10	24.1	57.8	621	0.8	46	0.39	23.96	0.05
7/19/2019 10:19	12	10	22	51.8	589	1	92	0.78	21.9	0.01
7/19/2019 10:09	13	10	21.2	54.1	583	1	96	0.87	21.89	0.02
7/19/2019 10:38	14	10	22.7	49.3	501	0.9	82	0.57	23.76	0.01
7/19/2019 11:02	15	10	22.7	51.1	597	0.8	82	0.81	22.45	0.02
7/19/2019 11:22	16	10	22.8	49.5	561	0.7	80	0.59	21.49	0.03
7/19/2019 11:49	17	10	22.9	50.3	592	0.9	84	0.51	23.75	0.01
7/19/2019 12:01	18	10	19.4	58.1	605	0.9	85	0.51	22.78	0.31
7/19/2019 12:30	19	10	23.2	50	486	0.6	78	0.45	19.87	0
7/19/2019 12:56	20	10	22.8	50	556	0.9	80	0.69	24.09	0.02
7/19/2019 13:18	21	10	23	50.8	561	0.8	80	0.53	22.76	0.06
7/19/2019 13:52	22	10	23.3	49.9	618	0.8	77	0.74	22.94	0.01
7/19/2019 14:10	23	10	22.6	49.8	511	0.7	71	0.64	23.09	0.05
7/19/2019 14:29	24	10	24.8	45.3	601	0.8	75	0.42	26.06	0.04

Table A.1: UBC field dataset- Part 1-I

Date-Time	WSID	building#	Temp	RH	CO2	CO	VOC	CLO	MRT	V
7/19/2019 14:49	25	10	24	47.1	596	0.8	75	0.39	24.24	0.01
7/17/2019 10:22	26	10	21.3	65.6	627	0.9	63	0.64	21.29	0.01
7/17/2019 10:41	27	10	21.7	63.3	574	0.8	57	0.53	21.79	0
7/15/2019 12:09	28	11	25.1	55.8	584	0.9	40	0.59	25.75	0.04
7/15/2019 13:59	29	11	26	52.9	643	0.9	44	0.57	24.35	0.08
7/16/2019 11:38	30	11	24.9	64.7	504.3	0.915	42.4	0.59	25.41	0.09
7/16/2019 12:11	31	11	24.8	57.1	566.34	0.89	40	0.51	24.9	0.03
7/16/2019 12:22	32	11	25.5	56.3	549.58	0.84	37.58	0.57	25.39	0.06
7/16/2019 12:33	33	11	25.3	56.3	534.41	0.841	40	0.43	25.28	0.15
7/16/2019 12:54	34	11	24.8	57.4	578.64	0.772	40	0.51	24.92	0.13
7/16/2019 13:24	35	11	24	58.8	621.47	0.861	52.67	0.37	24.16	0.02
7/16/2019 14:08	36	11	25.6	58	530.24	0.78	40	0.59	25.51	0.03
7/18/2019 10:40	37	11	20.6	50	555	0.9	37	0.6	22.26	0.04
7/18/2019 11:00	38	11	23	49.2	643	0.9	61	0.42	23.01	0.01
7/18/2019 11:18	39	11	23.7	48.2	571	0.8	40	0.6	21.96	0.03
7/18/2019 14:01	40	11	25.1	58.5	592	0.8	40	0.53	24.63	0.05
7/18/2019 14:26	41	11	24.5	46.2	623	0.9	35	0.54	23.95	0.02
7/18/2019 15:01	42	11	24.3	44.5	518	0.9	28	0.44	24.79	0.1
7/18/2019 16:04	43	11	24.4	43.8	569	0.8	29	0.79	25.02	0.04
7/16/2019 13:52	44	11	26.9	60	511.27	0.73	40	0.51	25.71	0.08
7/15/2019 15:28	45	12	24.7	55.8	592	0.7	39	0.53	24.58	0.01
7/15/2019 15:41	46	12	26.1	51.7	630	0.7	40	0.47	25.06	0.04
7/15/2019 12:22	47	12	25.4	52.6	612	1	40	0.87	26.16	0.03
7/16/2019 9:51	48	12	24.1	55.2	537.07	1.046	49.93	0.57	23.93	0.04
7/16/2019 10:03	49	12	23.3	59.4	529.7	0.85	47.725	0.57	23.92	0.08
7/16/2019 10:10	50	12	23.8	58.2	555.45	0.8	46.25	0.59	23.58	0.07
7/16/2019 10:31	51	12	22.5	62.8	558.13	0.66	46	0.53	22.86	0.16
7/15/2019 15:18	52	12	25	54	625	0.8	43	0.31	25.4	0.04
7/15/2019 14:51	53	13	25.1	55.5	518	0.8	40	0.53	27.31	0.18
8/13/2019 9:46	54	14	23.1	56.5	666	0.8	85	0.33	23.46	0.01
8/13/2019 9:54	55	14	23.7	53.3	566	0.8	80	0.57	22.89	0.04
8/13/2019 10:09	56	14	24	53	616	0.8	79	0.57	22.82	0.04
8/13/2019 10:22	57	14	24	54.1	634	0.8	80	0.43	22.95	0.02
8/13/2019 10:37	58	14	23.7	55	756	0.8	80	0.47	22.26	0.02
8/13/2019 10:53	59	14	23.9	54.9	758	0.8	85	0.39	23.24	0.01
8/13/2019 11:14	60	14	24.1	54.8	755	0.8	83	0.57	22.88	0.04
8/13/2019 11:20	61	14	23.8	55.6	797	0.8	85	0.6	22.54	0.02
8/13/2019 11:24	62	14	23.5	55.9	765	0.8	80	0.67	21.96	0.03
8/13/2019 11:30	63	14	22	59.8	731	0.7	80	0.75	21.95	0.1
8/13/2019 11:34	64	14	22.3	59	750	0.7	80	0.51	22.3	0.02
8/13/2019 11:43	65	14	23.1	57	732	0.7	80	0.43	22.27	0.08

Table A.2: UBC field dataset- Part 1-II

Date-Time	WSID	building#	Temp	RH	CO2	CO	VOC	CLO	MRT	V
8/13/2019 11:52	66	14	23.2	57.3	744	0.8	84	0.59	22.22	0.03
8/13/2019 12:03	67	14	23.3	57.3	660	0.7	83	0.51	23.12	0.01
8/15/2019 10:47	68	14	25.2	53.8	569	1	106	0.53	25.54	0.19
8/15/2019 10:28	69	14	24.8	55	552	1.1	113	0.51	25.71	0.06
8/15/2019 11:04	70	14	23.6	57.8	702	0.8	102	0.6	22.63	0.02
8/15/2019 11:15	71	14	23.5	59.5	667	0.9	102	0.55	23.46	0.06
8/15/2019 11:26	72	14	23.8	59.1	619	0.9	99	0.3	23.87	0.06
8/15/2019 11:35	73	14	23.3	60.2	632	0.9	97	0.74	23.33	0.01
8/15/2019 11:39	74	14	23.3	59.7	624	0.9	96	0.74	23.17	0.01
8/15/2019 11:54	75	14	23.4	59.4	602	0.8	96	0.56	22.95	0.01
8/15/2019 11:58	76	14	23.8	58.1	635	0.9	93	0.59	22.88	0.02
8/15/2019 13:14	77	14	23.8	58.4	677	0.8	95	0.87	23.73	0.03
8/15/2019 13:30	78	14	23.9	58.3	672	0.9	91	0.59	23.62	0.13
8/15/2019 13:50	79	14	24.6	56.2	585	0.8	91	0.77	23.11	0.06
8/15/2019 13:58	80	14	24.2	58	655	0.9	93	0.81	23.16	0.07
8/15/2019 14:05	81	14	24.2	56.8	663	0.8	93	0.44	22.24	0.03
8/15/2019 14:12	82	14	23.9	59	632	0.8	91	0.53	22.69	0.02
8/15/2019 14:18	83	14	23	62	632	0.8	92	0.53	22.71	0.05
8/15/2019 14:26	84	14	23.7	57.9	653	0.8	91	0.37	22.25	0.02
8/15/2019 15:01	85	14	23	60.7	611	0.8	91	0.51	22.96	0.01
8/15/2019 14:49	86	14	22.7	61.6	615	0.9	91	0.44	23.03	0.02
8/15/2019 14:54	87	14	23.1	60.8	580	0.9	91	0.59	23.28	0.02
8/15/2019 15:07	88	14	23.4	59.8	627	0.8	91	0.79	23.43	0.01
8/15/2019 15:19	89	14	23	59.7	661	0.8	91	0.45	22.89	0.02
8/15/2019 15:26	90	14	23	58.8	715	0.8	90	0.44	23.3	0.02
8/15/2019 15:43	91	14	22.9	58.9	728	0.8	87	0.81	22.8	0.05
8/15/2019 15:47	92	14	22.6	58.6	566	0.9	85	0.51	22.84	0.12
8/15/2019 15:58	93	14	20.8	64.6	589	0.7	87	0.51	22.72	0.03
8/15/2019 16:04	94	14	21.9	61	582	0.7	85	0.49	23.15	0.06
1/27/2020 10:35	95	10	21.5	43.6	941	0.9	138	0.5	21.18632358	0.01
1/27/2020 10:58	96	10	23	42	1026	0.9	134	0.63	22.4249528	0.01
1/27/2020 11:43	97	10	23.9	40.8	1258	0.8	140	0.55	22.50225339	0.08
1/27/2020 12:15	98	10	23.5	41	1271	0.7	139	0.57	22.44094256	0.03
1/27/2020 12:31	99	10	23.1	38.5	1084	0.7	134	0.66	21.45489183	0.05
1/27/2020 12:36	100	10	23.5	38.3	961	0.6	129	0.8	22.27178049	0

Table A.3: UBC field dataset- Part 1-III

Date-Time	WSID	building#	Temp	RH	CO2	CO	VOC	CLO	MRT	V
1/27/2020 12:53	101	10	22.8	37	821	0.6	124	0.5	20.99535004	0.02
1/27/2020 14:47	102	10	22.5	41.2	908	0.7	139	0.5	21.714506	0.04
1/27/2020 15:07	103	10	22.9	41.8	1117	0.8	146	0.6	21.41283334	0.04
1/27/2020 15:29	104	10	21.4	41.5	734	0.7	134	0.5	20.47944026	0.02
1/27/2020 15:41	105	10	21.9	41	744	0.7	134	0.75	20.63378804	0.04
1/27/2020 15:53	106	10	21.9	41.1	756	0.7	134	0.6	20.66599076	0.02
1/29/2020 13:48	107	10	23.2	47	1211	0.9	157	0.6	21.36189604	0.01
1/29/2020 14:19	108	10	24.1	42.8	1356	1	164	0.61	22.55760708	0.12
1/29/2020 14:36	109	10	24.6	41.7	891	0.9	148	0.66	23.71156934	0.01
1/29/2020 14:56	110	10	23.3	41.8	822	0.8	142	0.63	22.10147105	0.03
1/31/2020 10:44	111	10	22.3	43	791	1.3	158	0.87	20.92110853	0.02
1/31/2020 11:05	112	10	22.2	43.1	1032	1.2	158	0.63	20.83145714	0.03
1/31/2020 11:22	113	10	22.6	42.9	971	1.3	158	0.52	21.53172489	0.08
1/31/2020 11:40	114	10	22.3	42.5	830	1.1	154	0.48	20.96506393	0.07
1/31/2020 11:36	115	10	21.9	42.7	815	1.1	154	0.48	20.62964493	0.05
1/31/2020 13:32	116	10	22.3	44.6	1260	1.2	166	0.48	21.3470241	0.01
1/31/2020 13:38	117	10	22.7	44.3	1342	1.2	170	0.55	21.69447343	0.09
1/31/2020 14:07	118	10	23.1	44.2	1114	1.2	164	0.57	22.05991033	0.08
1/31/2020 14:43	119	10	23.3	44	966	1.1	152	0.49	22.28645231	0.05
1/31/2020 14:43	120	10	22.3	47.9	966	1.1	152	0.63	21.344	0.02
2/3/2020 10:50	121	11	22.1	27.9	694	0.8	123	0.82	21.6715	0.04
2/3/2020 11:10	122	11	22.7	26.3	725	0.7	116	0.9	22.3812	0.07
2/3/2020 11:14	123	11	23.1	25.6	666	0.7	116	0.75	22.4427	0.02
2/3/2020 11:23	124	11	23	25.7	712	0.7	115	0.55	22.786	0.02
2/7/2020 10:50	125	11	20.6	37	642	1.1	146	0.82	19.4342	0.07
2/7/2020 11:01	126	11	21.1	36.9	663	1	146	0.35	21.7848	0.03
2/7/2020 11:13	127	11	22.7	36.5	677	1.1	146	0.57	23.2356	0.01
2/7/2020 11:31	128	11	22.1	36.3	772	1	165	0.49	22.7173	0.05
2/7/2020 11:34	129	11	22.8	36.4	770	1	164	0.44	23.0152	0.02
2/7/2020 11:43	130	11	23.3	36.1	837	1.1	158	0.81	23.011	0.08
2/10/2020 12:10	131	12	20.5	39.8	619	1	146	0.63	21.1477	0.09
2/10/2020 12:24	132	12	22	35.2	704	0.9	146	0.4	19.6658	0.03
2/10/2020 12:35	133	12	24.2	34.6	707	1	147	0.7	24.4567	0.01
2/10/2020 13:24	134	12	18.7	36.2	660	0.6	147	0.58	19.7867	0.04
2/11/2020 10:19	135	12	20.7	36.2	605	0.9	140	0.5	20.9008	0.04
2/11/2020 10:35	136	12	21.3	37.9	646	1	140	0.71	20.7856	0.02
2/11/2020 11:17	137	12	20.9	34.7	725	1.1	146	0.55	19.8402	0.01
2/12/2020 14:01	138	12	22.1	36.6	706	0.6	146	0.63	23.1858	0.03
2/12/2020 14:20	139	12	22.1	35.7	678	0.5	152	0.63	22.0742	0.02
2/12/2020 14:26	140	12	22.7	35.3	652	0.6	150	0.52	23.120072	0.02
2/12/2020 14:41	141	12	23.6	31.2	685	0.6	147	0.58	23.239459	0.03

Table A.4: UBC field dataset- Part 1-IV

Date-Time	WSID	building	Noise	PART	HL Lux	HR Lux	V Lux	Lux	Sat Temp	RAIN
7/15/2019 10:39	1	10	41.7	0	1400	1900	603	1650	6	No
7/15/2019 10:59	2	10	46.8	0	749	951	345	850	7	No
7/17/2019 10:03	3	10	54.5	1.26	59.1	71	98.5	65.05	4	Yes
7/17/2019 10:14	4	10	54.2	1.26	324	529	160.9	426.5	5	Yes
7/17/2019 10:28	5	10	56.3	1.26	266	159.9	45.3	212.95	4	Yes
7/17/2019 11:08	6	10	42.3	1.36	236	240	218	238	4	Yes
7/17/2019 11:17	7	10	52.9	1.26	232	760	144	496	6	Yes
7/17/2019 11:45	8	10	46.1	1.47	788	580	451	684	7	Yes
7/17/2019 11:54	9	10	44.2	0	256	277	162	266.5	6	Yes
7/17/2019 15:54	10	10	47.5	0	676	435	236	555.5	6	No
7/17/2019 16:02	11	10	47.6	1.36	255	245	507	250	5	No
7/19/2019 10:19	12	10	46.5	1.26	233	249	990	241	6	No
7/19/2019 10:09	13	10	49.8	1.26	226	163	438	194.5	4	No
7/19/2019 10:38	14	10	45.6	0	1050	950	1775	1000	6	No
7/19/2019 11:02	15	10	50.7	0	363	210	253	286.5	4	No
7/19/2019 11:22	16	10	44.6	0	330	353	170.5	341.5	4	No
7/19/2019 11:49	17	10	44.6	1.25	650	436	176.2	543	6	No
7/19/2019 12:01	18	10	52.6	0	1160	630	379	895	6	No
7/19/2019 12:30	19	10	44.9	1.25	177.5	211	410	194.25	7	No
7/19/2019 12:56	20	10	43.2	1.08	315	193	316	254	3	No
7/19/2019 13:18	21	10	43.3	1.26	637	570	323	603.5	5	No
7/19/2019 13:52	22	10	44.4	0	357	325	285	341	5	No
7/19/2019 14:10	23	10	51.2	1.25	1050	830	835	940	6	No
7/19/2019 14:29	24	10	48.1	0	667	930	818	798.5	5	No
7/19/2019 14:49	25	10	46.2	0	880	610	330	745	1	No
7/17/2019 10:22	26	10	55.1	1.26	261	178	141.1	261	6	Yes
7/17/2019 10:41	27	10	53.6	0	158.3	105.2	51.2	158.3	6	Yes
7/15/2019 12:09	28	11	44.6	1.7	441	595	478	518	4	No
7/15/2019 13:59	29	11	47.2	1.04	512	430	286	471	2	No
7/16/2019 11:38	30	11	55.8	1.8	460	400	203	430	4	No
7/16/2019 12:11	31	11	59.6	0	628	422	272	525	6	No
7/16/2019 12:22	32	11	46.2	1.8	450	468	278	459	3	No
7/16/2019 12:33	33	11	51	1.8	1623	1230	810	1426.5	6	No
7/16/2019 12:54	34	11	55.4	0	806	766	197	786	7	No
7/16/2019 13:24	35	11	45.7	0	338	296	129	317	2	No
7/16/2019 14:08	36	11	50	1.8	583	610	415	596.5	6	No
7/18/2019 10:40	37	11	44.6	1.77	600	810	240	705	3	No
7/18/2019 11:00	38	11	41.2	1.77	420	636	293	528	4	No

Table A.5: UBC field dataset- Part 2-I

Date-Time	WSID	building	Noise	PART	HL Lux	HR Lux	V Lux	Lux	Sat Temp	RAIN
7/18/2019 11:18	39	11	34.5	0	318	361	170	339.5	5	No
7/18/2019 14:01	40	11	46.3	0	283	329	125	306	5	No
7/18/2019 14:26	41	11	43.7	0	570	596	192	583	4	No
7/18/2019 15:01	42	11	41.1	0	425	344	314	384.5	6	No
7/18/2019 16:04	43	11	46.4	0	339	214	1300	276.5	3	No
7/16/2019 13:52	44	11	43.2	0	1121	903	432	1121	7	No
7/15/2019 15:28	45	12	34.8	0	759	569	425	664	5	No
7/15/2019 15:41	46	12	38	0	266	168	186.2	217	3	No
7/15/2019 12:22	47	12	42.4	1.37	815	504	211	659.5	6	No
7/16/2019 9:51	48	12	46.2	0	1356	920	840	1138	6	No
7/16/2019 10:03	49	12	48	0	1237	616	473	926.5	6	No
7/16/2019 10:10	50	12	48.3	0	455	425	507	440	5	No
7/16/2019 10:31	51	12	47.9	0	820	960	287	890	4	No
7/15/2019 15:18	52	12	43.6	0	166	131	100	166	2	No
7/15/2019 14:51	53	13	46.4	0	805	960	592	882.5	6	No
8/13/2019 9:46	54	14	57.3	0	275	195	143	235	7	No
8/13/2019 9:54	55	14	50.4	1.43	281	248	249	264.5	6	No
8/13/2019 10:09	56	14	51.1	0	193	294	258	243.5	6	No
8/13/2019 10:22	57	14	51.8	0	372	418	397	395	6	No
8/13/2019 10:37	58	14	53.1	1.43	390	350	300	370	4	No
8/13/2019 10:53	59	14	49.7	0	211	177	176	194	5	No
8/13/2019 11:14	60	14	51.5	1.43	546	525	407	535.5	7	No
8/13/2019 11:20	61	14	48.1	1.43	85	101	307	93	4	No
8/13/2019 11:24	62	14	50.1	1.43	462	320	272	391	4	No
8/13/2019 11:30	63	14	48.6	1.43	365	438	266	401.5	2	No
8/13/2019 11:34	64	14	48.2	1.43	144	96	317	120	6	No
8/13/2019 11:43	65	14	50.7	1.43	222	350	505	286	4	No
8/13/2019 11:52	66	14	49.6	1.43	296	160	401	228	7	No
8/13/2019 12:03	67	14	50.1	1.43	237	475	618	356	5	No
8/15/2019 10:47	68	14	51	1.16	820	782	406	801	5	No
8/15/2019 10:28	69	14	54.3	1.16	820	1300	1073	1060	3	No
8/15/2019 11:04	70	14	48.7	0	210	211	92	210.5	4	No
8/15/2019 11:15	71	14	48.5	1.34	139	148	300	143.5	6	No

Table A.6: UBC field dataset- Part 2-II

Date-Time	WSID	building	Noise	PART	HL Lux	HR Lux	V Lux	Lux	Sat Temp	RAIN
8/15/2019 11:26	72	14	44.3	1.34	1213	630	1573	921.5	7	No
8/15/2019 11:35	73	14	45.6	1.34	88	100	144	94	3	No
8/15/2019 11:39	74	14	48.2	1.34	98	85	123	91.5	5	No
8/15/2019 11:54	75	14	51.3	1.34	236	306	435	271	4	No
8/15/2019 11:58	76	14	47.8	1.34	135	144	457	139.5	2	No
8/15/2019 13:14	77	14	47.5	0	1483	2130	335	1806.5	6	No
8/15/2019 13:30	78	14	46.5	0	308	147	176	227.5	5	No
8/15/2019 13:50	79	14	47.3	1.34	1146	742	1527	944	2	No
8/15/2019 13:58	80	14	49.7	0	650	430	582	540	7	No
8/15/2019 14:05	81	14	49.7	1.34	242	1058	131	650	5	No
8/15/2019 14:12	82	14	47.9	1.34	128	132	108	130	5	No
8/15/2019 14:18	83	14	48	1.34	341	268	347	304.5	3	No
8/15/2019 14:26	84	14	48	1.34	61	65	85	63	4	No
8/15/2019 15:01	85	14	50.7	1.34	101	106	202	103.5	6	No
8/15/2019 14:49	86	14	51.2	1.34	119	74	162	96.5	5	No
8/15/2019 14:54	87	14	50.4	0	396	443	464	419.5	6	No
8/15/2019 15:07	88	14	50.8	1.34	91	101	102	96	4	No
8/15/2019 15:19	89	14	53.2	1.34	167	232	134	199.5	5	No
8/15/2019 15:26	90	14	49.2	0	314	319	507	316.5	3	No
8/15/2019 15:43	91	14	47.3	0	328	308	383	318	5	No
8/15/2019 15:47	92	14	45.1	0	102	151	106	126.5	6	No
8/15/2019 15:58	93	14	49.8	0	230	258	149	244	3	No
8/15/2019 16:04	94	14	54.5	0	124	168	101	146	6	No
1/27/2020 10:35	95	10	47.4	0	159	185	204	172	2	No
1/27/2020 10:58	96	10	52.7	0	2170	3100	15040	2635	4	No
1/27/2020 11:43	97	10	48.6	1.25	323	252	741	287.5	4	No
1/27/2020 12:15	98	10	48.5	1.25	256	159	137	207.5	3	No
1/27/2020 12:31	99	10	40.1	0	680	1140	355	910	4	No
1/27/2020 12:36	100	10	40.1	0	180	200	180	190	6	No
1/27/2020 12:53	101	10	50	0	590	1083	495	836.5	6	No
1/27/2020 14:47	102	10	43.8	1.25	482	375	185	428.5	6	No
1/27/2020 15:07	103	10	43.2	0	166	261	95	213.5	5	No
1/27/2020 15:29	104	10	51.8	0	376	354	186.7	365	7	No
1/27/2020 15:41	105	10	51	0	175.7	182	129.6	178.85	6	No
1/27/2020 15:53	106	10	54	0	192	275	164.6	233.5	4	No
1/29/2020 13:48	107	10	42.7	0	1462	1007	353	1234.5	5	No

Table A.7: UBC field dataset- Part 2-III

Date-Time	WSID	building	Noise	PART	HL Lux	HR Lux	V Lux	Lux	Sat Temp	RAIN
1/29/2020 14:19	108	10	39	1.25	587	1676	537	1131.5	4	No
1/29/2020 14:36	109	10	54.9	0	2840	1060	1759	1950	6	No
1/29/2020 14:56	110	10	41.1	1.25	252	420	712	336	6	No
1/31/2020 10:44	111	10	41.5	1.26	142	255	250	198.5	6	Yes
1/31/2020 11:05	112	10	56.4	0	237	288	193	262.5	3	Yes
1/31/2020 11:22	113	10	54.4	1.7	195.5	252	120	223.75	6	Yes
1/31/2020 11:40	114	10	44	0	180	107	98	143.5	5	Yes
1/31/2020 11:36	115	10	44	0	157.4	208	173.6	182.7	6	Yes
1/31/2020 13:32	116	10	40.1	1.25	103	136	117.5	119.5	5	Yes
1/31/2020 13:38	117	10	47.2	1.25	159	176	68	167.5	5	Yes
1/31/2020 14:07	118	10	47.1	1.25	167	285	127	226	3	Yes
1/31/2020 14:43	119	10	47.1	1.25	123.6	172	90	147.8	3	Yes
1/31/2020 14:43	120	10	40	0	215	163	85	189	3	Yes
2/3/2020 10:50	121	11	42	0	529	137	255	333	3	No
2/3/2020 11:10	122	11	44	1.68	335	393	234	364	1	No
2/3/2020 11:14	123	11	47.7	0	370	341	253	355.5	6	No
2/3/2020 11:23	124	11	49.4	0	312	409	105	360.5	6	No
2/7/2020 10:50	125	11	39	0	563	675	317	619	6	Yes
2/7/2020 11:01	126	11	46	0	985	796	330	890.5	3	Yes
2/7/2020 11:13	127	11	46.1	1.68	743	637	405	690	2	Yes
2/7/2020 11:31	128	11	48.8	1.64	478	460	211	469	6	Yes
2/7/2020 11:34	129	11	48.8	1.64	194	261	300	227.5	4	Yes
2/7/2020 11:43	130	11	46.9	1.64	730	694	364	712	3	Yes
2/10/2020 12:10	131	12	53	0	1600	536	773	1068	3	No
2/10/2020 12:24	132	12	53.8	0	1152	1970	3170	1561	1	No
2/10/2020 12:35	133	12	45.7	0	458	426	740	442	6	No
2/10/2020 13:24	134	12	63	0	169.2	128	93.9	148.6	3	No
2/11/2020 10:19	135	12	47.5	0	156	133	142	144.5	4	No
2/11/2020 10:35	136	12	61	0	2190	635	515	1412.5	3	No
2/11/2020 11:17	137	12	49.3	0	890	758	566	824	2	No
2/12/2020 14:01	138	12	48.5	0	370	419	144.3	394.5	5	No
2/12/2020 14:20	139	12	46.4	1.37	260	241	1152	250.5	4	No
2/12/2020 14:26	140	12	49.9	1.07	285	264	134	274.5	4	No
2/12/2020 14:41	141	12	52	0	149	135	233	142	3	No

Table A.8: UBC field dataset- Part 2-IV

Date-Time	WSID	building	BLINDS	BL OPER	BL DWN	CLO	DESK MAT	DESK COL	CHAIR
7/15/2019 10:39	1	10	Yes	Yes	Yes	0.41	Laminate	White	Office chair
7/15/2019 10:59	2	10	Yes	Yes	No	0.59	Laminate	White	Office chair
7/17/2019 10:03	3	10	Yes	Yes	Yes	0.72	laminate	white	Office chair
7/17/2019 10:14	4	10	Yes	Yes	Yes	0.51	laminate	white	Office chair
7/17/2019 10:28	5	10	Yes	Yes	Yes	0.59	laminate	white	Office chair
7/17/2019 11:08	6	10	Yes	Yes	No	0.54	laminate	wood	Office chair
7/17/2019 11:17	7	10	Yes	Yes	No	0.51	laminate	white	Office chair
7/17/2019 11:45	8	10	Yes	Yes	No	0.9	laminate	white	Office chair
7/17/2019 11:54	9	10	Yes	Yes	Yes	0.54	laminate	white	Office chair
7/17/2019 15:54	10	10	Yes	Yes	No	0.81	laminate	wood	Office chair
7/17/2019 16:02	11	10	Yes	Yes	No	0.39	wood	wood	Office chair
7/19/2019 10:19	12	10	Yes	Yes	No	0.78	Laminate	Wood/light brown	Office chair
7/19/2019 10:09	13	10	Yes	Yes	No	0.87	Laminate	Wood/light brown	Office chair
7/19/2019 10:38	14	10	Yes	Yes	Yes	0.57	laminate	white	Office chair
7/19/2019 11:02	15	10	Yes	Yes	No	0.81	laminate	white	Office chair
7/19/2019 11:22	16	10	Yes	Yes	No	0.59	laminate	white	Office chair
7/19/2019 11:49	17	10	Yes	Yes	Yes	0.51	laminate	light brown/wood	Office chair
7/19/2019 12:01	18	10	Yes	Yes	No	0.51	laminate	white	Office chair
7/19/2019 12:30	19	10	Yes	Yes	half	0.45	laminate	light brown/wood	Office chair
7/19/2019 12:56	20	10	Yes	Yes	Yes	0.69	laminate	white	Office chair
7/19/2019 13:18	21	10	No	Yes	half	0.53	laminate	light brown/wood	Office chair
7/19/2019 13:52	22	10	Yes	Yes	Yes	0.74	laminate	white	Office chair
7/19/2019 14:10	23	10	Yes	Yes	half	0.64	laminate	white	Office chair
7/19/2019 14:29	24	10	Yes	Yes	half	0.42	laminate	white	Office chair
7/19/2019 14:49	25	10	Yes	Yes	half	0.39	wood	brown	Office chair
7/17/2019 10:22	26	10	Yes	Yes	half	0.64	wood	brown	Office chair
7/17/2019 10:41	27	10	Yes	Yes	Yes	0.53	laminate	white	Office chair
7/15/2019 12:09	28	11	Yes	Yes	Yes	0.59	Laminate	White	Office chair
7/15/2019 13:59	29	11	No	No	No	0.57	Laminate	White	Office chair
7/16/2019 11:38	30	11	Yes	Yes	half	0.59	laminate	white	Office chair
7/16/2019 12:11	31	11	Yes	Yes	half	0.51	laminate	wood/light brown	Office chair
7/16/2019 12:22	32	11	Yes	Yes	half	0.57	laminate	white	Office chair
7/16/2019 12:33	33	11	Yes	Yes	No	0.43	laminate	white	Office chair
7/16/2019 12:54	34	11	Yes	Yes	half	0.51	laminate	light brown	Office chair
7/16/2019 13:24	35	11	No	N/A	N/a	0.37	laminate	light brown	Office chair
7/16/2019 14:08	36	11	Yes	Yes	Yes	0.59	laminate	white	Office chair
7/18/2019 10:40	37	11	Yes	Yes	No	0.6	Wood	Brown	Office chair
7/18/2019 11:00	38	11	Yes	Yes	No	0.42	Wood	Brown	Office chair

Table A.9: UBC field dataset- Part 3-I

Date-Time	WSID	building	BLINDS	BL OPER	BL DWN	CLO	DESK MAT	DESK COL	CHAIR
7/18/2019 11:18	39	11	Yes	Yes	Yes	0.6	Wood	Brown	Office chair
7/18/2019 14:01	40	11	Yes	Yes	Yes	0.53	Wood	Brown	Office chair
7/18/2019 14:26	41	11	Yes	Yes	No	0.54	wood	brown	Office chair
7/18/2019 15:01	42	11	Yes	Yes	Yes	0.44	wood	brown	yoga ball
7/18/2019 16:04	43	11	No	No	No	0.79	wood	brown	Office chair
7/16/2019 13:52	44	11	Yes	Yes	half	0.51	laminate	brown	Office chair
7/15/2019 15:28	45	12	No	No	No	0.53	Laminate	Black	Office chair
7/15/2019 15:41	46	12	Yes	Yes	Half	0.47	Laminate	Beige	Stool
7/15/2019 12:22	47	12	Yes	Yes	No	0.87	laminate	browm	Office chair
7/16/2019 9:51	48	12	Yes	Yes	Yes	0.57	laminate	Beige	Office chair
7/16/2019 10:03	49	12	Yes	Yes	Yes	0.57	laminate	brown	Office chair
7/16/2019 10:10	50	12	Yes	Yes	No	0.59	laminate	white	Office chair
7/16/2019 10:31	51	12	Yes	Yes	No	0.53	laminate	white	Office chair
7/15/2019 15:18	52	12	Yes	Yes	Yes	0.31	Laminate	Black	None
7/15/2019 14:51	53	13	Yes	Yes	No	0.53	Laminate	White	Office chair
8/13/2019 9:46	54	14	Yes	Yes	No	0.33	Wood	light brown	None
8/13/2019 9:54	55	14	Yes	Yes	No	0.57	Wood	light brown	Office chair
8/13/2019 10:09	56	14	Yes	Yes	No	0.57	Wood	light brown	Office chair
8/13/2019 10:22	57	14	Yes	Yes	No	0.43	Wood	light brown	Office chair
8/13/2019 10:37	58	14	Yes	Yes	No	0.47	Wood	light brown	Office chair
8/13/2019 10:53	59	14	Yes	Yes	No	0.39	Wood	light brown	Office chair
8/13/2019 11:14	60	14	Yes	Yes	No	0.57	Wood	light brown	Office chair
8/13/2019 11:20	61	14	Yes	Yes	No	0.6	Wood	light brown	Office chair
8/13/2019 11:24	62	14	Yes	Yes	No	0.67	Wood	light brown	Office chair
8/13/2019 11:30	63	14	Yes	Yes	No	0.75	Wood	light brown	Office chair
8/13/2019 11:34	64	14	Yes	Yes	No	0.51	Wood	light brown	Office chair
8/13/2019 11:43	65	14	Yes	Yes	No	0.43	Wood	light brown	Office chair
8/13/2019 11:52	66	14	Yes	Yes	No	0.59	Wood	light brown	Office chair
8/13/2019 12:03	67	14	Yes	Yes	No	0.51	Wood	light brown	Office chair
8/15/2019 10:47	68	14	Yes	Yes	Yes	0.53	Wood	Light brown	Office chair
8/15/2019 10:28	69	14	Yes	Yes	Yes	0.51	Wood	Light brown	Office chair
8/15/2019 11:04	70	14	No	N/A	N/A	0.6	Wood	Light brown	Office chair
8/15/2019 11:15	71	14	Yes	Yes	Partially	0.55	Wood	Light brown	Office chair

Table A.10: UBC field dataset- Part 3-II

Date-Time	WSID	building	BLINDS	BL OPER	BL DWN	CLO	DESK MAT	DESK COL	CHAIR
8/15/2019 11:26	72	14	Yes	Yes	Partially	0.3	Wood	Light brown	Office chair
8/15/2019 11:35	73	14	Yes	Yes	Partially	0.74	Wood	Light brown	Office chair
8/15/2019 11:39	74	14	Yes	Yes	Partially	0.74	Wood	Light brown	Office chair
8/15/2019 11:54	75	14	Yes	Yes	No	0.56	Wood	Light brown	Office chair
8/15/2019 11:58	76	14	Yes	Yes	No	0.59	Wood	Light brown	Office chair
8/15/2019 13:14	77	14	Yes	Yes	Partially	0.87	Wood	Light brown	Office chair
8/15/2019 13:30	78	14	Yes	Yes	No	0.59	Laminate	White	Office chair
8/15/2019 13:50	79	14	Yes	Yes	No	0.77	Wood	Light brown	Office chair
8/15/2019 13:58	80	14	Yes	Yes	No	0.81	Wood	Light brown	Office chair
8/15/2019 14:05	81	14	N/A	N/A	N/A	0.44	Wood	Light brown	Office chair
8/15/2019 14:12	82	14	Yes	Yes	Yes	0.53	Wood	Light brown	Office chair
8/15/2019 14:18	83	14	Yes	Yes	Yes	0.53	Wood	Light brown	Office chair
8/15/2019 14:26	84	14	N/A	N/A	N/A	0.37	Wood	Light brown	Office chair
8/15/2019 15:01	85	14	N/A	N/A	N/A	0.51	Wood	Light brown	Office chair
8/15/2019 14:49	86	14	N/A	N/A	N/A	0.44	Wood	Light brown	Office chair
8/15/2019 14:54	87	14	N/A	N/A	N/A	0.59	Wood	Light brown	Office chair
8/15/2019 15:07	88	14	N/A	N/A	N/A	0.79	Wood	Light brown	Office chair
8/15/2019 15:19	89	14	N/A	N/A	N/A	0.45	Wood	Light brown	Office chair
8/15/2019 15:26	90	14	Yes	Yes	Partially	0.44	Wood	Light brown	Office chair
8/15/2019 15:43	91	14	Yes	Yes	No	0.81	Wood	Light brown	Office chair
8/15/2019 15:47	92	14	Yes	Yes	Yes	0.51	Wood	Light brown	Office chair
8/15/2019 15:58	93	14	Yes	Yes	Yes	0.51	Wood	Light brown	Office chair
8/15/2019 16:04	94	14	Yes	Yes	Yes	0.49	Wood	Light brown	Office chair
1/27/2020 10:35	95	10	Yes	Yes	No	0	Laminate	White	Office chair
1/27/2020 10:58	96	10	No	No	No	0	Laminate	White	Office chair
1/27/2020 11:43	97	10	Yes	Yes	Yes	0	wood	brown	Office chair
1/27/2020 12:15	98	10	Yes	Yes	No	0	wood	wood	Office chair
1/27/2020 12:31	99	10	Yes	Yes	No	0	Laminate	White	Office chair
1/27/2020 12:36	100	10	Yes	Yes	No	0	Laminate	White	Office chair
1/27/2020 12:53	101	10	Yes	Yes	No	0	laminate	white	Office chair
1/27/2020 14:47	102	10	No	No	No	0	laminate	white	Office chair
1/27/2020 15:07	103	10	Yes	Yes	Yes	0	laminate	white	Office chair
1/27/2020 15:29	104	10	Yes	Yes	No	0	laminate	white	Office chair
1/27/2020 15:41	105	10	Yes	Yes	Yes	0	laminate	white	Office chair
1/27/2020 15:53	106	10	Yes	Yes	Yes	0	laminate	white	Office chair
1/29/2020 13:48	107	10	No	No	No	0	laminate	white	Office chair

Table A.11: UBC field dataset- Part 3-III

Date-Time	WSID	building	BLINDS	BL OPER	BL DWN	CLO	DESK MAT	DESK COL	CHAIR
1/29/2020 14:19	108	10	Yes	Yes	Yes	0	laminate	white	Office chair
1/29/2020 14:36	109	10	No	No	No	0	laminate	white	Office chair
1/29/2020 14:56	110	10	Yes	Yes	Yes	0	laminate	white	Office chair
1/31/2020 10:44	111	10	Yes	Yes	No	0	Laminate	White	Office chair
1/31/2020 11:05	112	10	Yes	Yes	Yes	0	Laminate	White	Office chair
1/31/2020 11:22	113	10	Yes	Yes	Yes	0	Laminate	White	Office chair
1/31/2020 11:40	114	10	Yes	Yes	HALF	0	Laminate	White	Office chair
1/31/2020 11:36	115	10	Yes	Yes	HALF	0	Laminate	White	Office chair
1/31/2020 13:32	116	10	Yes	Yes	No	0	Wood	Brown	Office chair
1/31/2020 13:38	117	10	Yes	Yes	No	0	Wood	Brown	Office chair
1/31/2020 14:07	118	10	Yes	Yes	HALF	0	laminate	white	Office chair
1/31/2020 14:43	119	10	Yes	Yes	HALF	0	laminate	white	Office chair
1/31/2020 14:43	120	10	Yes	Yes	No	0	laminate	White	Office chair
2/3/2020 10:50	121	11	Yes	Yes	No	0	Wood	Brown	Office chair
2/3/2020 11:10	122	11	Yes	Yes	Yes	0	Laminate	White	Office chair
2/3/2020 11:14	123	11	Yes	Yes	Half	0	wood	brown	Standing
2/3/2020 11:23	124	11	Yes	Yes	Yes	0	Laminate	White	Office chair
2/7/2020 10:50	125	11	Yes	Yes	Half	0	wood	brown	Office chair
2/7/2020 11:01	126	11	Yes	Yes	No	0	wood	brown	Office chair
2/7/2020 11:13	127	11	Yes	Yes	No	0	Laminate	White	Office chair
2/7/2020 11:31	128	11	Yes	Yes	No	0	Wood	Brown	Office chair
2/7/2020 11:34	129	11	Yes	Yes	No	0	Wood	Brown	Office chair
2/7/2020 11:43	130	11	Yes	Yes	No	0	Wood	Brown	Office chair
2/10/2020 12:10	131	12	Yes	Yes	Half	0	Wood	Brown	Office chair
2/10/2020 12:24	132	12	No	NA	No	0	Wood	Brown	Office chair
2/10/2020 12:35	133	12	Yes	Yes	Half	0	Wood	Brown	Office chair
2/10/2020 13:24	134	12	No	NA	NA	0	Wood	Brown	Office chair
2/11/2020 10:19	135	12	Yes	Yes	half	0	wood	Black	Office chair
2/11/2020 10:35	136	12	Yes	Yes	half	0	wood	browm	Office chair
2/11/2020 11:17	137	12	No	NA	NA	0	wood	browm	Office chair
2/12/2020 14:01	138	12	No	No	No	0	wood	Brown	Office chair
2/12/2020 14:20	139	12	No	No	No	0	wood	Brown	Office chair
2/12/2020 14:26	140	12	Yes	Yes	No	0	wood	Brown	Office chair
2/12/2020 14:41	141	12	Yes	Yes	Yes	0	wood	Light brown	Office chair

Table A.12: UBC field dataset- Part 3-IV

Date-Time	WSID	building#	CHAIR MAT	FLOOR MAT	FLOOR COL	WALL COL	CEILING COL
7-15-19 10:39 AM	1	10	mesh	Carpet	Grey	White	White + wood
7-15-19 10:59 AM	2	10	mesh	Carpet	Grey	White	White + wood
7-17-19 10:03 AM	3	10	mesh	Carpet	Grey	White	wood
7-17-19 10:14 AM	4	10	mesh	Carpet	Grey	White	wood
7-17-19 10:28 AM	5	10	mesh	Carpet	Grey	White	wood
7-17-19 11:08 AM	6	10	Cloth	Carpet	Grey	White	wood
7-17-19 11:17 AM	7	10	mesh	Carpet	Grey	White	White
7-17-19 11:45 AM	8	10	mesh	Carpet	Grey	White	White+ wood
7-17-19 11:54 AM	9	10	mesh	Carpet	Grey	White	White
7-17-19 3:54 PM	10	10	mesh	Carpet	Grey	White	wood+ White
7-17-19 4:02 PM	11	10	mesh	Carpet	Grey	White	wood+ White
7-19-19 10:19 AM	12	10	Cloth	Carpet	Grey	White	Wood
7-19-19 10:09 AM	13	10	Cloth	Carpet	Grey	White	Wood
7-19-19 10:38 AM	14	10	mesh	Carpet	Grey	wood	White
7-19-19 11:02 AM	15	10	mesh	Carpet	Grey	White	White
7-19-19 11:22 AM	16	10	mesh	Carpet	Grey	White	White+ wood
7-19-19 11:49 AM	17	10	mesh	Carpet	Grey	White	White+ wood
7-19-19 12:01 PM	18	10	mesh	Carpet	Grey	White	White+ wood
7-19-19 12:30 PM	19	10	Cloth	Carpet	Grey	White	White+ wood
7-19-19 12:56 PM	20	10	mesh	Carpet	Grey	White	White+ wood
7-19-19 1:18 PM	21	10	Cloth	Carpet	Grey	White	White + wood
7-19-19 1:52 PM	22	10	mesh	Carpet	Grey	White and wood	White+ wood
7-19-19 2:10 PM	23	10	mesh	Carpet	Grey	White	White
7-19-19 2:29 PM	24	10	mesh	Carpet	Grey	White	wood
7-19-19 2:49 PM	25	10	mesh	Carpet	Grey	White and wood	White+ wood
7-17-19 10:22 AM	26	10	Office chair	mesh	Carpet	Grey	White
7-17-19 10:41 AM	27	10	Office chair	mesh	Carpet	Grey	White
7-15-19 12:09 PM	28	11	leather	Carpet	Grey	White	Grey
7-15-19 1:59 PM	29	11	mesh	Carpet	Grey	White	Grey
7-16-19 11:38 AM	30	11	leather	Carpet	dark Grey	White	White
7-16-19 12:11 PM	31	11	Leather	Carpet	Gray	White	White
7-16-19 12:22 PM	32	11	leather	Carpet	Grey	White	White
7-16-19 12:33 PM	33	11	mesh	Carpet	Grey	White	White
7-16-19 12:54 PM	34	11	mesh	Carpet	Grey	White	White
7-16-19 1:24 PM	35	11	fabric	Carpet	Grey	White	White
7-16-19 2:08 PM	36	11	Leather	Carpet	Grey	White	White
7-18-19 10:40 AM	37	11	Cloth	Carpet	Grey	Beige	White
7-18-19 11:00 AM	38	11	Mesh	Carpet	Grey	Beige	White
7-18-19 11:18 AM	39	11	Mesh	Carpet	Grey	White	White
7-18-19 2:01 PM	40	11	Cloth	Carpet	Grey	White	White

Table A.13: UBC field dataset- Part 4-I

Date-Time	WSID	building#	CHAIR MAT	FLOOR MAT	FLOOR COL	WALL COL	CEILING COL
7-18-19 2:26 PM	41	11	leather	Carpet	Grey	White	White
7-18-19 3:01 PM	42	11	rubber	Carpet	Blue	White	White
7-18-19 4:04 PM	43	11	mesh	Carpet	blue/Grey	White	White
7-16-19 1:52 PM	44	11	Office chair	mesh	Carpet	dark blue and beige	White
7-15-19 3:28 PM	45	12	Cloth	Carpet	Grey	Grey and Chartruse	N/A
7-15-19 3:41 PM	46	12	Cloth	Linoleum	Grey	White	Brown
7-15-19 12:22 PM	47	12	mesh	linoleum	Grey	White	White
7-16-19 9:51 AM	48	12	mesh	lionelum	Grey	White	wood
7-16-19 10:03 AM	49	12	mesh	linoleum	Grey	White	wood
7-16-19 10:10 AM	50	12	Cloth	linoleum	Grey	White	wood
7-16-19 10:31 AM	51	12	plastic	linoleum	Grey	White	White
7-15-19 3:18 PM	52	12	None	N/A	Carpet	Darkish Gray	Grey
7-15-19 2:51 PM	53	13	mesh	Carpet	Grey	White	White
8-13-19 9:46 AM	54	14	#N/A	Carpet	Grey	White	White
8-13-19 9:54 AM	55	14	Cloth	Carpet	Grey	White	White
8-13-19 10:09 AM	56	14	Cloth	Carpet	Grey	White	White
8-13-19 10:22 AM	57	14	Cloth	Carpet	Grey	White	White
8-13-19 10:37 AM	58	14	Cloth	Carpet	Grey	Blue (navy-ish)	White
8-13-19 10:53 AM	59	14	Cloth	Carpet	Grey	White	White
8-13-19 11:14 AM	60	14	mesh	Carpet	Grey	White	White
8-13-19 11:20 AM	61	14	Cloth	Carpet	Grey	White	White
8-13-19 11:24 AM	62	14	Cloth	Carpet	Grey	Blue (navy-ish)	White
8-13-19 11:30 AM	63	14	Cloth	Carpet	Grey	Blue (navy-ish)	White
8-13-19 11:34 AM	64	14	Cloth	Carpet	Grey	White	White
8-13-19 11:43 AM	65	14	Cloth	Carpet	Grey	White	White
8-13-19 11:52 AM	66	14	Cloth	Carpet	Grey	Blue (navy-ish)	White
8-13-19 12:03 PM	67	14	Cloth	Carpet	Grey	White	White
8-15-19 10:47 AM	68	14	Cloth	Carpet	Dark Grey	White	White+ wood
8-15-19 10:28 AM	69	14	Cloth	Carpet	Dark Grey	White	White wood
8-15-19 11:04 AM	70	14	Cloth	Carpet	Grey	White	White
8-15-19 11:15 AM	71	14	Cloth	Carpet	Grey	White	White
8-15-19 11:26 AM	72	14	Cloth	Carpet	Grey	White	White
8-15-19 11:35 AM	73	14	Cloth	Carpet	Grey	White	White
8-15-19 11:39 AM	74	14	Cloth	Carpet	Grey	White	White
8-15-19 11:54 AM	75	14	Cloth	Carpet	Grey	White	White
8-15-19 11:58 AM	76	14	Cloth	Carpet	Grey	White	White
8-15-19 1:14 PM	77	14	Cloth	Carpet	Grey	White	White
8-15-19 1:30 PM	78	14	Cloth	Carpet	Grey	White	White
8-15-19 1:50 PM	79	14	Cloth	Carpet	Grey	White	White
8-15-19 1:58 PM	80	14	Cloth	Carpet	Grey	White	White

Table A.14: UBC field dataset- Part 4-II

Date-Time	WSID	building#	CHAIR MAT	FLOOR MAT	FLOOR COL	WALL COL	CEILING COL
8-15-19 2:05 PM	81	14	Cloth	Carpet	Grey	White	White
8-15-19 2:12 PM	82	14	Cloth	Carpet	Grey	White	White
8-15-19 2:18 PM	83	14	Cloth	Carpet	Grey	White	White
8-15-19 2:26 PM	84	14	Cloth	Carpet	Grey	White	White
8-15-19 3:01 PM	85	14	Cloth	Carpet	Grey	White	White
8-15-19 2:49 PM	86	14	Cloth	Carpet	Grey	White	White
8-15-19 2:54 PM	87	14	Cloth	Carpet	Grey	White	White
8-15-19 3:07 PM	88	14	Cloth	Carpet	Grey	White	White
8-15-19 3:19 PM	89	14	Cloth	Carpet	Grey	White	White
8-15-19 3:26 PM	90	14	Cloth	Carpet	Grey	White	White
8-15-19 3:43 PM	91	14	Cloth	Carpet	Grey	White	White
8-15-19 3:47 PM	92	14	Cloth	Carpet	Grey	White	White
8-15-19 3:58 PM	93	14	Cloth	Carpet	Grey	White	White
8-15-19 4:04 PM	94	14	Cloth	Carpet	Grey	White	White
1-27-20 10:35 AM	95	10	Office chair	Cloth	Carpet	Grey	White
1-27-20 10:58 AM	96	10	Office chair	Cloth	Carpet	Grey	White
1-27-20 11:43 AM	97	10	Office chair	Cloth	Carpet	Grey	Cream
1-27-20 12:15 PM	98	10	Office chair	Cloth	Carpet	Grey	Cream
1-27-20 12:31 PM	99	10	Office chair	Cloth	Carpet	Grey	Cream
1-27-20 12:36 PM	100	10	Office chair	Cloth	Carpet	Grey	Cream
1-27-20 12:53 PM	101	10	Office chair	Cloth	Carpet	Grey	Cream
1-27-20 2:47 PM	102	10	Office chair	Cloth	Carpet	Grey	White
1-27-20 3:07 PM	103	10	Office chair	Cloth	Carpet	Grey	White
1-27-20 3:29 PM	104	10	Office chair	Cloth	Carpet	Grey	wood
1-27-20 3:41 PM	105	10	Office chair	Cloth	Carpet	Grey	wood
1-27-20 3:53 PM	106	10	Office chair	Cloth	Carpet	Grey	wood
1-29-20 1:48 PM	107	10	Office chair	Cloth	Carpet	Grey	White
1-29-20 2:19 PM	108	10	Office chair	Cloth	Carpet	Grey	White
1-29-20 2:36 PM	109	10	Office chair	Cloth	Carpet	Grey	White
1-29-20 2:56 PM	110	10	Office chair	Cloth	Carpet	Grey	White
1-31-20 10:44 AM	111	10	Office chair	Cloth	Carpet	Grey	White
1-31-20 11:05 AM	112	10	Office chair	Cloth	Carpet	Grey	White
1-31-20 11:22 AM	113	10	Office chair	Cloth	Carpet	Grey	wood
1-31-20 11:40 AM	114	10	Office chair	Cloth	Carpet	Grey	wood

Table A.15: UBC field dataset- Part 4-III

Date-Time	WSID	building#	CHAIR MAT	FLOOR MAT	FLOOR COL	WALL COL	CEILING COL
1-31-20 11:36 AM	115	10	Office chair	Cloth	Carpet	Grey	wood
1-31-20 1:32 PM	116	10	Office chair	Cloth	Carpet	Grey	White
1-31-20 1:38 PM	117	10	Office chair	Cloth	Carpet	Grey	White
1-31-20 2:07 PM	118	10	Office chair	Cloth	Carpet	Grey	White
1-31-20 2:43 PM	119	10	Office chair	Cloth	Carpet	Grey	White
1-31-20 2:43 PM	120	10	Office chair	Cloth	Carpet	Grey	White
2002-03-20 10:50	121	11	Office chair	Cloth	Carpet	Grey	White
2002-03-20 11:10	122	11	Office chair	Cloth	Carpet	Grey	White
2002-03-20 11:14	123	11	standing	NA	Carpet	Grey	White
2002-03-20 11:23	124	11	Office chair	Cloth	Carpet	Grey	White
2002-07-20 10:50	125	11	Office chair	Cloth	Carpet	blue	White
2002-07-20 11:01	126	11	Office chair	Cloth	Carpet	BLUE	White
2002-07-20 11:13	127	11	Office chair	Cloth	Carpet	Grey	White
2002-07-20 11:31	128	11	Office chair	Cloth	Carpet	Brown	White
2002-07-20 11:34	129	11	Office chair	Cloth	Carpet	Brown	White
2002-07-20 11:43	130	11	Office chair	Cloth	Carpet	Brown	White
2002-10-20 12:10	131	12	Office chair	Plastic	Concrete	Grey	White
2002-10-20 12:24	132	12	Office chair	Mesh	Concrete	Grey	White
2002-10-20 12:35	133	12	Office chair	Clotj	Concrete	Grey	White
2002-10-20 13:24	134	12	stool	wood	Concrete	Grey	White
2002-11-20 10:19	135	12	Standing	NA	Carpet	Grey	green+ Grey
2002-11-20 10:35	136	12	Office chair	plastic	concrete	Grey	White+ blue
2002-11-20 11:17	137	12	Office chair	Cloth	Carpet	Grey	White
2002-12-20 14:01	138	12	Office chair	Cloth	concrete	Grey	White
2002-12-20 14:20	139	12	Office chair	Cloth	concrete	Grey	White
2002-12-20 14:26	140	12	Office chair	plastic	concrete	Grey	White
2002-12-20 14:41	141	12	Office chair	Cloth	concrete	Grey	White

Table A.16: UBC field dataset- Part 4-IV

Date-Time	WSID	building	PART COL	LAMP	LAMP ON	CEIL H	DIST PLANT	WIND OPEN	DIST WIND1	WIND LOC1
7/15/19 10:39 AM	1	10	N/A	No	No	3.7	N/A	Yes	0	S
7/15/19 10:59 AM	2	10	N/A	No	No	3.7	N/A	Yes	0	N
7/17/19 10:03 AM	3	10	grey	No	No	3.6	N/A	No	3.9	SE
7/17/19 10:14 AM	4	10	grey	No	No	3.6	1.6	No	0.4	SE
7/17/19 10:28 AM	5	10	grey	No	No	3.6	2.44	No	2.46	SE
7/17/19 11:08 AM	6	10	grey	No	No	3.7	0.1	No	3	SE
7/17/19 11:17 AM	7	10	grey	No	No	3.09	0.1	Yes	1.4	NW
7/17/19 11:45 AM	8	10	grey	No	No	3.53	N/A	No	0.3	SE
7/17/19 11:54 AM	9	10	N/A	No	No	2.74	N/A	No	2.8	SE
7/17/19 3:54 PM	10	10	N/A	No	No	3.52	1.6	Yes	1.6	SE
7/17/19 4:02 PM	11	10	grey	No	No	3.68	1.5	Yes	3.5	NW
7/19/19 10:19 AM	12	10	grey	No	No	3.6	2.47	No	3.1	NW
7/19/19 10:09 AM	13	10	grey	Yes	No	3.14	3.8	No	3.8	NW
7/19/19 10:38 AM	14	10	N/A	No	No	3.68	2.6	Yes	0	SE
7/19/19 11:02 AM	15	10	N/A	No	Yes	3.11	1.8	Yes	1.8	NW
7/19/19 11:22 AM	16	10	N/A	No	No	3.7	N/A	No	0.7	NW
7/19/19 11:49 AM	17	10	grey	No	No	3.6	0	Yes	3.2	E
7/19/19 12:01 PM	18	10	N/A	No	No	3.53	0	Yes	0	NW
7/19/19 12:30 PM	19	10	grey	No	No	3.6	0	Yes	3.3	NW
7/19/19 12:56 PM	20	10	N/A	No	No	3.6	N/A	No	2.87	S
7/19/19 1:18 PM	21	10	grey	Yes	No	3.7	0	No	2.9	SE
7/19/19 1:52 PM	22	10	N/A	No	No	3.09	0	No	2.05	W
7/19/19 2:10 PM	23	10	grey	No	No	2.7	N/A	Yes	0	SE
7/19/19 2:29 PM	24	10	N/A	Yes	Yes	3.6	N/A	No	0	N
7/19/19 2:49 PM	25	10	N/A	Yes	No	3.7	N/A	Yes	0	SE
7/17/19 10:22 AM	26	10	wood	grey	No	3.6	0.3	No	0.4	SE
7/17/19 10:41 AM	27	10	wood	N/A	No	3.6	N/A	Yes	2.5	SE
7/15/19 12:09 PM	28	11	gray	No	No	2.6	5	No	2.58	NW
7/15/19 1:59 PM	29	11	gray	No	No	2.6	1	No	7	NW
7/16/19 11:38 AM	30	11	brown	No	No	2.59	1.7	No	3	NW
7/16/19 12:11 PM	31	11	N/a	No	No	2.58	N/A	No	2	NW
7/16/19 12:22 PM	32	11	brown	No	No	2.59	1.37	No	0	NW
7/16/19 12:33 PM	33	11	white	No	No	2.59	1.4	No	2.1	NW
7/16/19 12:54 PM	34	11	N/A	No	No	2.92	N/A	Yes	0.79	NE
7/16/19 1:24 PM	35	11	N/A	No	No	2.92	N/A	N/A	N/A	NE
7/16/19 2:08 PM	36	11	grey	No	No	2.59	N/A	No	1.31	NW
7/18/19 10:40 AM	37	11	White	No	No	2.712	3	No	3.3	W
7/18/19 11:00 AM	38	11	White	No	No	2.712	3.3	No	3.5	SW
7/18/19 11:18 AM	39	11	N/a	No	No	2.9	N/A	No	0	SW
7/18/19 2:01 PM	40	11	N/A	Yes	No	2.922	N/A	No	N/A	N/A
7/18/19 2:26 PM	41	11	N/A	No	No	2.6	1.2	No	1.2	NW
7/18/19 3:01 PM	42	11	N/A	No	No	2.7	N/A	Yes	0.2	SE
7/18/19 4:04 PM	43	11	N/A	No	No	2.75	2.5	Yes	0.2	NW
7/16/19 1:52 PM	44	11	white	N/A	No	2.72	1.46	Yes	1.37	S
7/15/19 3:28 PM	45	12	N/A	No	No	3.1	N/A	Yes	N/A	N/A
7/15/19 3:41 PM	46	12	N/a	No	No	2.45	1.9	No	0.9	NW
7/15/19 12:22 PM	47	12	brown	No	No	2.37	1.07	Yes	1.25	NW
7/16/19 9:51 AM	48	12	N/A	Yes	Yes	2.3	1.7	Yes	0.48	SW
7/16/19 10:03 AM	49	12	N/A	Yes	Yes	2.3	0	Yes	0.48	SW
7/16/19 10:10 AM	50	12	N/A	Yes	No	2.3	0.4	No	0	North
7/16/19 10:31 AM	51	12	N/A	No	No	3.54	N/A	No	2.08	E
7/15/19 3:18 PM	52	12	N/A	N/A	No	3.1	N/A	Yes	2.1	SE
7/15/19 2:51 PM	53	13	N/A	No	No	3.1	N/A	No	0	NW
8/13/19 9:46 AM	54	14	none	Yes	No	2.1	0.1	No	2	S
8/13/19 9:54 AM	55	14	white	No	No	2.74	0.3	No	7	SE

Table A.17: UBC field dataset- Part 5-I

Date-Time	WSID	building	PART COL	LAMP	LAMP ON	CEIL H	DIST PLANT	WIND OPEN	DIST WIND1	WIND LOC1
8/13/19 10:09 AM	56	14	white	No	No	2.74	0.5	No	8.5	SE
8/13/19 10:22 AM	57	14	white	No	No	2.74	0.5	No	8.5	SE
8/13/19 10:37 AM	58	14	White	No	No	2.74	4.5	No	9	S
8/13/19 10:53 AM	59	14	N/A	No	No	2.74	N/A	No	6.3	SW
8/13/19 11:14 AM	60	14	N/A	No	No	2.74	N/A	No	N/A	N/A
8/13/19 11:20 AM	61	14	N/A	No	No	2.74	N/A	No	N/A	N/A
8/13/19 11:24 AM	62	14	White	No	No	2.74	0.3	No	N/A	N/A
8/13/19 11:30 AM	63	14	White	No	No	2.74	N/A	No	N/A	N/A
8/13/19 11:34 AM	64	14	N/A	No	No	2.74	N/A	No	N/A	N/A
8/13/19 11:43 AM	65	14	N/A	No	No	2.74	N/A	No	N/A	N/A
8/13/19 11:52 AM	66	14	White	No	No	2.74	N/A	No	N/A	N/A
8/13/19 12:03 PM	67	14	White	No	No	2.74	0.2	No	N/A	N/A
8/15/19 10:47 AM	68	14	grey	No	No	3.08	2	No	0.3	W
8/15/19 10:28 AM	69	14	grey	No	No	3.08	3	No	0.3	W
8/15/19 11:04 AM	70	14	N/A	No	No	2.74	N/A	No	N/A	N/A
8/15/19 11:15 AM	71	14	grey	No	No	2.74	0.5	No	3	N/A
8/15/19 11:26 AM	72	14	grey	No	No	2.74	0.5	No	0.5	N
8/15/19 11:35 AM	73	14	grey	No	No	2.74	1	No	4	N
8/15/19 11:39 AM	74	14	grey	No	No	2.74	0.5	No	4	N
8/15/19 11:54 AM	75	14	grey	No	No	2.74	1	No	7.38	N
8/15/19 11:58 AM	76	14	grey	No	No	2.74	1	No	3	N
8/15/19 1:14 PM	77	14	N/A	No	No	2.7	0.2	No	0.3	SW
8/15/19 1:30 PM	78	14	N/A	No	No	2.7	N/A	No	0.3	SW
8/15/19 1:50 PM	79	14	grey	No	No	2.74	3	No	0.3	NE
8/15/19 1:58 PM	80	14	grey	No	No	2.74	2	No	0.3	N
8/15/19 2:05 PM	81	14	grey	Yes	Yes	2.74	0.5	N/A	N/A	N/A
8/15/19 2:12 PM	82	14	grey	No	No	2.74	N/A	No	2.5	E
8/15/19 2:18 PM	83	14	grey	No	No	2.74	N/A	No	0.5	E
8/15/19 2:26 PM	84	14	grey	Yes	No	2.74	1.5	N/A	N/A	N/A
8/15/19 3:01 PM	85	14	grey	No	No	2.74	0.5	N/A	N/A	N/A
8/15/19 2:49 PM	86	14	grey	Yes	Yes	2.74	0.5	N/A	N/A	N/A
8/15/19 2:54 PM	87	14	grey	No	No	2.74	4.5	N/A	N/A	N/A
8/15/19 3:07 PM	88	14	grey	No	No	2.74	N/A	N/A	N/A	N/A
8/15/19 3:19 PM	89	14	grey	No	No	2.74	1.5	N/A	N/A	N/A
8/15/19 3:26 PM	90	14	N/A	No	No	2.74	N/A	No	0.5	SW
8/15/19 3:43 PM	91	14	N/A	Yes	No	2.74	0.5	No	0.5	SW
8/15/19 3:47 PM	92	14	N/A	Yes	No	2.74	1	No	0.5	SW
8/15/19 3:58 PM	93	14	N/A	Yes	No	2.74	0.7	No	3	SW
8/15/19 4:04 PM	94	14	N/A	Yes	No	2.74	1	No	0.5	SW
1/27/20 10:35 AM	95	10	White	N/A	No	3	0.4	No	3	SW
1/27/20 10:58 AM	96	10	White	N/A	No	3	6.3	No	0.4	SE
1/27/20 11:43 AM	97	10	Wood	Grey	No	3.7	N/A	No	2.76	NW
1/27/20 12:15 PM	98	10	White + wood	Grey	No	3.7	3.3	No	2.5	NW
1/27/20 12:31 PM	99	10	White + wood	N/A	No	3.5	N/A	No	1.4	N/A
1/27/20 12:36 PM	100	10	White + wood	N/A	No	3.5	N/A	No	3.7	N/A
1/27/20 12:53 PM	101	10	Wood	N/A	Yes	3.7	N/A	No	1.1	NW
1/27/20 2:47 PM	102	10	White + wood	Grey	No	3.7	1	No	1.4	N
1/27/20 3:07 PM	103	10	White + wood	N/A	No	3.7	1.4	Yes	1	SE
1/27/20 3:29 PM	104	10	White + wood	N/A	No	3.7	1.3	No	1.3	W

Table A.18: UBC field dataset- Part 5-II

Date-Time	WSID	building	PART COL	LAMP	LAMP ON	CEIL H	DIST PLANT	WIND OPEN	DIST WIND1	WIND LOC1
1/27/20 3:41 PM	105	10	White + wood	N/A	No	3.7	0	No	2.2	SW
1/27/20 3:53 PM	106	10	White + wood	N/A	No	3.7	0.7	No	1.2	W
1/29/20 1:48 PM	107	10	White + wood	N/A	No	3.7	N/A	No	1.4	NW
1/29/20 2:19 PM	108	10	White	N/A	No	3.5	1.7	No	1.4	NW
1/29/20 2:36 PM	109	10	White	N/A	No	3.1	N/A	No	2.3	SW
1/29/20 2:56 PM	110	10	White + wood	gREY	No	3.7	N/A	No	N/A	N/A
1/31/20 10:44 AM	111	10	White + wood	Grey	Yes	3.7	N/A	No	2.5	NW
1/31/20 11:05 AM	112	10	White + wood	N/A	No	3.5	N/A	No	0.4	SE
1/31/20 11:22 AM	113	10	White + wood	Grey	No	3.5	N/A	No	0.4	SE
1/31/20 11:40 AM	114	10	White + wood	N/A	No	3.7	5.5	No	3.3	SE
1/31/20 11:36 AM	115	10	White + wood	N/A	No	3.7	7.3	No	3.3	SE
1/31/20 1:32 PM	116	10	White + wood	Grety	No	3.7	0.1	No	3.3	NW
1/31/20 1:38 PM	117	10	White + wood	Grey	No	3.7	1	No	1	NW
1/31/20 2:07 PM	118	10	White + wood	Grey	No	3.7	N/A	No	1	SE
1/31/20 2:43 PM	119	10	White + wood	Grey	No	3.7	N/A	No	2	SE
1/31/20 2:43 PM	120	10	White + wood	Grey	No	3.7	N/A	No	0.6	E
2/3/20 10:50 AM	121	11	White	N/A	Yes	2.58	N/A	No	3.5	W
2/3/20 11:10 AM	122	11	White	Grey	No	2.58	N/A	No	1	NW
2/3/20 11:14 AM	123	11	White	N/A	No	2.58	N/A	No	1.5	W
2/3/20 11:23 AM	124	11	White	N/A	No	2.58	2.9	No	1	NW
2/7/20 10:50 AM	125	11	white	N/A	No	2.44	0.3	No	1	SE
2/7/20 11:01 AM	126	11	WHITE	N/A	No	2.44	2.7	No	0.5	NW
2/7/20 11:13 AM	127	11	White	Grey	No	2.6	1.2	No	1.8	NW
2/7/20 11:31 AM	128	11	White	Grey	No	2.7	3	No	3	NE
2/7/20 11:34 AM	129	11	White	Grey	No	2.7	N/A	No	3	NE
2/7/20 11:43 AM	130	11	White	Grey	No	2.7	N/A	No	N/A	N/A
2/10/20 12:10 PM	131	12	White + wood	N/A	Yes	2.3	0.6	Yes	0.3	N
2/10/20 12:24 PM	132	12	Wood	N/A	No	2.5	N/A	No	0.3	NW
2/10/20 12:35 PM	133	12	Wood	N/A	No	2.5	N/A	No	0.4	SE
2/10/20 1:24 PM	134	12	white	N/A	No	2.25	N/A	N/A	N/A	N/A
2/11/20 10:19 AM	135	12	white	N/A	No	3.12	N/A	No	2	SE
2/11/20 10:35 AM	136	12	white	N/A	No	3	N/A	No	1	SE
2/11/20 11:17 AM	137	12	white	N/A	No	3.12	N/A	N/A	N/A	N/A
2/12/20 2:01 PM	138	12	white	N/A	No	2.3	N/A	N/A	N/A	N/A
2/12/20 2:20 PM	139	12	white	wood	No	2.3	1	N/A	0.8	W
2/12/20 2:26 PM	140	12	white	wood	No	2.4	N/A	No	3.1	E
2/12/20 2:41 PM	141	12	Wood	N/A	No	2.4	1.4	No	0.8	W

Table A.19: UBC field dataset- Part 5-III

Date-Time	WSID	building	DIST WIND2	WIND DIR2	DESK Width	DESK Len	DESK H	PART H
7/15/2019 10:39	1	10	N/A	N/A	1.5	0.8	0.8	N/A
7/15/2019 10:59	2	10	N/A	N/A	1.5	0.8	0.8	N/A
7/17/2019 10:03	3	10	4.5	NW	1.51	0.75	0.84	1.26
7/17/2019 10:14	4	10	8.08	NW	1.52	0.61	0.84	1.26
7/17/2019 10:28	5	10	6.78	NW	1.52	0.61	0.84	1.26
7/17/2019 11:08	6	10	N/A	N/A	1.82	0.6	0.85	1.36
7/17/2019 11:17	7	10	N/A	N/A	1.8	0.76	0.76	1.26
7/17/2019 11:45	8	10	N/A	N/A	1.52	0.76	0.77	1.47
7/17/2019 11:54	9	10	N/A	N/A	1.52	0.79	0.74	N/A
7/17/2019 15:54	10	10	N/A	N/A	2	0.85	0.86	N/A
7/17/2019 16:02	11	10	N/A	N/A	0.95	0.86	0.85	1.36
7/19/2019 10:19	12	10	N/A	N/A	0.94	0.76	0.74	1.26
7/19/2019 10:09	13	10	N/A	N/A	1.25	0.76	0.74	1.26
7/19/2019 10:38	14	10	6.55	SW	1.52	0.75	0.75	N/A
7/19/2019 11:02	15	10	8.775	SE	1.17	0.6	0.86	N/A
7/19/2019 11:22	16	10	6.6	SE	2	0.49	0.72	N/A
7/19/2019 11:49	17	10	6.4	N	2.82	0.6	0.75	1.25
7/19/2019 12:01	18	10	8.6	SE	3.43	0.75	0.84	N/A
7/19/2019 12:30	19	10	5.9	SE	2.8	0.6	0.74	1.25
7/19/2019 12:56	20	10	6.03	N	1.16	0.6	0.74	1.08
7/19/2019 13:18	21	10	7.2	NW	2.7	0.61	0.74	1.26
7/19/2019 13:52	22	10	N/A	N/A	2.24	0.75	0.83	N/A
7/19/2019 14:10	23	10	N/A	N/A	1.52	0.75	0.73	1.25
7/19/2019 14:29	24	10	5.9	N/A	1.52	0.75	0.75	N/A
7/19/2019 14:49	25	10	N/A	N/A	1.52	0.76	0.86	N/A
7/17/2019 10:22	26	10	8.08	NW	1.52	0.61	0.84	1.26
7/17/2019 10:41	27	10	N/A	N/A	1.52	0.77	0.74	N/A
7/15/2019 12:09	28	11	N/A	N/A	1.2	0.6	0.74	1.7
7/15/2019 13:59	29	11	N/A	N/A	2.4	0.61	0.7	1.04
7/16/2019 11:38	30	11	N/A	NW	1.8	0.59	0.84	1.8
7/16/2019 12:11	31	11	N/A	N/A	1.52	0.76	0.75	N/A
7/16/2019 12:22	32	11	N/A	N/A	1.53	0.76	0.74	1.8
7/16/2019 12:33	33	11	N/A	N/A	1.53	0.76	1.04	1.8
7/16/2019 12:54	34	11	N/A	N/A	1.5	0.66	0.78	N/A
7/16/2019 13:24	35	11	N/A	N/A	1.53	0.6	0.7	N/A
7/16/2019 14:08	36	11	N/A	N/A	1.52	0.6	0.67	1.8
7/18/2019 10:40	37	11	N/A	N/A	3.5	0.75	0.74	1.77
7/18/2019 11:00	38	11	N/A	N/A	3.5	0.75	0.74	1.77
7/18/2019 11:18	39	11	N/A	N/A	2.75	0.76	0.75	N/A
7/18/2019 14:01	40	11	N/A	N/A	2.6	0.6	0.73	N/A
7/18/2019 14:26	41	11	N/A	N/A	1.6	0.76	0.7	N/A
7/18/2019 15:01	42	11	N/A	N/A	1.52	0.76	0.8	N/A
7/18/2019 16:04	43	11	N/A	N/A	1.94	0.76	0.72	N/A

Table A.20: UBC field dataset- Part 6-I

Date-Time	WSID	building	DIST WIND2	WIND DIR2	DESK Width	DESK Len	DESK H	PART H
7/16/2019 13:52	44	11	N/A	N/A	1.52	0.77	0.8	N/A
7/15/2019 15:28	45	12	N/A	N/A	1.6	0.8	0.71	N/A
7/15/2019 15:41	46	12	N/A	N/A	2.98	0.6	0.823	N/A
7/15/2019 12:22	47	12	N/A	N/A	1.53	0.76	0.84	1.37
7/16/2019 9:51	48	12	2.4	NW	1.6	0.8	0.84	N/A
7/16/2019 10:03	49	12	1.35	NW	1.6	0.8	0.83	N/A
7/16/2019 10:10	50	12	3	SW	1.5	0.76	0.83	N/A
7/16/2019 10:31	51	12	N/A	N/A	1.58	0.6	0.84	N/A
7/15/2019 15:18	52	12	N/A	N/A	1	0.6	1	N/A
7/15/2019 14:51	53	13	N/A	N/A	1.2	0.6	0.47	N/A
8/13/2019 9:46	54	14	N/A	N/A	1.67	0.73	0.97	N/A
8/13/2019 9:54	55	14	N/A	N/A	1.8	0.59	0.83	1.43
8/13/2019 10:09	56	14	N/A	N/A	1.8	0.59	0.83	N/A
8/13/2019 10:22	57	14	N/A	N/A	1.8	0.59	0.83	N/A
8/13/2019 10:37	58	14	N/A	N/A	1.8	0.59	0.83	1.43
8/13/2019 10:53	59	14	N/A	N/A	1.7	0.6	0.73	N/A
8/13/2019 11:14	60	14	N/A	N/A	1.8	0.59	0.83	1.43
8/13/2019 11:20	61	14	N/A	N/A	1.8	0.59	0.83	1.43
8/13/2019 11:24	62	14	N/A	N/A	1.8	0.59	0.83	1.43
8/13/2019 11:30	63	14	N/A	N/A	1.8	0.59	0.83	1.43
8/13/2019 11:34	64	14	N/A	N/A	1.8	0.59	0.83	1.43
8/13/2019 11:43	65	14	N/A	N/A	1.8	0.59	0.83	1.43
8/13/2019 11:52	66	14	N/A	N/A	1.8	0.59	0.83	1.43
8/13/2019 12:03	67	14	N/A	N/A	1.8	0.59	0.83	1.43
8/15/2019 10:47	68	14	2	N	1.83	0.58	0.73	1.16
8/15/2019 10:28	69	14	4.5	S	1.83	0.58	0.73	1.16
8/15/2019 11:04	70	14	N/A	N/A	1.54	0.58	0.74	N/A
8/15/2019 11:15	71	14	N/A	N/A	1.8	0.59	0.83	1.34
8/15/2019 11:26	72	14	N/A	N/A	1.8	0.59	0.83	1.34
8/15/2019 11:35	73	14	N/A	N/A	1.8	0.59	0.83	1.34
8/15/2019 11:39	74	14	N/A	N/A	1.8	0.59	0.83	1.34
8/15/2019 11:54	75	14	N/A	N/A	1.8	0.59	0.83	1.34
8/15/2019 11:58	76	14	N/A	N/A	1.8	0.59	0.83	1.34
8/15/2019 13:14	77	14	N/A	N/A	1.78	0.73	0.69	N/A
8/15/2019 13:30	78	14	N/A	N/A	1.52	0.6	0.73	N/A
8/15/2019 13:50	79	14	N/A	N/A	1.8	0.59	0.83	1.34
8/15/2019 13:58	80	14	N/A	N/A	1.8	0.59	0.83	N/A
8/15/2019 14:05	81	14	N/A	N/A	1.8	0.59	0.83	1.34
8/15/2019 14:12	82	14	N/A	N/A	1.8	0.59	0.83	1.34

Table A.21: UBC field dataset- Part 6-II

Date-Time	WSID	building	DIST WIND2	WIND DIR2	DESK Width	DESK Len	DESK H	PART H
8/15/2019 14:18	83	14	N/A	N/A	1.8	0.59	0.83	1.34
8/15/2019 14:26	84	14	N/A	N/A	1.8	0.59	0.83	1.34
8/15/2019 15:01	85	14	N/A	N/A	1.8	0.59	0.83	1.34
8/15/2019 14:49	86	14	N/A	N/A	1.8	0.59	0.83	1.34
8/15/2019 14:54	87	14	N/A	N/A	1.78	0.6	0.72	N/A
8/15/2019 15:07	88	14	N/A	N/A	1.8	0.59	0.83	1.34
8/15/2019 15:19	89	14	N/A	N/A	1.8	0.59	0.83	1.34
8/15/2019 15:26	90	14	N/A	N/A	1.67	0.73	0.72	N/A
8/15/2019 15:43	91	14	N/A	N/A	1.78	0.73	0.74	N/A
8/15/2019 15:47	92	14	N/A	N/A	1.8	0.58	0.75	N/A
8/15/2019 15:58	93	14	N/A	N/A	1.82	0.58	0.73	N/A
8/15/2019 16:04	94	14	N/A	N/A	1.84	0.6	0.74	N/A
1/27/2020 10:35	95	10	n/a	n/a	2.235	0.75	0.72	N/A
1/27/2020 10:58	96	10	N/A	N/A	2.12	0.75	0.75	N/A
1/27/2020 11:43	97	10	N/A	N/A	1.83	0.75	0.74	1.25
1/27/2020 12:15	98	10	N/A	N/A	1.75	0.75	0.75	1.25
1/27/2020 12:31	99	10	N/A	N/A	1.52	0.6	0.74	N/A
1/27/2020 12:36	100	10	N/A	N/A	1.52	0.6	0.71	N/A
1/27/2020 12:53	101	10	N/A	N/A	1.52	0.8	0.74	N/A
1/27/2020 14:47	102	10	N/A	N/A	1.5	0.75	0.78	1.25
1/27/2020 15:07	103	10	N/A	N/A	2.8	0.75	0.77	N/A
1/27/2020 15:29	104	10	1.43	E	1.54	0.75	1.13	N/A
1/27/2020 15:41	105	10	N/A	N/A	1.5	0.76	0.74	N/A
1/27/2020 15:53	106	10	N/A	N/A	1.52	0.75	0.83	N/A
1/29/2020 13:48	107	10	N/A	N/A	2.4	0.75	0.75	N/A
1/29/2020 14:19	108	10	N/A	N/A	2.21	0.75	0.74	1.25
1/29/2020 14:36	109	10	N/A	N/A	385	0.75	0.75	N/A
1/29/2020 14:56	110	10	N/A	N/A	1.5	0.75	0.75	1.25
1/31/2020 10:44	111	10	N/A	N/A	1.47	0.75	0.77	1.26
1/31/2020 11:05	112	10	1	SW	1.52	0.75	0.74	N/A
1/31/2020 11:22	113	10	N/A	N/A	1.52	0.75	0.76	1.7
1/31/2020 11:40	114	10	6.1	NW	1.9	0.92	0.74	N/A
1/31/2020 11:36	115	10	6.1	NW	1.9	0.92	0.74	N/A

Table A.22: UBC field dataset- Part 6-III

Date-Time	WSID	building	DIST WIND2	WIND DIR2	DESK Width	DESK Len	DESK H	PART H
1/31/2020 13:32	116	10	N/A	N/A	2.5	0.6	0.73	1.25
1/31/2020 13:38	117	10	N/A	N/A	1.84	0.76	0.85	1.25
1/31/2020 14:07	118	10	N/A	N/A	1.52	0.6	0.75	1.25
1/31/2020 14:43	119	10	N/A	N/A	1.52	0.6	0.75	1.25
1/31/2020 14:43	120	10	8.6	W	1.9	0.92	0.74	N/A
2/3/2020 10:50	121	11	N/A	N/A	1.52	0.61	0.77	N/A
2/3/2020 11:10	122	11	N/A	N/A	1.52	0.6	0.66	1.68
2/3/2020 11:14	123	11	N/A	N/A	1.52	0.76	1.1	N/A
2/3/2020 11:23	124	11	N/A	N/A	1.52	0.7	0.7	N/A
2/7/2020 10:50	125	11	N/A	N/A	1.52	0.76	0.8	N/A
2/7/2020 11:01	126	11	N/A	N/A	1.52	0.76	0.72	N/A
2/7/2020 11:13	127	11	N/A	N/A	1.52	0.76	0.67	1.68
2/7/2020 11:31	128	11	N/A	N/A	2.88	0.75	0.75	1.64
2/7/2020 11:34	129	11	N/A	N/A	2.88	0.75	0.75	1.64
2/7/2020 11:43	130	11	N/A	N/A	2.88	0.75	0.75	1.64
2/10/2020 12:10	131	12	0.7	W	0.8	1.53	0.76	N/A
2/10/2020 12:24	132	12	N/A	N/A	2.6	0.75	0.74	N/A
2/10/2020 12:35	133	12	N/A	N/A	1.84	0.92	0.75	N/A
2/10/2020 13:24	134	12	N/A	N/A	1.8	0.67	0.76	N/A
2/11/2020 10:19	135	12	N/A	N/A	1.6	0.8	1.3	N/A
2/11/2020 10:35	136	12	N/A	N/A	1.6	0.8	0.72	N/A
2/11/2020 11:17	137	12	N/A	N/A	1.82	0.68	0.73	N/A
2/12/2020 14:01	138	12	N/A	N/A	2.18	0.59	0.74	N/A
2/12/2020 14:20	139	12	N/A	N/A	1.52	0.76	0.74	1.37
2/12/2020 14:26	140	12	4.6	S	2.1	0.76	0.79	1.07
2/12/2020 14:41	141	12	N/A	N/A	2.4	0.6	0.74	N/A

Table A.23: UBC field dataset- Part 6-IV

Appendix B

Sample of the Bayesian Model Code[1]

[1] A version of this Appendix has been published: Extending the Fanger PMV model to include the effect of non-thermal conditions on thermal comfort", In Proceedings of eSIM 2020: Building Simulation meets a global pandemic, Vancouver, Canada.

Extending the Fanger PMV model to include the effect of non-thermal conditions on thermal comfort

Build Run-Time Environment

[2]:
```
#create a clone (a local copy) of the repository containing the dataset on Github
%cd /content
!git clone https://CrosbySarah/eta-lab/COPE-PLUS_UBC_winter_and_summer-data.git

#update all origin branches on Git and merge them
%cd COPE-PLUS_UBC_winter_and_summer-data

!git pull
```

[ ]:
```
#Install Python libraries and packages
!pip install arviz==0.6.1
!pip install pymc3==3.8
!pip install Theano==1.0.4
```

[ ]:
```
#Import Python libraries and packages
import pickle
import matplotlib.pyplot as plt
import pandas as pd
import seaborn
import numpy as np
```

```
import pymc3 as pm
import warnings
from matplotlib.patches import Patch
from matplotlib.lines import Line2D
warnings.filterwarnings('ignore')
pd.set_option('display.float_format', lambda x: '%.3f' % x)
from collections import OrderedDict
from sklearn import preprocessing
```

1 Introduction

Occupant's thermal comfort has become one of the most important aspects of the built environment nowadays. The potential implications of the occupant's thermal dissatisfaction and its effect on cognitive performance as well as overall satisfaction have been the focus of recent research studies (Ferreira et al., 2012; Visher, 2008; Jamrozik et al., 2017; Jensen et al., 2009; Int-Hout, 2013). The American Society for Heating and Refrigeration Engineers (ASHRAE) defines thermal comfort as "the condition of the mind in which satisfaction is expressed with the thermal environment"(ASHRAE Standard 55, 2013). The judgment of thermal comfort is, therefore, a cognitive process influenced by occupant's well-being and overall satisfaction (Djongyang et al., 2010).

In recent years, it has been increasingly observed that occupant's thermal comfort can be influenced by differences in mood, culture and other parameters of well-being Jamrozik et al., 2017; Jokl et al, 2006; and Djongyang et al., 2010). Until now, in current building codes and standards, both physical and psychological indicators of indoor environmental quality are not yet taken into consideration in performance criteria for IEQ. However, in 2014, the International WELL Building Institute launched the WELL building standard, a building performance accreditation scheme focusing only on the comfort and wellness of building occupants (The WELL Building Standard, 2014).

While the WELL building standard has been emerging, many studies in the literature that have explored the interdependencies between thermal comfort, IEQ, and building design have occurred in parallel. For example, in prior studies by Al Horr et al. (2017); Huang et al. (2011); Kamaruzzaman et al. (2011); and Rupp et al. (2015), the authors found that acoustic and visual comfort, biophilia, indoor air quality, and office layout can affect overall occupant comfort and satisfaction. Wagner et al.(2007) found that occupants who are generally satisfied with many non-thermal metrics of IEQ are more likely to be satisfied with thermal conditions as well. Jamrozik et al. (2017) studied the effect of six well-being factors on the occupant's satisfaction in a living lab experimental setup. They found that the perceptions of IEQ conditions that were not varied were also affected. This finding suggests that building occupants' perceptions of environmental conditions may be holistic: dissatisfaction with one set of environmental conditions may affect occupants' perception of the whole environment.

Crosby et al. (2019) and (2021) took advantage of the increasing awareness in the research of the interdependencies between perceived thermal comfort and overall IEQ. The authors have recently developed a novel methodology that is able to evaluate the correlations between non-thermal parameters of IEQ, such as indoor air quality and noise levels, and perceived thermal satisfaction of occupants of open-plan offices. They have applied advanced Bayesian inference techniques on a prior field dataset in order to predict the relationships between surveyed occupant thermal dissatisfaction, thermal IEQ parameters and several measurable non-thermal, "well-being"-type indoor environmental conditions. They found that there exist statistically significant independent correlations between thermal comfort, as perceived by occupants of open-plan offices, and some non-thermal IEQ parameters. The most significant finding is that a modest increase in measured indoor CO2 concentrations, from 500 ppm to 900 ppm, is found to be correlated with a decrease in perceived thermal satisfaction by 30%. The results have been tested for statistical significance using three different model comparison approaches.

It is, then, the purpose of this paper to make use of the recently developed method by Crosby et al. (2021) and update their results using a more recent field study database. A large IEQ field study carried out at the University of British Columbia (UBC) across 2019/2020 is presented in this paper. The new IEQ field dataset consists of survey responses from occupants across offices of the UBC campus in Vancouver, Canada, as well as indoor environmental physical measurements from around 150 workstations.

Recently, several studies in the literature have explored the deficiency of standard models of thermal comfort in accurately predicting true occupant thermal comfort, despite being used in international building codes and standards for some time (Humphreys et al., 2002 and Oseland et al., 1994). This study seeks to update the recent findings by Crosby et al. (2021) while addressing the most prevailing research gap with respect to thermal comfort models by proposing an extension to the Fanger PMV model which consists of a more holistic method of thermal comfort evaluation.

The goal of updating the COPE database with the newly collected data in UBC is to validate and test the Bayesian method proposed by Crosby et al. (2021) while updating the 'relatively old' IEQ COPE field dataset with a more recent one. If these results are found to be repeatable under different datasets types and a variety of model input conditions and by using a new set of experimental procedures and methodologies, it will indicate that the previous results and findings are robust and significant.

2 Methodologies

2.1 Bayesian Modelling of Thermal Comfort

Bayesian inference techniques have been applied in many recent studied to improve predictions of thermal comfort using new observational data (Jensen et al., 2008 and Langevin et al. 2013)

One of the greatest advantages of adopting Bayesian methodologies in thermal comfort predictions is the ability to incorporate different datasets into one thermal comfort model while updating prior knowledge on thermal comfort's distributions from past research into the current estimation of model parameters which provides a robust manner of updating these parameters as more data becomes available.

As discussed in the introduction to this paper, a prevailing research gap with respect to standard models of thermal comfort models is that these models have not always accurately predicted true thermal comfort observations found in the field and only consider thermal parameters of IEQ when evaluating occupant's thermal comfort. However, Crosby et al. (2021) have recently proposed a method that seeks to evaluate the correlations between non-thermal IEQ conditions, such as indoor lighting levels, and CO2, on metrics and perceived thermal comfort. The research underlying this paper involves applying this novel Bayesian methodology to a recent field IEQ study while updating and verifying these previous findings. This paper introduces two new indoor environmental parameters that are not included in Crosby et al.: Desktop partition heights, and indoor concentrations of carbon monoxide.

In this paper, a Bayesian logistic regression model is applied and used as an expansion of the Fanger PMV model in a manner that allows for the probability of occupant's thermal dissatisfaction, p(D), or thermal satisfaction, p(S) to be related not only to the original model terms of the Fanger PMV-PPD model, F = {T, R, M, V}, but also to several non-thermal well-being IEQ parameters defined by a separate set of terms, W. The thermal parameters, F, are specifically: T = air temperature (ºC), R =relative humidity (%), M = mean radiant temperature (ºC), and V = air velocity (m/s). The measurable non-thermal well-being IEQ metrics used in this study, defined by the set W = {C, N, L} are: C = indoor air CO2 levels (ppm), N = A-weighted indoor noise levels (dBA), L = Measured desktop illuminance (lux).

The Bayesian network diagram of the model is shown in Figure 1, where posterior predictions of thermal satisfaction p(S | F, W) can be determined as follows:

$$p(D \mid \mathbf{F}, \mathbf{W}) \propto p(\mathbf{W} \mid D) \cdot \; p(\mathbf{F} \mid D) \cdot \; p(\mathbf{D}) \tag{1}$$

Such that:

$$p(D \mid \mathbf{F}, \mathbf{W}) = \frac{p(\mathbf{W} \mid D) \cdot \ p(\mathbf{F} \mid D) \cdot \ p(D)}{p(\mathbf{F}, \mathbf{W})} \tag{2}$$

For the equations above $p(D \mid \mathbf{F})$ is the probabilistic model of thermal dissatisfaction given the thermal and non-thermal well-being IEQ parameters. $p(\mathbf{F} \mid D)$ is the likelihood distribution of observing the well-being IEQ parameters, $\mathbf{W}$ given the observed thermal dissatisfaction, D.

[ ]:
```
#import image from Github

file=('BN1.png')
from IPython.display import Image
Image(file)
```

[ ]:

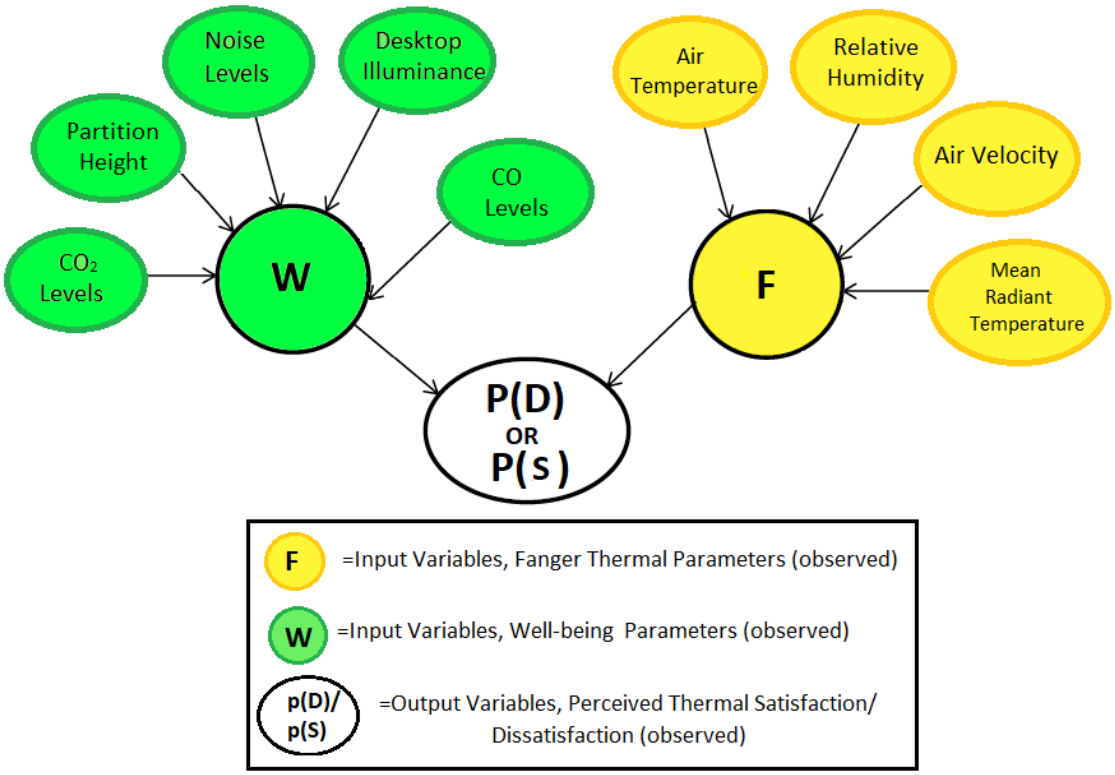

Figure 1: A Bayesian Network of the developed thermal comfort model incorporating both thermal Fanger and non-thermal wellbeing parameters

2.2 IEQ Field Data

2.2.1 Updating the COPE dataset

The Cost-effective Open-Plan Environment or the 'COPE' study, is a field IEQ data collected by the National Research Council Canada (NRC). The COPE dataset consists of IEQ data collected from 779 workstations and their occupants in nine buildings between 2000 and 2002 across Canada and the United States. The IEQ data contains instantaneous physical measurements of IEQ conditions as well as data from satisfaction survey responses. The physical measurements were taken using a cart-and-chair system for each visited

workstation (Newsham et al., 2008). The IEQ parameters covered by the COPE dataset consist of over four physical measurements of thermal conditions, such as temperature, relative humidity, and air velocity, and additionally include 12 measurements of non-thermal conditions, such as noise levels, desktop illuminance, and CO2 concentrations.

Crosby et al.(2021) have recently developed a novel approach for the prediction of thermal comfort, in which they have evaluated the quantifiable correlations between measurable non-thermal "well-being"-type parameters of IEQ and perceived thermal satisfaction. The authors took advantage of the COPE dataset, made available for their study, and incorporated it into the developed Bayesian methodology. A total of 28 different Bayesian models were developed to examine the relationship between thermal satisfaction, thermal dissatisfaction, and different combinations of thermal and non-thermal IEQ parameters drawn from the COPE dataset.

The results revealed that there is evidence to suggest that some of the non-thermal IEQ parameters do have independent correlations with perceived thermal dissatisfaction/satisfaction, at least in the open-plan offices of the COPE dataset. It is also important to note that some of these novel findings have proven to show substantial evidence of a correlation despite the limited range of data. It is also important to note that this is the first known study to suggest that there exists a correlation between perceived thermal satisfaction and non-thermal parameters of IEQ.

In this study, the COPE dataset will be used along with the newly collected data, explained in detail in the following section, to update the 'relatively old' IEQ field data and verify and test the robustness of the previous findings.

2.2.2 2.2.2. ESTEBAN-UBC IEQ data collection

2.2.2.1. Experimental Setup Such observable correlations have revealed the need for developing an updated version of the data collected. A large IEQ field study has been conducted on the University of British Columbia campus in Vancouver, Canada. The field study has been designed to examine the potential correlations between the occupant's perceived thermal satisfaction/dissatisfaction and many thermal and non-thermal parameters of IEQ. The database consists of instantaneous physical measurement of IEQ, and spatial and manual measurements coupled with responses from an IEQ questionnaire collected from 150 workstations in five buildings between 2019 and 2020 from open-plan offices. Each building is visited twice, in summer (between July and August 2019) and wintertime (between January and February 2020). Participants are mostly University staff and faculty members from different schools and disciplines. The 14 buildings used in this study are mainly open-plan offices, except for a few semi-private offices (2-5 desks). All potential participants were contacted by the research team beforehand and provided with an information package and asked to sign a consent form. Upon consent, each participant is scheduled to participate in the experiment on a particular date and time. The procedure was approved by UBC's Research ethics board.

2.2.2.2. IEQ Satisfaction Questionnaire A satisfaction survey questionnaire is developed to collect data from occupants asking them about their "right-here-right-now" satisfaction with the thermal environment as well as many environmental features in their workspace and their overall environmental satisfaction. Occupants were ensured to have been at a sedentary position at their workstation at least one hour before data collection to make sure each participant was at a certain level of equilibrium with the surrounding environment.

Participants were asked to fill out an anonymized questionnaire based on how they are feeling 'right now'. Participants completed a 41-questions survey at their workstation while the physical IEQ measurements are collected. Questions about IEQ satisfaction included 3 different perception categories for each IEQ parameter: perceived levels (How would you rate the temperature at your workspace right now?), satisfaction with perceived levels (How would you rate your satisfaction with the air temperature at your workspace right now?), and preference to perceived levels (How would rate your preference with respect to the air temperature at your workspace right now?). These questions were answered on a 7-point Likert scale (1

= Extremely dissatisfied and 7 = Extremely satisfied). The survey covered satisfaction with the thermal environment, job and workplace satisfaction, long term satisfaction with the air temperature, satisfaction with background noise levels, lighting levels, daylight availability, glare, view to the outside, quality and quantity of artificial lighting, air quality, air movement, humidity, IEQ controllability, and other individual features of the workspace.

2.2.2.3. IEQ physical measurement A high precision IEQ sensors cart, ESTEBAN (Exceptional Sensing Testbed for Environment, Biophilia, Air-quality, and Nippiness), is designed and built for this study. The cart carries all the IEQ sensors required for the local microclimate measurement conducted at each workstation. For each participant, the occupant's office chair is removed and replaced with the measurement cart. ESTEBAN, holds a variety of sensors and measurement tools to record indoor air tem-perature, relative humidity, mean radiant temperature, CO2 concentrations, CO concentrations, TVOCs concentrations, air velocity, A-weighted noise levels, horizontal and vertical illuminance levels, and NO concentrations.

[ ]:

Manual and spatial measurements included participant clothing levels, workspace size, raining conditions, blinds and windows presence and operability, desktop partition height, distance to the nearest indoor plant, distance to the nearest windows, presence of a desk lamp, and 360° images of workspaces for future processing of percentage indoor greenery. Materials, types, and colours of office chair, walls, desktop partition and ceiling, primary and secondary windows orientation, and ceiling height are also recorded. The IEQ physical measurement and manual data collection take a total of about 10 minutes.

Preparing the training data

[ ]:
```
# loading dataset file for COPE  and UBC data

file = ('COPE_and_UBC-Summer+winter-Data_II.xlsx')
xl = pd.ExcelFile(file)
pulled_data = xl.parse('alldata')
data=pulled_data
```

[ ]:
```
#Detect missing (null) values within the dataset

data = data[~pd.isnull(data['SAT_TEMP'])]
data = data[~pd.isnull(data['Temp'])]
data = data[~pd.isnull(data['RH'])]
data = data[~pd.isnull(data['MRT'])]
data = data[~pd.isnull(data['V'])]
data = data[~pd.isnull(data['CO2'])]
data = data[~pd.isnull(data['Lux'])]
data = data[~pd.isnull(data['Noise'])]
data = data[~pd.isnull(data['CO'])]
data = data[~pd.isnull(data['PART'])]
```

[ ]:

Estaiblishing a criteria variable for simulation

```
#The variable `criteria` requires user input.
#CLASS is p(D) or p(S) depending on the criteria input
#Enter 'pD' for probability of thermal dissatisfaction
#Enter 'pS' for probability of thermal satisfaction
#pS is thermal satisfaction votes (Sat_temp>4)
#pD is thermal didsatisfaction votes (Sat_temp<4)
criteria = "pS"
if criteria == "pS":
  data['CLASS'] = 1*(data['SAT_TEMP'] > 4)
else:
  data['CLASS'] = 1*(data['SAT_TEMP'] < 4)
```

2.3 Bayesian Models Descriptions

A Bayesian logistic regression model is developed to represent eq.2 against observations found in the field IEQ datasets. The model predicts the linear relationship between thermal dissatisfaction (D), Fanger's thermal conditions ($\mathbf{F}$), and non-thermal well-being IEQ parameters ($\mathbf{W}$) drawn from the dataset. By considering as a set of regression model coefficients, $p(D|F,W)$, is estimated as:

$$p(D \mid \mathbf{F},\ \mathbf{W}) = \frac{1}{1 + e^{-}[\Sigma_{\mathbf{F}}\ (\beta_{\mathbf{F}}\ .\ \mathbf{F}) + \Sigma_{\mathbf{W}}\ (\beta_{\mathbf{W}}\ .\mathbf{W}) +\ \beta\ ]} \tag{3}$$

Similarly, the developed Bayesian logistic regression model is applied to observed thermal satisfaction data to evaluate the correlation between thermal satisfaction $p(S)$, Fanger thermal conditions ($\mathbf{F}$) and non-thermal wellbeing IEQ parameters ($\mathbf{W}$) drawn from the COPE dataset. The posterior probability distribution of thermal satisfaction given thermal and non-thermal parameters can be written as follows:

$$p(S \mid \mathbf{F},\ \mathbf{W}) = \frac{1}{1 + e^{-}[\Sigma_{\mathbf{F}}\ (\beta_{\mathbf{F}}\ .\ \mathbf{F}) + \Sigma_{\mathbf{W}}\ (\beta_{\mathbf{W}}\ .\mathbf{W}) +\ \beta\ ]} \tag{4}$$

Physical measurements of thermal and non-thermal parameters of IEQ are drawn from the UBC and COPE databases. Data are then plugged into the Bayesian model in order to infer the posterior probability distri-butions of occupant's thermal satisfaction given thermal and non-thermal parameters of IEQ.

The following non-thermal 'wellbeing-type' measurements are used in this study to characterize conditions of air quality, lighting, acoustics, and interior design. These are defined by the set W = {C, N, L, CO, P}. Where C is the measured indoor CO2 levels (ppm), N is the A-weighted indoor noise levels(dBA), L is the measured illuminance at each workstation (Lux), CO is the measured CO concentrations(ppm), and P is the desktop partition height (m). The thermal IEQ conditions drawn from the field database, defined by the set **F**, are **F**= {T, R, M, V}. Where: T = air temperature (ºC), R =relative humidity (%), M = mean radiant temperature (ºC), and V = air velocity (m/s).

Occupant survey responses to the perceived thermal satisfaction question ("How satisfied are you with the temperature at your workspace right now?"), TS, are drawn from the COPE and UBC datasets. The probability of thermal dissatisfaction, $p(D)$, is modelled as a dichotomous dependent and parameters of F and W sets are modelled as continuous independent variables. Observed data for $p(D)$ is inferred from the COPE and the UBC datasets by assuming that for each survey response, i:

$$D_i = \begin{cases} 1 \text{ (or dissatisfied) } if\ TS < 4 \\ 0 \text{ (or satisfied) } if\ TS > 4 \end{cases} \tag{5}$$

Similarly, the probability of satisfaction, p(S), is modelled as a dichotomous dependent variable and observed data for p(S) is inferred from the COPE and UBC datasets by assuming that for each survey response, i:

$$S_i = 1 - D_i \tag{6}$$

Neutral responses made by occupants with regards to thermal satisfaction (i.e., TS = 4) are ultimately excluded as they are neither clear indicators of thermal satisfaction nor dissatisfaction. Distributions of thermal satisfaction responses received, TS, for each of the COPE and UBC databases are shown in Figure 3.

Calculated probabilities for p(D) and p(S) for both the COPE and UBC dataset are shown in Figure 4. Calculated probabilities for p(D) and p(S) for both the COPE and UBC dataset are shown in Figure 4.

[ ]:
```
#Distribution of thermal satisfaction across the COPE and UBC datasetsplt.
 ↪figure(figsize=(9,6))
plt.figure(figsize=(9,6))
bins=[0,0.5,1.5,2.5,3.5,4.5,5.6,6.5,7.5]
seaborn.distplot( data[data['Dataset'] == 'COPE']["SAT_TEMP"],bins=bins,
 ↪norm_hist=True,color='orange'␣
 ↪,hist=True,label='COPE',hist_kws=dict(rwidth=1,align="mid",edgecolor="orange",␣
 ↪linewidth=1,alpha=0.7))
seaborn.distplot( data[data['Dataset'] == 'UBC']["SAT_TEMP"],bins=bins,
 ↪norm_hist=True,color='cornflowerblue'␣
 ↪,hist=True,label='UBC',hist_kws=dict(rwidth=1,align="mid",edgecolor="cornflowerblue",␣
 ↪linewidth=1,alpha=0.6))
plt.xlabel('Thermal Satisfaction Responses, TS', fontsize=12)
plt.ylabel('Probability Density Function', fontsize=12)
plt.legend()
plt.xticks(ticks=[1,2,3,4,5,6,7])
plt.show()
```

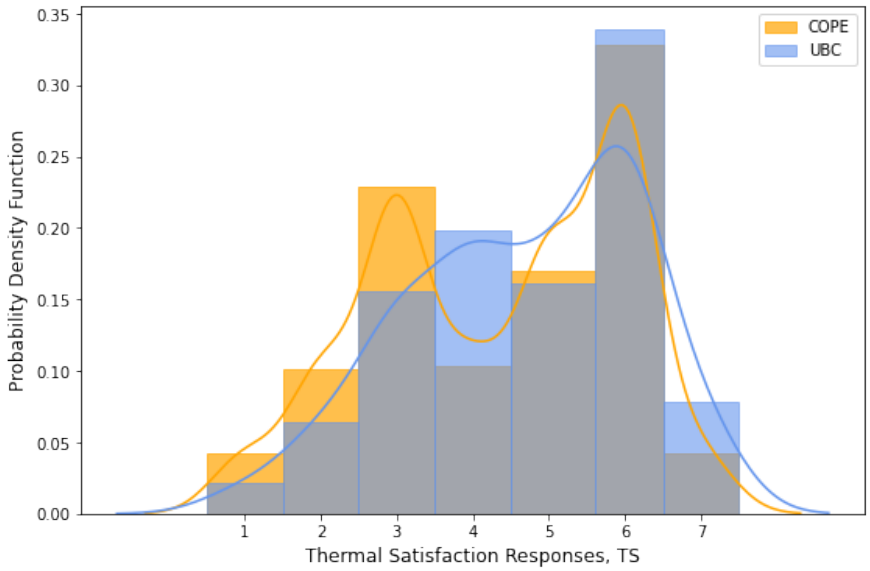

Figure 2: Distributions of thermal satisfaction responses received, TS, for COPE and UBC databases

[ ]:
```
#calculate p(D) and p(S) for COPE, UBC, and COPE+UBC datasets
UBC_p_S_Count=data[data['Dataset'] == 'UBC'].loc[data.SAT_TEMP>4,'SAT_TEMP'].count()
UBC_p_D_Count=data[data['Dataset'] == 'UBC'].loc[data.SAT_TEMP<4,'SAT_TEMP'].count()
COPE_p_S_Count=data[data['Dataset'] == 'COPE'].loc[data.SAT_TEMP>4,'SAT_TEMP'].count()
COPE_p_D_Count=data[data['Dataset'] == 'COPE'].loc[data.SAT_TEMP<4,'SAT_TEMP'].count()
UBC_data_count=data.loc[data.Dataset=='UBC','Dataset'].count()
COPE_data_count=data.loc[data.Dataset=='COPE','Dataset'].count()
p_S_count=data.loc[data.SAT_TEMP>4,'SAT_TEMP'].count()
p_D_count=data.loc[data.SAT_TEMP<4,'SAT_TEMP'].count()
prob_pS_UBC= UBC_p_S_Count/UBC_data_count
prob_pD_UBC=UBC_p_D_Count/UBC_data_count
prob_pS_COPE=COPE_p_S_Count/COPE_data_count
prob_pD_COPE=COPE_p_D_Count/COPE_data_count
prob_pS_all=p_S_count/data.shape[0]
prob_pD_all=p_D_count/data.shape[0]
```

[ ]:
```
p_S_ = (prob_pS_COPE,prob_pS_UBC, prob_pS_all)
p_D_ = (prob_pD_COPE,prob_pD_UBC, prob_pD_all)

fig, ax = plt.subplots(figsize=(8,4.5))
index = np.arange(3)
bar_width = 0.35
opacity = 0.9

rects1 = plt.bar(index, p_D_, bar_width,
alpha=0.7,
color='#E04D2D',
label='p(D)')

rects2 = plt.bar(index + bar_width, p_S_, bar_width,
```

```
alpha=0.7,
color='#23946A',
label='p(S)')

plt.ylabel('p(D) / p(S)')
plt.title('Probability of Thermal Satisfaction/Dissatisfaction across COPE and UBC␣
 ↪Offices')
plt.xticks(index+(bar_width/2) , ('COPE', 'UBC', 'COPE+UBC'))
plt.legend()
plt.tight_layout()
plt.show()
```

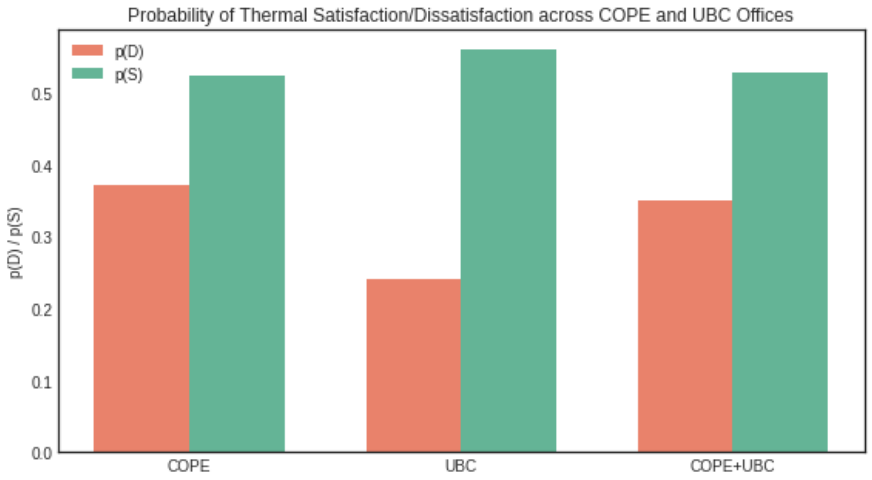

Figure 3: Probability of thermal satisfaction/dissatisfaction across COPE and UBC offices

Probability distributions derived from both the COPE and UBC datasets for both **F** and,**W** IEQ datasets are shown in Figure 5 and Figure 6 respectively.

[ ]:
```
Dataset = ['COPE', 'UBC']
colours = ['orange', 'cornflowerblue']
plt.subplots(figsize=(14,7))
plt.subplot(231)
seaborn.distplot( data[data['Dataset'] == 'UBC']["Noise"], color='cornflowerblue'␣
 ↪,hist=False,label='UBC')
seaborn.distplot( data[data['Dataset'] == 'COPE']["Noise"], color='orange'␣
 ↪,hist=False,label='COPE')
plt.xlabel("Noise Levels [dBA]")
plt.ylabel("Probabilitiy Density Function")
plt.legend()
plt.tight_layout()

plt.subplot(232)
seaborn.distplot( data[data['Dataset'] == 'UBC']["CO2"], color='cornflowerblue'␣
 ↪,hist=False,label='UBC')
seaborn.distplot( data[data['Dataset'] == 'COPE']["CO2"], color='orange'␣
 ↪,hist=False,label='COPE')
```

```
plt.ylabel("Probabilitiy Density Function")
plt.xlabel("CO2 Concentrations [ppm]")
plt.legend()
plt.tight_layout()
plt.subplot(233)
seaborn.distplot( data[data['Dataset'] == 'UBC']["Lux"], color='cornflowerblue',␣
 ↪,hist=False,label='UBC')
seaborn.distplot( data[data['Dataset'] == 'COPE']["Lux"], color='orange',␣
 ↪,hist=False,label='COPE')
plt.xlabel("Desktop Illuminance [Lux]")
plt.legend()
plt.ylabel("Probabilitiy Density Function")
plt.tight_layout()
plt.subplot(234)
seaborn.distplot( data[data['Dataset'] == 'UBC']["PART"], color='cornflowerblue',␣
 ↪,hist=False,label='UBC')
seaborn.distplot( data[data['Dataset'] == 'COPE']["PART"], color='orange',␣
 ↪,hist=False,label='COPE')
plt.ylabel("Probabilitiy Density Function")
plt.xlabel("Partition Height [m]")
plt.legend()
plt.tight_layout()
plt.subplot(235)
seaborn.distplot( data[data['Dataset'] == 'UBC']["CO"], color='cornflowerblue',␣
 ↪,hist=False,label='UBC')
seaborn.distplot( data[data['Dataset'] == 'COPE']["CO"], color='orange',␣
 ↪,hist=False,label='COPE')
plt.xlabel("CO Concentrations [ppm]")
plt.ylabel("Probabilitiy Density Function")
plt.legend()
```

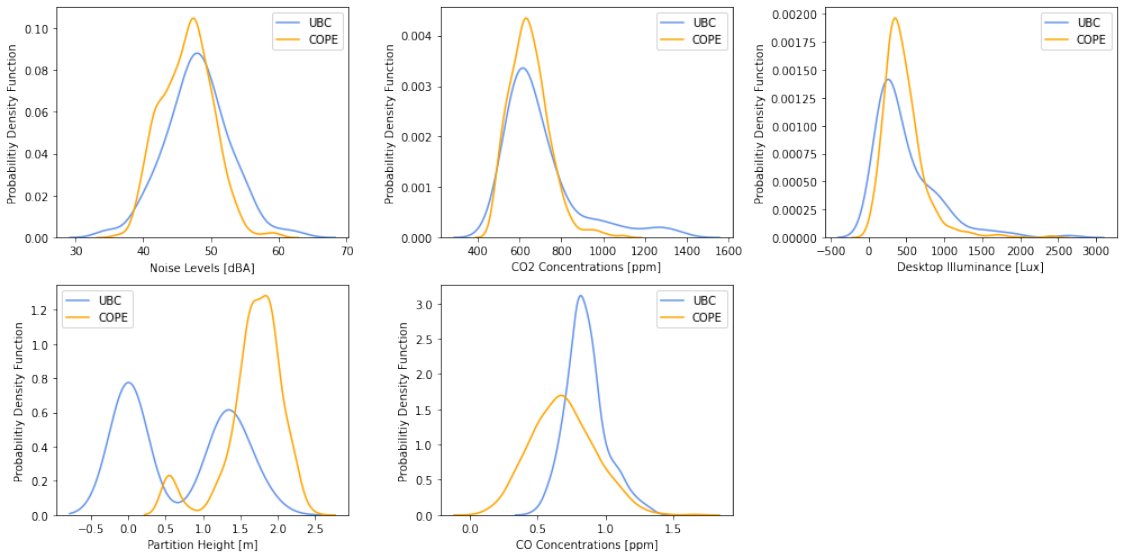

Figure 4: Probability distributions of non-thermal IEQ parameters across COPE and UBC offices

[ ]:
```
Dataset = ['COPE', 'UBC']
colours = ['orange', 'cornflowerblue']
plt.subplots(figsize=(10,7))
plt.subplot(221)
seaborn.distplot( data[data['Dataset'] == 'UBC']["Temp"], color='cornflowerblue'␣
 ↪,hist=False,label='UBC')
seaborn.distplot( data[data['Dataset'] == 'COPE']["Temp"], color='orange'␣
 ↪,hist=False,label='COPE')
plt.xlabel("Indoor Air Temperature [C]")
plt.ylabel("Probabilitiy Density Function")
plt.legend()
plt.tight_layout()

plt.subplot(222)
seaborn.distplot( data[data['Dataset'] == 'UBC']["RH"], color='cornflowerblue'␣
 ↪,hist=False,label='UBC')
seaborn.distplot( data[data['Dataset'] == 'COPE']["RH"], color='orange'␣
 ↪,hist=False,label='COPE')
plt.ylabel("Probabilitiy Density Function")
plt.xlabel("Relative Humidity [%]")
plt.legend()
plt.tight_layout()

plt.subplot(223)
seaborn.distplot( data[data['Dataset'] == 'UBC']["V"], color='cornflowerblue'␣
 ↪,hist=False,label='UBC')
seaborn.distplot( data[data['Dataset'] == 'COPE']["V"], color='orange'␣
 ↪,hist=False,label='COPE')
plt.xlabel("Air Velocity [m/s]")
plt.legend()
plt.tight_layout()

plt.subplot(224)
seaborn.distplot( data[data['Dataset'] == 'UBC']["MRT"], color='cornflowerblue'␣
 ↪,hist=False,label='UBC')
seaborn.distplot( data[data['Dataset'] == 'COPE']["MRT"], color='orange'␣
 ↪,hist=False,label='COPE')
plt.xlabel("Mean Radiant Temperature [C]")
plt.ylabel("Probabilitiy Density Function")
plt.legend()
plt.tight_layout()
```

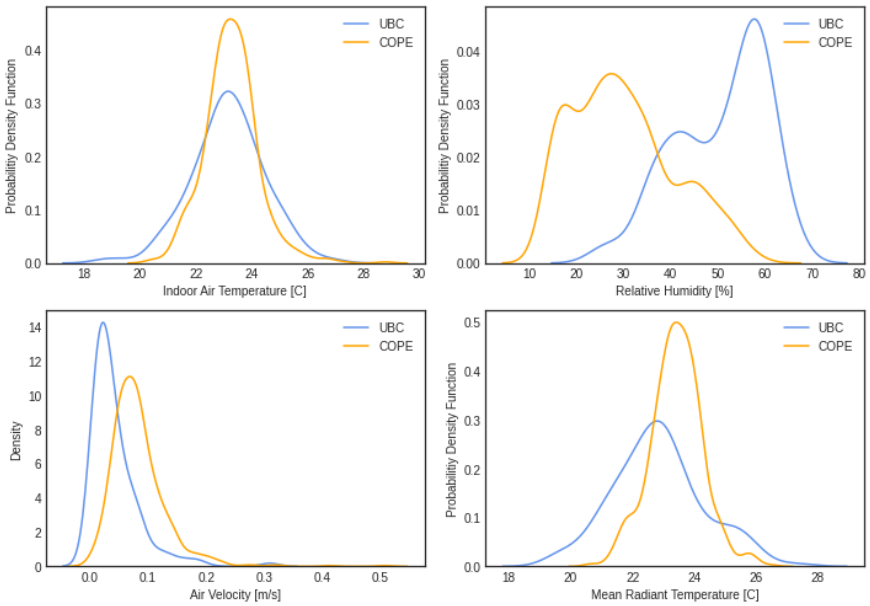

Figure 5: Probability distributions of thermal IEQ parameters across COPE and UBC offices

2.3.1 Description of candidate models

As this paper seeks to determine whether non-thermal parameters can improve the prediction of indoor thermal comfort, there is no single provisional model that is to be explored, but several.Table 1 lays out a total of 15 different cases of thermal and non-thermal variables, **F** and, **W**. Each combination is used to define a logistic regression model that seeks to relate the probability of an occupant feeling thermally satisfied or dissatisfied to parameters indicated. For example, case 3 establishes the probability of an occupant feeling thermally satisfied or dissatisfied as a function of indoor air temperature, mean radiant temperature, air velocity, and relative humidity. The respective long-form notation of the models attributed to this case would be respectively $p(S|T, M, R, V)$ and $p(D|T, M, R, V)$. In all, a total of 30 different models are established out of the 15 cases of Table 1. It should be noted that the list presented in this table already reflects the outcome of some trial and error analysis that was undertaken to determine candidate cases of interest. It is by no means an exhaustive list of all possible models that can be created out of the COPE and UBC datasets.

Normalize data

[ ]:
```
#The goal of this following script is to scale each column of data into a range of 0␣
↪and 1.
#The package function `MinMaxScaler()` holds memory of the scaling functions per␣
↪variable,
#so one can later inverse scaled results to their original values, or find the␣
↪scaled-equivalent of a new test values.
```

Table 1: List of cases evaluated for generating models of p(D) and p(S)

Case	Parameters of F
1	T, M
2	T, M,R
3	T, M,R,V
4	T, M
5	T, M
6	T,M
7	T, M, R,V
8	T,M,R
9	T, M,R,V
10	T, M, R, V
11	-
12	-
13	-
14	-
15	-

[ ]:
```
#array of data to be normalized
data = data[['WSID','building#','CLASS','Temp','PART','MRT', 'RH',␣
↪'CO2','Lux','CO','Noise','V','SAT_TEMP']]
```

[ ]:
```
from sklearn import preprocessing

data_original = data.copy()
data_original['OT']=(data_original['Temp']+data_original['MRT'])/2
data_numpy = data.values # Converts 'data' dataframe to numpy array
min_max_scaler = preprocessing.MinMaxScaler() # Establishes a normalizing function␣
↪that will normalize each column of data to a scale between 0 and 1
data_numpy_scaled = min_max_scaler.fit_transform(data_numpy) # Transform 'data' numpy␣
↪array into normalized values
data_archive = data # Archive original 'data' dataframe
data = pd.DataFrame(data_numpy_scaled,columns=data.columns) # Replace 'data' with␣
↪normalized values
data_inverse=min_max_scaler.inverse_transform(data.values)
```

[ ]:
```
#prepare data

Sat_Temp=data['SAT_TEMP']
building_no=data['building#']

# Thermal parameters
#1st order
Temp= data['Temp']
RH = data['RH']
MRT = data['MRT']
V =data['V']

#2nd order
data['Temp2'] = np.square(data['Temp'])
```

```
Temp2= data['Temp2']
data['RH2'] = np.square(data['RH'])
RH2= data['RH2']
data['MRT2'] = np.square(data['MRT'])
MRT2= data['MRT2']
data['V2'] = np.square(data['V'])
V2= data['V2']

# WELL/non-thermal Parameters
#1st order
CO2=data['CO2']
CO= data['CO']   #Carbon Monoxide
Noise= data['Noise']
PART=data['PART']
Lux = data['Lux']  # Average of 2 desk illuminace measurements
#2nd order
data['CO2_2'] = np.square(data['CO2'])
CO2_2=data['CO2_2']
data['Lux_2'] = np.square(data['Lux'])
```

[ ]:
```
#Finish preparing dataset
classified_data = data['CLASS']#(1 if Sat_Temp<4 and 0 if SAT_Temp>4)
classified_data.value_counts()
classified_data=np.array(classified_data,dtype=bool)
```

Bayesian Models

[ ]:
```
N_models = 18 #rows
N_terms = 15  #columns
key =np.zeros((N_models,N_terms))

#          T MRT RH  V  T^2 MRT^2 RH^2 CO2 CO2^2 CO part SII Noise Lux  VOC
key[0] = [1,1  ,0,  0, 1  ,1    ,0   ,0  ,0    ,0 ,0   ,0  ,0    ,0    ,0] # P(D/S|
 ↪T,MRT,T2,MRT2)
key[1] = [1,1  ,1,  0, 1  ,1    ,1   ,0  ,0    ,0 ,0   ,0  ,0    ,0    ,0] # P(D/S|
 ↪T,MRT,RH,T2,MRT2,RH2)
key[2] = [1,1  ,1,  1, 1  ,1    ,1   ,0  ,0    ,0 ,0   ,0  ,0    ,0    ,0] # P(D/S|
 ↪T,MRT,RH,T2,MRT2,RH2,V)
key[3] = [0,0  ,0,  0, 0  ,0    ,0   ,1  ,0    ,0 ,0   ,0  ,0    ,0    ,0] # P(D/S|
 ↪CO2)
key[4] = [0,0  ,0,  0, 0  ,0    ,0   ,1  ,1   ,0 ,0   ,0  ,0    ,0    ,0] # P(D/S|
 ↪CO2,CO2_2)
key[5] = [1,1  ,0,  0, 1  ,1    ,0   ,1  ,0    ,0 ,0   ,0  ,0    ,0    ,0] # P(D/S|
 ↪T,MRT,T2,MRT2,CO2)
key[6] = [1,1  ,1,  0, 1  ,1    ,1   ,1  ,0    ,0 ,0   ,0  ,0    ,0    ,0] # P(D/S|
 ↪T,MRT,RH,T2,MRT2,RH2,CO2)
key[7] = [1,1  ,1,  1, 1  ,1    ,1   ,1  ,0    ,0 ,0   ,0  ,0    ,0    ,0] # P(D/S|
 ↪T,MRT,RH,T2,MRT2,RH2,V,CO2)
key[8] = [1,1  ,1,  1, 1  ,1    ,1   ,1  ,1    ,0 ,0   ,0  ,0    ,0    ,0] # P(D/S|
 ↪T,MRT,RH,T2,MRT2,RH2,V,CO2,CO2_2)
```

```
key[9] = [1,1  ,1,  1, 1  ,1    ,1   ,1  ,0    ,1 ,1   ,0  ,1    ,1    ,0] # P(D/S|
 ↪T,MRT,RH,T2,MRT2,RH2,V,Lux,CO2,CO,part,N))
key[10]= [0,0  ,0,  0, 0  ,0    ,0   ,1  ,0    ,1 ,1   ,0  ,1    ,1    ,0] # P(D/S|
 ↪CO2,CO,part,Lux,Noise)
key[11]= [1,1  ,1,  1, 1  ,1    ,1   ,1  ,0    ,0 ,0   ,0  ,1    ,1    ,0] # P(D/S|
 ↪T,MRT,RH,T2,MRT2,RH2,V,CO2,Lux,Noise)
key[12]= [0,0  ,0,  0, 0  ,0    ,0   ,1  ,0    ,0 ,0   ,0  ,1    ,1    ,0] # P(D/S|
 ↪CO2,Noise,Lux)
key[13]= [0,0  ,0,  0, 0  ,0    ,0   ,1  ,1    ,0 ,0   ,0  ,1    ,1    ,0] # P(D/S|
 ↪CO2,CO2_2,Noise,Lux)
key[14]= [1,1  ,0,  0, 1  ,1    ,0   ,0  ,0    ,0 ,0   ,0  ,0    ,1    ,0] # P(D/S|
 ↪T,MRT,T2,MRT2,Lux)
key[15]= [1,1  ,0,  0, 1  ,1    ,0   ,0  ,0    ,0 ,0   ,0  ,1    ,0    ,0] # P(D/S|
 ↪T,MRT,T2,MRT2,Noise)
key[16] = [0,0  ,0,  0, 0  ,0    ,0   ,0  ,0    ,1 ,0   ,0  ,0    ,0    ,0] # P(D/S|
 ↪CO)
key[17] = [0,0  ,0,  0, 0  ,0    ,0   ,0  ,0    ,0 ,1   ,0  ,0    ,0    ,0] # P(D/S|
 ↪part)
#key[18] = [0,0  ,0,  0, 0  ,0    ,0   ,0  ,0    ,0 ,0   ,0  ,0    ,0    ,1] # P(D/S|
 ↪VOC)
```

[ ]:
```
#Load models and traces from pickles
#models are pickled in a separate notebook pushed into Git and loaded here
models, traces = OrderedDict(), OrderedDict()
model_names = []
for j in range(N_models):
  model_name=criteria
  for i in range(N_terms):
    if key[j,i]==1:
      if i == 0:
        model_name=model_name+"_T"
      elif i==1:
        model_name=model_name+"_M"
      elif i==2:
        model_name=model_name+"_R"
      elif i==3:
        model_name=model_name+"_V"
      elif i==4:
        model_name=model_name+"_T2"
      elif i==5:
        model_name=model_name+"_M2"
      elif i==6:
        model_name=model_name+"_R2"
      elif i==7:
        model_name=model_name+"_C"
      elif i==8:
        model_name=model_name+"_C2"
      elif i==9:
        model_name=model_name+"_CO"
      elif i==10:
        model_name=model_name+"_P"
      #  elif i==11:
```

```
    #    model_name=model_name+"_S"
    elif i==12:
      model_name=model_name+"_N"
    elif i==13:
      model_name=model_name+"_L"
nm = j
model_names.append(model_name)
model_file='model_'+model_name+'.pkl'
models[nm]=pickle.load( open( model_file, "rb" ) )
trace_file='trace_'+model_name+'.pkl'
traces[nm]=pickle.load( open( trace_file, "rb" ) )
```

2.3.2 Sampling of posterior distributions

The Bayesian statistics Python library, PyMC3, is used to infer posterior distribution of logistic regression model coefficients for all of the models p resented. 5000 samples are drawn from the posteriors using the Sequential Monte Carlo (SMC) sampling method, which is a type of Markov Chain Monte Carlo (MCMC) sampling method. Weakly informative priors for the model regression parameters are used, as recommended by Gelman et al. (2008). Based upon initial trial and error undertaken prior to this paper, the following model coefficients are modelled as having a first order linear relationship with $p(D)$ and $p(S)$: L, N, V, and C. The following parameters are modelled as having a quadratic relationship with $p(D)$ and $p(S)$: T, M,and R.

3 Bayesian methods for Models' comparison and evaluation criteria

Two different cross-validation and model selection approaches applicable to Bayesian analysis are applied in this work: the Watanabe-Akaike Information Criteria and Odds Ratios. The model com-parison methods aim to evaluate the predictive accuracy of the proposed Bayesian models and compare each model amongst one another to determine the best model that fits the observable data.

3.1 Watanabe-Akaike Information Criteria

The widely applicable, or Watanabe-Akaike, information criterion (WAIC) method, developed by Watanabe (Watanabe, 2010), is a Bayesian validation approach that measures how well a Bayesian model fits the data. The WAIC method does this by estimating the out-of-sample prediction error of a model, such as a Bayesian regression model as proposed in this paper. The WAIC method, which resolves itself to calculating a 'score' of a model, takes into account in-sample accuracy (i.e., an evaluation of how well a model fits observed data) and out-of-sample prediction (i.e., how well a model can predict unobserved/future data) (Gelman et al., 2014). WAIC is fully Bayesian in that it uses the entire posterior distribution, and it is asymptotically equal to Bayesian cross-validation (Vehtari et al., 2012). As a WAIC score of an individual model is a probabilistic metric and a model is viewed to be a better fit to data than another if its mean WAIC score is lower.

3.2 Odds Ratio

The odds ratio is a statistical model validation approach that quantifies the strength of the association between the dependent variables of a model (model predictors) and the independent variable (model outcome). Therefore, the odds ratio is used to test the sensitivity of a model to its independent variables, by evaluating the ratio between the odds of the dependent variable with and without the independent parameter under test.

In logistic regression, the odds ratio represents the constant effect of an independent model parameter on the likelihood that an outcome will occur (Chen et al., 2010). Since the logistic regression model uses a logistic function to estimate the relationship between a binary dependent variable and a group of predictor variables, the log of the odds ratio is often an easier approach to interpret. For the sake of validating our developed models for both $p(D \mid \mathbf{F}, \mathbf{W})$ and $p(S \mid \mathbf{F}, \mathbf{W})$, the log odds ratio for each of the **W** regression parameters for both $p(D)$ and $p(S)$ models are evaluated. The log odds ratio for $p(D)$ and $p(S)$ models evaluates the effect of each of the non-thermal **W** parameters on both thermal satisfaction and thermal dissatisfaction.

If the odds ratio of a model's parameter deviates from 1, it indicates that the model's outcome is sensitive to this particular model parameter (i.e. it either increases or decreases the odds of thermal satisfaction or dissatisfaction.). If the mean of the posterior traces odds ratio is close to 1, this means that the model parameter has nearly no effect on the dependent variable of interest. In other words, the dependent variable is the same in both the presence and absence of this parameter.

4 Results

4.1 The effect of non-thermal parameters of IEQ on occupant's thermal satisfaction

Figure 7 shows the first set results produced by Bayesian logistic regression models predicting the correlations between thermal dissatisfaction/satisfaction, $p(D)$ /$p(S)$, against individual non-thermal parameters of **W** assuming a fixed set of values for all other parameters. The Bayesian models used are $p(D \mid T, M, R, V, C, L, N, CO, P)$ and $p(S \mid T, M, R, V, C, L, N, CO, P)$. Unless otherwise specified, the fixed set of parameters are: T=23.3 (C), R=33%, M=23.4(C), V=0.08 m/s, C = 650 ppm, N = 47 dBA, CO = 0.7 ppm , P = 1.5 m ,and L = 463 lux. These values are chosen as they are the mean observed values for each parameter out of the dataset.

The results yield a potential relationship between surveyed thermal dissatisfaction, surveyed thermal sat-isfaction, and several non-thermal W IEQ parameters. For example, it is observed that the indoor air CO2 levels and thermal dissatisfaction, p(D), are positively correlated; which suggests that more occupants are found to be thermally dissatisfied at higher measured indoor CO2 levels than at lower indoor CO2 levels. As might be expected from this observation, the results suggest that p(S) is negatively correlated to CO2 concentrations.

The results are showing that the indoor noise levels are correlated with thermal satisfaction, as well as thermal dissatisfaction, such that the higher the measured indoor noise levels, the more likely occupants will feel thermally dissatisfied. The measured desktop illuminance appears to have a weaker correlation with perceived thermal comfort than the other measured non-thermal parameters, for both p(S) and p(D). One of the interesting parameters is that of indoor air carbon monoxide concentrations,*CO*, which show an inverse correlation with thermal dissatisfaction, as shown in Figure 7.

The occupant's desktop partition height has a strong positive correlation between thermal dissatisfaction and a negative correlation between occupant's thermal satisfaction, as perceived by occupants of open-plan offices. This indicates that the higher the desktop partition, the more the occupants will report thermal dissatisfaction with their indoor environments. Whether these results imply that the high desktop partition disrupts thermal satisfaction, or that the partition height is a proxy for another set of indoor environmental parameters: this finding along with the other conclusions drawn from the Bayesian sampling, need further verification of its statistical significance, which will be presented in detail in the following section. The observed correlations between occupants' perceived thermal comfort and CO concentration levels as well as desktop partition heights are consistent with the previous study conducted by Crosby et al. (2019). It is not clear whether the indoor carbon monoxide is a proxy of another physical issue driving its impact on occupants' thermal comfort. The source of CO should be first identified to conclude that thermal discomfort is correlated directly to CO.

The number of occurrences of each non-thermal **W** parameter across all the buildings is added to the regression models' plots in Figure 7. The results in Figure 7 are driven by both the quality and quantity of observable data, which means that if the data is a sparse certain range of values, the standard deviation of the model traces will tend to be high.

```
#generate a grid plot for the posterior probability of p(S) given F & W parameters

fig, ax = plt.subplots(5,1,figsize=(4,14))
plt.style.use('seaborn-white')
```

[ ]:

```
plot_vars = ['CO2','Noise','PART','Lux','CO']
model_id = 9 # See code above for description and ID of models

for p in range(len(plot_vars)):




  plot_var = plot_vars[p]

  # Number of individual trials to plot
  N_Samples = 500
  N_Traces = len(traces[0])

  # Default values

  Temp = 23.26
  MRT = 23.37
  RH = 33
  V= 0.076
  CO2 = 650
  Lux =463
```

```
Noise = 46.69
CO=        0.71
PART=        1.53

# Generate plot data
# ------------------------------------------
plot_trace=traces[model_id]

y_model_name_pref = 'p (' + criteria[1] + ' | '

if plot_var == 'Temp':
  C1= np.arange(20,26,0.2)
  xlabel_text = 'Operative Temperature [$^\circ$C]'
  data_key = 'OT'
elif plot_var == 'CO2':
  C1= np.arange(450,965,5)
  xlabel_text = 'CO$_2$ Levels [ppm]'
  data_key = 'CO2'
elif plot_var == 'PART':
  C1= np.arange(1.0,2.5,0.02)
  xlabel_text = 'Partition Height [m]'
  data_key = 'PART'
elif plot_var == 'Noise':
  C1= np.arange(40,57,0.25)
  xlabel_text = 'Noise levels [dBA]'
  data_key = 'Noise'
elif plot_var == 'Lux':
  C1= np.arange(100,1300,5)
  xlabel_text = 'Desktop Illuminance [lux]'
  data_key = 'Lux'
elif plot_var == 'CO':
  C1= np.arange(0.01,1.7,0.01)
  xlabel_text = 'CO Concentrations [ppm]'
  data_key = 'CO'

plot_results = np.zeros((len(C1),N_Samples))
means = []
upper_95 = []
lower_95 = []
j=0

for xN in C1:
  if plot_var == 'Temp':
    Temp = xN
    MRT = xN
  elif plot_var == 'CO2':
    CO2 = xN
  elif plot_var == 'PART':
    PART = xN
  elif plot_var == 'Noise':
    Noise = xN
  elif plot_var == 'Lux':
    Lux = xN
```

```
    elif plot_var == 'CO':
      CO = xN

#data = data[['WSID','building#','CLASS','Temp','PART','MRT', 'RH',␣
↪'CO2','Lux','CO','Noise','V','SAT_TEMP']]
    CoF_NP = np.zeros((1,13))

    CoF_NP[0,0] = 0 # WSID | 0
    CoF_NP[0,1] = 0 # building | 1
    CoF_NP[0,2] = 0 # CLASS | 2
    CoF_NP[0,3] = Temp # Temp | 3
    CoF_NP[0,4] = PART # PART | 4
    CoF_NP[0,5] = MRT # MRT| 5
    CoF_NP[0,6] = RH # RH | 6
    CoF_NP[0,7] = CO2 # CO2 | 7
    CoF_NP[0,8] = Lux # Lux | 8
    CoF_NP[0,9] = CO # CO | 9
    CoF_NP[0,10] = Noise # Noise | 10
    CoF_NP[0,11] = V # V | 11
    CoF_NP[0,12] = 0 # SAT_TEMP | 12

    CoF_Sc = min_max_scaler.transform(CoF_NP)

    CoF_Sq = np.zeros((1,5))
    CoF_Sq[0,0] = np.square(CoF_Sc[0,3]) # T2 | 13
    CoF_Sq[0,1] = np.square(CoF_Sc[0,5])  # MRT2 | 14
    CoF_Sq[0,2] = np.square(CoF_Sc[0,7])  # CO2_2 | 15
    CoF_Sq[0,3] = np.square(CoF_Sc[0,8])  # Lux_2 | 16
    CoF_Sq[0,4] = np.square(CoF_Sc[0,6])  # RH_2 | 17

    CoF_Sc=np.append(CoF_Sc,CoF_Sq,axis=1)

    regress_Eq = traces[model_id].Beta_0[:N_Samples]
    y_model_name = ''
    const_value_script = '$\mathbf{Fixed\, values:}$\n'

    for i in range(N_terms):
      if key[model_id,i]==1:
        if i == 0:
          regress_Eq = regress_Eq+traces[model_id].Beta_1[:N_Samples]*CoF_Sc[0,3]
          y_model_name = y_model_name + 'T,'
          if plot_var!='Temp':
            const_value_script = const_value_script + "T = " + str(Temp) + ' $^\circ$C␣
↪\n'

        elif i==1:
          regress_Eq = regress_Eq+traces[model_id].Beta_2[:N_Samples]*CoF_Sc[0,5]
          y_model_name = y_model_name + 'M,'

        elif i==2:
          regress_Eq = regress_Eq+traces[model_id].Beta_3[:N_Samples]*CoF_Sc[0,6]
          y_model_name = y_model_name + 'R,'
          if plot_var!='RH':
```

```
            const_value_script = const_value_script + "RH = " + str(RH) + '% \n'

        elif i==3:
          regress_Eq = regress_Eq+traces[model_id].Beta_4[:N_Samples]*CoF_Sc[0,11]
          y_model_name = y_model_name + 'V,'
          if plot_var!='vel':
            const_value_script = const_value_script + "V = " + str(V) + ' m/s \n'

        elif i==4:
          regress_Eq = regress_Eq+traces[model_id].Beta_5[:N_Samples]*CoF_Sc[0,13]

        elif i==5:
          regress_Eq = regress_Eq+traces[model_id].Beta_6[:N_Samples]*CoF_Sc[0,14]

        elif i==6:
          regress_Eq = regress_Eq+traces[model_id].Beta_7[:N_Samples]*CoF_Sc[0,17]

        elif i==7:
          regress_Eq = regress_Eq+traces[model_id].Beta_8[:N_Samples]*CoF_Sc[0,7]

          y_model_name = y_model_name + 'C,'
          if plot_var!='CO2':
            const_value_script = const_value_script + "CO$_2$ = " + str(CO2) + ' ppm␣
↪\n'

        elif i==8:
          regress_Eq = regress_Eq+traces[model_id].Beta_9[:N_Samples]*CoF_Sc[0,15]

        elif i==9:
          regress_Eq = regress_Eq+traces[model_id].Beta_10[:N_Samples]*CoF_Sc[0,9]
          y_model_name = y_model_name + 'CO,'
          if plot_var!='CO':
            const_value_script = const_value_script + "CO = " + str(CO) + ' ppm \n'

        elif i==10:
          regress_Eq = regress_Eq+traces[model_id].Beta_11[:N_Samples]*CoF_Sc[0,4]
          y_model_name = y_model_name + 'P,'
          if plot_var!='PART':
            const_value_script = const_value_script + "H$_{part}$ = " + str(PART) + '␣
↪m \n'

        elif i==12:
          regress_Eq = regress_Eq+traces[model_id].Beta_13[:N_Samples]*CoF_Sc[0,10]
          y_model_name = y_model_name + 'N,'
          if plot_var!='Noise':
            const_value_script = const_value_script + "N = " + str(Noise) + ' dBA \n'

        elif i==13:
          regress_Eq = regress_Eq+traces[model_id].Beta_14[:N_Samples]*CoF_Sc[0,8]
          y_model_name = y_model_name + 'L,'
          if plot_var!='Lux':
            const_value_script = const_value_script + "L = " + str(Lux) + ' lux \n'
```

```
  plot_results[j,:] = (1 / (1+np.exp(-(regress_Eq))))
  means.append(np.mean(plot_results[j,:]))
  stdev = np.std(plot_results[j,:])
  upper_95.append(means[j]+stdev)
  lower_95.append(means[j]-stdev)
  j=j+1

# Configure and generate plots
# ------------------------------------
# Display fixed-variable data in top-left
#ax.text(0.03, 0.95,const_value_script, ha='left', va='top', transform=ax.transAxes)

ax2 = ax[p].twinx()
ax2=seaborn.distplot(data_original[data_key], hist=True, kde=False,
            color = 'black',
           hist_kws={'edgecolor':'grey','alpha':0.1},
         kde_kws={'shade':False,'linewidth':0.5})

mean_val=data_original[data_key].mean()
upper_std1=(data_original[data_key].mean()+data_original[data_key].std())
lower_std2=(data_original[data_key].mean()-data_original[data_key].std())

h1=ax[p].axvline(mean_val, color='black', linestyle='-.', linewidth=1.1)
#h2=plt.axvline(upper_std1, color='brown', linestyle='--', linewidth=0.8)
#h3=plt.axvline(lower_std2, color='brown', linestyle='--', linewidth=0.8)

#ax2.axvspan(upper_std1, lower_std2, alpha=0.05, color='black')

# Generate fill of standard deviation above and below mean
N_name_chars = len(y_model_name)
y_model_name = y_model_name_pref + y_model_name[0:N_name_chars-1] + ')'
ax[p].fill_between(C1, lower_95, upper_95, facecolor='red' ,alpha=0.2)

# Generate individual plots of each trace
for k in range(N_Samples):
  ax[p].plot(C1 ,plot_results[:,k],'blue',linewidth=0.04,alpha=0.04)

# Plot lines of mean and standard deviation
ax[p].plot(C1 ,means,'red')
ax[p].plot(C1 ,lower_95,':',color='red',alpha=1)
ax[p].plot(C1 ,upper_95,':',color='red',alpha=1)

# Configure grid and set plot limits
ax[p].grid(color='black',linestyle=':',linewidth=1,alpha=0.1)

if criteria == 'pD':
  ax[p].set_yticks(np.arange(.2, .61, 0.1))
  ax[p].set_ylim((0.2,0.55))
else:
  ax[p].set_yticks(np.arange(.3, .8, 0.1))
  ax[p].set_ylim((0.39,0.8))

ax2.set_yticks([0,50,100])
```

```
  ax[p].set_xlim((min(C1),max(C1)))
  ax[p].set_xlabel(xlabel_text)

  for tl in ax2.get_yticklabels():
      tl.set_color('darkgrey')

  ax2.set_ylabel('Number of \noccurences', color='darkgrey', fontsize=12)
 # ax[p].set_ylabel('P (D|F,W)', fontsize=9)
  if criteria == 'pD':
     ax[p].set_ylabel('P (D | F, W)', fontsize=12)
  else:
      ax[p].set_ylabel('P (S | F, W)', fontsize=12)
#" + r"$\bf{" + str(number) + "}$"
  # Write axis labels
  ax[p].set_xlabel(xlabel_text,fontsize=12)
  ax[0].set_title(y_model_name+'\n',fontsize=13,fontweight="bold")

legend_elements = [Line2D([0], [0], color='r', lw=1, label='Mean of posterior traces'),
                   Line2D([0], [0], color='r', ls=':',lw=1, label='Standard deviation␣
 ↪of posterior traces'),
                   Line2D([0], [0], color='blue',lw=0.2, label='Individual posterior␣
 ↪traces'),
                   Patch(facecolor='lightgrey', edgecolor='grey',
                         label='Histogram of occurrences')]
#ax[p].legend(handles=legend_elements,loc=9, bbox_to_anchor=(0.35, -0.6),fontsize=9)

fig.tight_layout(pad=0.4, w_pad=0.5, h_pad=1.0)
```

[8]:

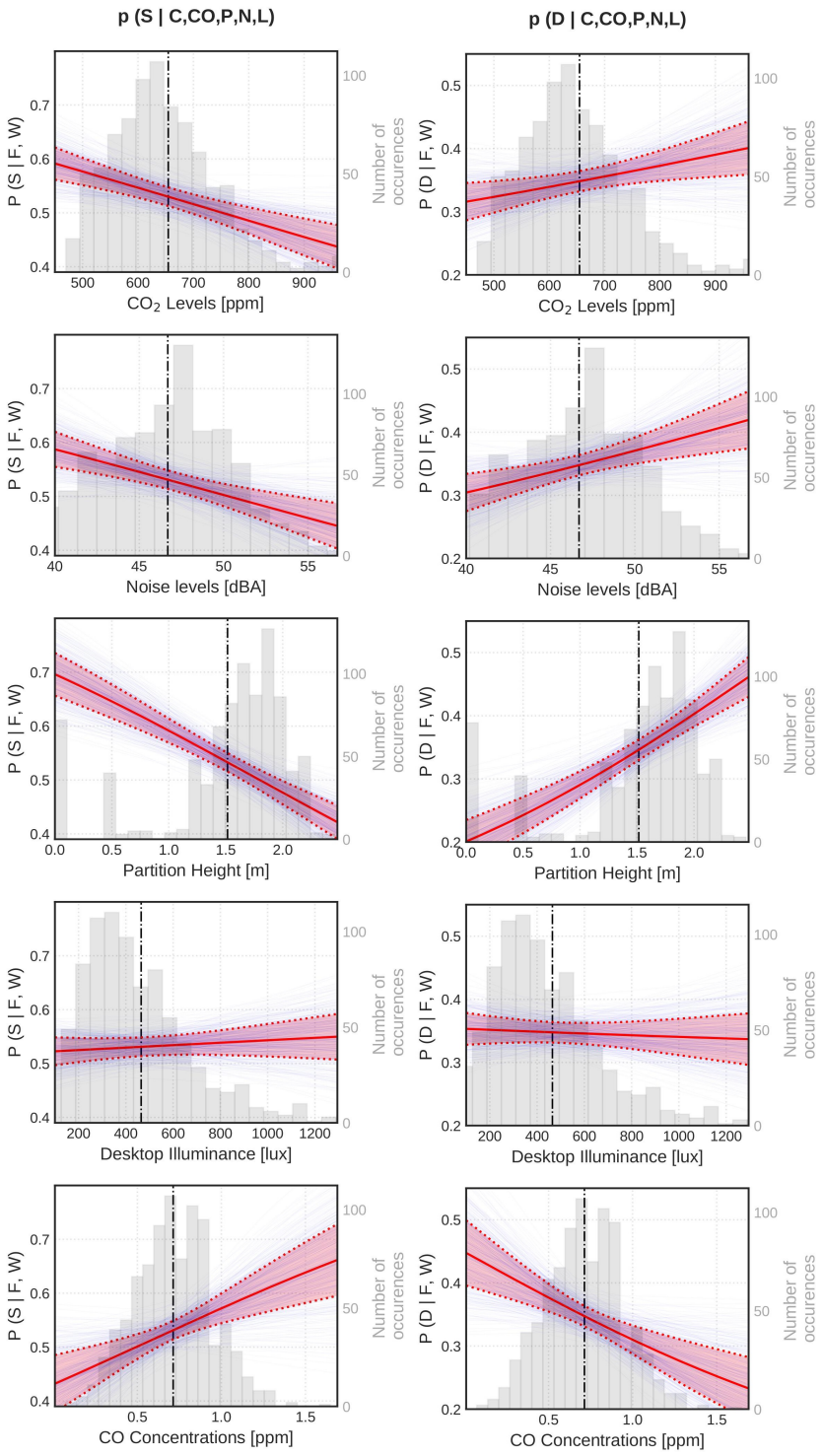

Figure 6: Probability p(D | **F**,**W**) and p(S | **F**,**W**) with thin blue lines indicating individual sample traces, solid red lines indicate mean predicted value from all traces, dashed red bands indicate the standard error of traces, grey bars indicate the probability distribution of each independent parameter, and black dashed centre lines are the mean values of observations

4.2 Models comparison and selection

In this section, model comparison and cross-validation methods are used to answer 3 main research questions. The first question is: "What is the best 'thermal' model, analogous to Fanger classic model, that represents the data in the most accurate way?" To determine which thermal model is the best, all the candidate models for both p(D) and p(S) are compared to Fanger model. Bayesian Factors and WAIC scores are evaluated for all the candidate models to select the best thermal model accordingly. Answering the first question is crucial for the second research question, which is: "Do non-thermal parameters of IEQ improve thermal comfort predictions in open-plan offices?" In order to answer this question, all the candidate models that take into account both thermal and non-thermal parameters of IEQ (**F** and **W**) are compared with the best thermal model selected in step 1.

The first phase of models selection's results revealed that the thermal model that best represents thermal comfort for the COPE dataset for both $p(S)$ and $p(D)$ is the model that includes air temperature, T, and mean radiant temperature, M (excluding air velocity and relative humidity from the original Fanger model) i.e. $p(S/D|T, M)$. While for the UBC + COPE dataset, the best thermal model for p(S) is also $p(S|T, M)$ and the best thermal model for $p(D)$ is $P(D|T, M, R, V)$ which is analogous to Fanger model. Having different thermal models representing $p(S)$ and $p(D)$ might not be very practical for real field applications but maybe or maybe not physically more correct. This difference might be due to the lack of p(D) data, especially in the newly collected data, as shown in Figure 4. This is also supported by the fact that the best thermal model representing the COPE data is the same for both p(S) and p(D) as there is enough data on thermal dissatisfaction to draw such a conclusion, while the UBC data lacks having enough p(D) data that is sufficient to select the best thermal model. Physically speaking, it might be also best to represent thermal dissatisfaction using air velocity and relative humidity in addition to temperature and mean radiant temperature, while T and MRT might be sufficient for p(S) modelling, as the results suggest. It is also important to note that the COPE dataset collected 17 years ago by the NRC doesn't include data on occupants' clothing insulation (clo) and metabolic rates (met), for this reason, the prior results didn't include clo and met. However, the recently collected data in UBC has clothing insulation and metabolic rate data, which is guaranteed to be included in the future analysis on the UBC dataset alone. Since, this study mainly aims to update the prior field IEQ dataset while strengthening the previous results, including clo and met for the UBC data is out of this paper's scope.

To answer the second research question, the models are compared to the 'Null hypothesis' to evaluate the effects of adding non-thermal IEQ parameters to the 'best thermal model'. Tables 2 and 3 show the WAIC scores for the COPE dataset $p(S)$ and $p(D)$ models. Tables 4 and 5 show the WAIC scores for the COPE +UBC dataset $p(S)$ and $p(D)$ models. The null hypothesis is shown in red for each of the studied cases.

Table 2: WAIC Scores generated for the p(S) models for the COPE dataset. The Null hypothesis shown in red

Bayesian Model	WAIC Score	WAIC standard error
P(S $\mid C$)	1002.13	4.06
p(S $\mid T, M, C$)	999.65	4.26
P(S $\mid C, S, N, L$)	1002.82	5.29
p(S $\mid T, M, L$)	1002.14	3.35
P(S $\mid T, M, R, V, C$)	1000.27	4.79
P(S $\mid T, M, N$)	1002.25	3.66
$P(S \mid T, M)$	1004.38	3.06
P(S $\mid T, M, R, C$)	1002.15	4.285
P(S $\mid T, M, R, V, L$)	1002.43	3.89
P(S $\mid T, M, R, V, S$)	1003.38	4.42
P(S $\mid T, M, R, V$)	1003.14	3.57
P(S $\mid T, M, R, V, C, S, N, L$)	1003.03	5.35
P(S $\mid T, M, R$)	1004.07	3.12

Bayesian Model	WAIC Score	WAIC standard error
P(S $\mid T, M, R, V, N$)	1004.32	4.04

Table 3: WAIC Scores generated for the p(D) models for the COPE dataset. The Null hypothesis shown in red

Bayesian Model	WAIC Score	WAIC standard error
P(D $\mid T, M, N$)	955.01	13.934
p(D $\mid T, M, C$)	954.95	13.89
p(D $\mid T, M, L$)	955.89	13.78
P(D $\mid C$)	958.26	13.76
P(D $\mid C, S, N, L$)	958.14	14.08
P(D $\mid T, M, R, V, S$)	955.56	14.11
$P(D \mid T, M)$	960.22	13.62
P(D $\mid T, M, R, C$)	956.63	13.88
P(D $\mid T, M, R, V$)	956.81	13.82
P(D $\mid T, M, R, V, N$)	957.05	14.03
P(D $\mid T, M, R$)	957.53	13.66
P(D $\mid T, M, R, V, C$)	956.78	13.58
P(D $\mid T, M, R, V, L$)	958.81	13.83
P(D $\mid T, M, R, V, C, S, N, L$)	958.73	14.49

Table 4: WAIC Scores generated for the p(S) models for the COPE + UBC datasets. The Null hypothesis shown in red

Bayesian Model	WAIC Score	WAIC standard error
P(S $\mid C, N, L, CO, P$)	1182.0	10.12
P(S $\mid T, M, R, V, C, N, L, CO, P$)	1184.46	10.6
P(S $\mid P$)	1192.55	7.82
p(S $\mid CO$)	1199.2	5.67
P(S $\mid C$)	1201.97	4.92
P(S $\mid T, M, C$)	1200.65	5.38
P(S $\mid C, N, L$)	1202.69	5.46
P(S $\mid T, M, R, V, C$)	1201.25	6.21
P(S $\mid T, M, R, V, C, N, L$)	1200.93	6.79
P(S $\mid T, M, R, C$)	1201.63	5.65
$P(S \mid T, M)$	1204.48	4.19
P(S $\mid T, M, N$)	1204.51	4.64
P(S $\mid T, M, L$)	1205.69	4.26
P(S $\mid T, M, R$)	1205.11	4.76
P(S $\mid T, M, R, V$)	1205.37	5.02

Table 5: WAIC Scores generated for the p(D) models for the COPE + UBC datasets. The Null hypothesis shown in red

Bayesian Model	WAIC Score	WAIC standard error
P(D $\mid C, N, L, CO, P$)	1110.96	19.26
P(D $\mid T, M, R, V, C, N, L, CO, P$)	1109.97	19.44
P(D $\mid P$)	1116.4	18.52
$P(D \mid T, M, R, V)$	1121.13	18

Bayesian Model	WAIC Score	WAIC standard error
p(D | CO)	1123.95	17.71
P(D | T, M, R, V, C)	1120.74	18.01
P(D | T, M, R, V, C, N, L)	1121.77	18.2
P(D | T, M, R)	1122.52	17.81
P(D | T, M, R, C)	1123	18.01
$P(D \mid T, M)$	1128.53	17.28 7
P(D | T, M, C)	1128.79	17.51
P(D | C)	1131.15	17.32
P(D | T, M, N)	1129.627	17.297
P(D | T, M, L)	1129.98	17.41
P(D | C, N, L)	1133.05	17.5

The COPE model comparison results suggested that many non-thermal parameters of IEQ appear to have an effect on thermal satisfaction and thermal dissatisfaction over the best thermal model, or the 'Null hypothesis'. For example, it is shown that adding CO_2 to the Null hypothesis improves thermal satisfaction predictions. Similarly, noise levels, desk-top illuminance are showing statistical significance over the null hypothesis, as suggested by the WAIC scores results.

Adding the UBC data to the COPE dataset, it is evident that some non-thermal parameters are showing an even more strong correlation with thermal comfort than the COPE dataset alone. It is important to note that, in this study, desktop partition, P, and CO levels are added to the other non-thermal parameters studied in the previous COPE study. The results are suggesting that all the studied non-thermal parameters have an independent effect on thermal satisfaction varying in significance, with CO levels and desktop partition being the most significant ones, as the results suggest. While $p(D)$ results are showing weaker significance for some of the non-thermal parameters, this is due to the lack of $p(D)$ data collected from the UBC dataset.

The third and apparently the most important question to answer now is: "what is the best model to predict thermal comfort in open-plan offices?" The results above are suggesting that the best model to predict thermal satisfaction and thermal dissatisfaction should include CO_2, CO levels as well as noise levels, desktop illuminance and partition height. The results are also suggesting that the second-best model representing thermal comfort (both p(S) and p(D)) includes all the non-thermal and thermal parameters, i.e. $p(S/D|$ T, M, R, V, C, N, L, CO, P). This might lead to the conclusion that to best model oc-cupant's thermal comfort in open-plan offices, parameters of non-thermal IEQ should be considered along with the Fanger thermal parameters.

It might be interesting to conclude that the best thermal comfort model doesn't seem to have any thermal parameters included. It might be also that some of these non-thermal parameters are a proxy of some other thermal parameters embedded in them or maybe those non-thermal parameters, which are also a proxy for an occupant's well-being and mood, are doing a better job in predicting thermal comfort on their own. Either way, this is such an interesting and novel conclusion to find which is aligning with the previous results drawn from the COPE dataset making it robust and more significant.

```
[ ]: #generate a grid plot for odds ratio for p(S) models
fig, ax = plt.subplots(5,1,figsize=(4,12))
plt.style.use('seaborn-white')
#CO2
plt.subplot(5, 1, 1)
model_id = 5
b = traces[model_id]['Beta_8']
plot_label = r'$\beta_{CO2}$, $CO_2$, concentrations'
seaborn.distplot(np.exp(b), hist=False, kde=True, label=plot_label,
```

```
            hist_kws={'edgecolor':'grey','alpha':0.1},
            kde_kws={'shade':False,'linewidth':1})

h1=plt.axvline(1, color='black', linestyle='-.', linewidth=1.1)
plt.ylabel("Probability (-)")
plt.xscale("log")
plt.xlim(0.1,10)
plt.grid(True,which="both",ls=':',c='gray',alpha=0.5)
plt.legend(loc='upper right')
#CO
model_id = 9
plt.subplot(5, 1, 2)
b = traces[model_id]['Beta_10']
plot_label = r'$\beta_{CO}$, COconcentrations'
seaborn.distplot(np.exp(b), hist=False, kde=True, label=plot_label,
            hist_kws={'edgecolor':'grey','alpha':0.1},
            kde_kws={'shade':False,'linewidth':1})

h1=plt.axvline(1, color='black', linestyle='-.', linewidth=1.1)
plt.ylabel("Probability (-)")
plt.xscale("log")
plt.xlim(0.1,10)
plt.grid(True,which="both",ls=':',c='gray',alpha=0.5)
plt.legend(loc='upper left')
#Lux
model_id = 9
plt.subplot(5, 1, 3)
b = traces[model_id]['Beta_14']
plot_label = r'$\beta_{L}$, desktop illuminance'
seaborn.distplot(np.exp(b), hist=False, kde=True, label=plot_label,
            hist_kws={'edgecolor':'grey','alpha':0.1},
            kde_kws={'shade':False,'linewidth':1})

h1=plt.axvline(1, color='black', linestyle='-.', linewidth=1.1)
plt.ylabel("Probability (-)")
plt.xscale("log")
plt.xlim(0.1,10)
plt.grid(True,which="both",ls=':',c='gray',alpha=0.5)
plt.legend(loc='upper left')
#Noise
model_id=11
plt.subplot(5,1,4)
b = traces[model_id]['Beta_13']
plot_label = r'$\beta_{N}$, noise levels'
seaborn.distplot(np.exp(b), hist=False, kde=True, label=plot_label,
            hist_kws={'edgecolor':'grey','alpha':0.1},
            kde_kws={'shade':False,'linewidth':1})

h1=plt.axvline(1, color='black', linestyle='-.', linewidth=1.1)
plt.ylabel("Probability (-)")
plt.xscale("log")
plt.xlim(0.1,10)
plt.grid(True,which="both",ls=':',c='gray',alpha=0.5)
```

```
plt.legend(loc='upper right')
#PART
model_id=9
plt.subplot(5,1,5)
b = traces[model_id]['Beta_11']
plot_label = r'$\beta_{P}$, Partition Height'
seaborn.distplot(np.exp(b), hist=False, kde=True, label=plot_label,
             hist_kws={'edgecolor':'grey','alpha':0.1},
             kde_kws={'shade':False,'linewidth':1})

h1=plt.axvline(1, color='black', linestyle='-.', linewidth=1.1)
plt.ylabel("Probability (-)")
plt.xscale("log")
plt.xlim(0.1,10)
plt.grid(True,which="both",ls=':',c='gray',alpha=0.5)
plt.legend(loc='upper right')
#########################################################################
plt.xlabel("Odds Ratio of Posterior Traces-P(S) Models")
fig.tight_layout(pad=0.4, w_pad=0.5, h_pad=1.0)
```

[ ]:

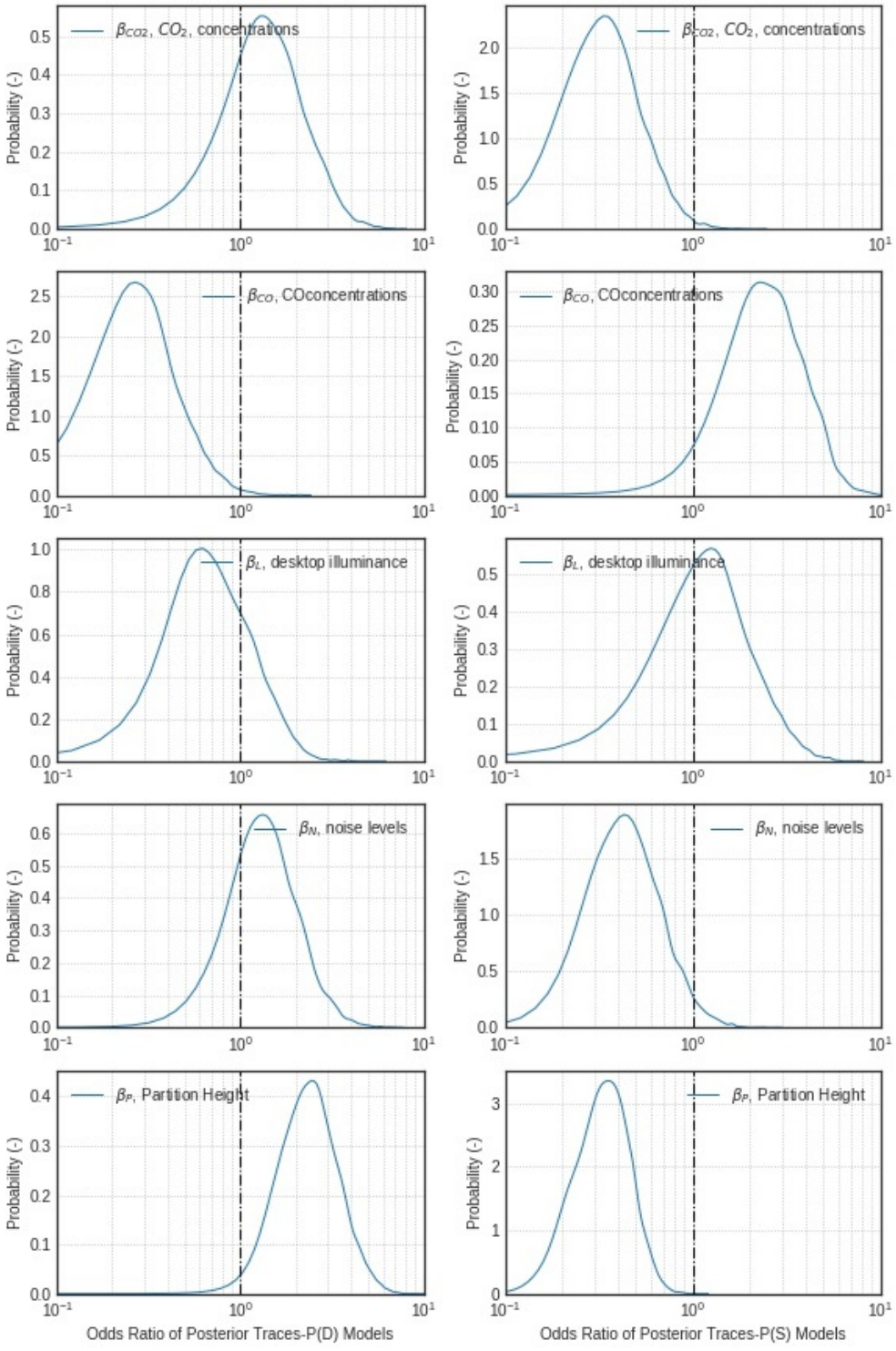

Figure 7: Odds Ratio of posterior traces of non-thermal IEQ parameters for $p(S)$ and $p(D)$ Bayesian models

The log odds ratio for each of the non-thermal **W** parameters under study, for both $p(D)$ and $p(S)$ models, are evaluated and summarized in Figure 8. Each of these plots represents one of the studied **W** indoor en-

vironmental design conditions. It is observed from Figure 8 that some of the non-thermal parameters have a significant effect on thermal satisfaction and dissatisfaction; for example, the odds ratio of the posterior traces of the CO_2's regression parameter of the $p(S)$ model deviates significantly from 1, which means that there is a significant correlation between $C\ O_2$ levels and the occupant's thermal satisfaction.

It is also shown that both the partition height and the measured CO concentrations have strong effects on the occupant's thermal dissatisfaction and satisfaction, which is clear from the odds ratio curve being on the far left, and far-right respectively.

Similarly, it is apparent from Figure 8 that there is an effect of the noise levels on the occupant's thermal dissatisfaction and a weaker effect on thermal dissatisfaction. It is also apparent that the desktop illuminance doesn't seem to have much of an effect on both $p(D)$ and $p(S)$, which is consistent with the previous model comparison results.

These findings appear to also support the previous research studies which correlated the occupant's mood and overall satisfaction (represented here as the non-thermal parameters of IEQ, **W** dataset) with occupant's thermal comfort, as discussed earlier in the paper. The non-thermal parameters of IEQ which are a proxy for the occupant's overall mood and wellness are affecting thermal comfort, as suggested by previous studies in the literature. This might explain the underlying mechanism for the strong association between thermal comfort and some of the **W** parameters studied in this paper which appear to disrupt the occupant's thermal comfort, like the CO_2 levels and the partition height, in particular.

5 Conclusions

This paper presented an update to a prior work that examined the correlations between non-thermal indoor environmental conditions on the perceived thermal comfort of occupants in open-plan offices. The prior work made use of the COPE dataset, an IEQ field database of 800 occupants of open-plan offices in large Canadian and US cities in the early 2000s. It has been found that there exist statistically significant inde-pendent correlations between thermal comfort, as perceived by occupants of open-plan offices, and many measured non-thermal, wellbeing-type IEQ parameters. The most significant finding is that a modest in-crease in measured indoor CO2 concentrations, from 500 ppm to 900 ppm, was found to be correlated with a decrease in perceived thermal satisfaction by 30%.

This paper provides an update to that work based on a further field study of 150 offices carried out at the University of British Columbia campus across 2019/2020. A novel Bayesian inference methodology is applied to the expanded IEQ dataset to isolate meaningful correlations between perceived thermal comfort and non-thermal metrics. It culminates in a proposed extension to the Fanger PMV thermal comfort model to include the effects of meaningful non-thermal indoor conditions when predicting thermal satisfaction and dissatisfaction. To determine whether non-thermal parameters can improve the prediction of indoor thermal comfort and test the prior research findings, 15 different cases of thermal and non-thermal variables were studied. Each combination is used to define a logistic regression model that seeks to relate the proba-bility of an occupant feeling thermally satisfied, or dissatisfied, to parameters of thermal and non-thermal IEQ.

The posterior prediction results of the Bayesian models reveal that there exist significant correlations between surveyed thermal comfort and many non-thermal parameters of IEQ. Many of these significant correlations align with the previous results found by the 17-years old dataset. The significance of all the investigated models was tested using cross-validation methods: The Odds Ratios and WAIC scores. The results revealed that, in several cases, there is evidence to suggest that several non-thermal parameters do have an independent effect on perceived thermal dissatisfaction/satisfaction,

at least in the open-plan office. Having repeatable results under different datasets types and a variety of model input conditions and using a new set of experimental procedures and methodologies indicates that the previous results and findings are robust and significant.

This research advances a novel technique for improving current thermal comfort models while filling a prevailing research gap vis-à-vis standard models of thermal comfort is that these models have not always accurately predicted true thermal comfort observations found in the field. This is the first research to suggest that there exist correlations between measured non-thermal parameters of IEQ and perceived thermal comfort and that predictions of thermal comfort can be possibly improved by implementing metrics on non-thermal indoor environmental quality.

Appendix C

UBC Field IEQ Study- Survey Questions

Survey categories Office chair	**Individual survey questions**	**Occupant's response**
Background	What is your gender?	M; F; prefer not to say
	What is your age group?	25 or under; 30-50; Above 50
	How long have you been working in this office?	1. Less than 4 months; 2. 4 to 8 months; 3. 8 months to a year; 4.More than a year
	How many days per week do you usually spend in your office?	1. Under 3 days/week 2. Between 3-5 days/week 3. 5 days/week or more
	How many hours per day do you usually spend in your office?	1. Under 5 hours/day 2. Between 5-8 hours/day 3. 8 hours/day or more
Air temperature- Comfort	How would you rate the temperature at your workspace right now?	7-point scale (cold to hot)
Air temperature- Satisfaction	How would you rate your satisfaction with the air temperature at your workspace right now?	7-point scale (Very unsatisfied to very satisfied)
Air temperature- Preference	How would you rate your preference with respect to the air temperature at your workspace right now?	1. Prefer warmer; 2. Remain the same; 3. Prefer cooler)
Air temperature- Controllability	Do you feel you are able to control the temperature at your workplace (e.g., operable window, thermostat, etc.)?	3-point scale (Yes, definitely; Yes, somewhat; No)
Air temperature- Long term satisfaction	Over the PAST YEAR, how satisfied have you been with the overall air temperature at your workspace?	3 subcategories: Entire year; Summer; Winter: 7-point scale (Very unsatisfied to very satisfied)
	During a typical work week, how many times do you feel dissatisfied with the air temperature at your workspace?	1.At least once per day 2.3-5 times per week 3.Once or twice per week 4.Rarely

Table C.1: UBC field IEQ study- Survey questions-Part 1

Survey categories	Individual survey questions	Occupant's response
Air temperature- Long term preference	Over the PAST YEAR, how would you rate your preference with respect to the typical air temperature at your workstation?	2 subcategories: Summer; Winter: 1. Prefer warmer; 2. Remain the same 3. Prefer cooler
Workstation and Ergonomics	How would you rate your satisfaction with the following design aspects of your workstation? - The size of your workstation - The amount of privacy at your workstation -The overall layout and design of your office	7-point scale (Very unsatisfied to very satisfied)
	Is your desk height-adjustable?	Yes; No
	Is your chair/seat adjustable (height, tilt, etc.)?	Yes; No
	How would you rate your satisfaction with the comfort provided by your office chair?	7-point scale (Very unsatisfied to very satisfied)
	Is your computer screen adjustable (height, angle, etc.)?	Yes; No
	Do you use personal headphones while you work?	1. Always; 2. Sometimes; 3. Never
Noise levels	How would you rate the background noise levels around you at your workspace right now?	7-point scale (Very unsatisfied to very satisfied)
	How would you rate your long-term satisfaction with the background noise levels at your office ?	7-point scale (Very unsatisfied to very satisfied)
	Which factors do you find affect the background noise levels at your workstation the most?	1. Conversations and speech 2. Electronic devices and appliances 3. Traffic 4. Air-conditioning 5. Other
Biophilia	How would you rate the amount of indoor plants or greenery at your workspace right now?	7-point scale (Very unsatisfied to very satisfied)
Lighting intensity	How would you rate the amount of light received at your desktop at your workspace right now?	7-point scale (Very unsatisfied to very satisfied)
Daylight	How much daylight do you feel your workstation receives on an average workday?	1. A lot 2. A moderate amount 3. Very little 4. None at all
	Are you able to control the amount of daylight you receive?	1. Yes, 2. Definitely 3. Yes, somewhat No
	How would you rate your satisfaction with the daylight availability at your workspace right now?	7-point scale (Very unsatisfied to very satisfied)
Computer screen glare	How satisfied are you with the visibility of your computer screen in relation to glare (from sunlight or reflected light from outdoors) ?	7-point scale (Very unsatisfied to very satisfied)
View to the outside	Would you consider your workstation to have a good view to the outside?	Yes; No

Table C.2: UBC field IEQ study- Survey questions-Part 2

Survey categories	Individual survey questions	Occupant's response
Artificial lighting	With respect to artificial (electric) lighting, how would you rate your satisfaction with the amount of light at your workspace right now?	7-point scale (Very unsatisfied to very satisfied)
	How would you rate your preference with respect to light levels at your workspace?	1. Prefer darker 2. Keep the same 3. Prefer brighter
	Do you feel you have adequate control over the electric lights at your workspace (dimmable light switches, desk lamp, on/off control)?	1. Yes, definitely 2. Yes, somewhat 3. No
Lighting colour	How would you describe the colour tone around your workspace? (Check all that apply)	1. Warm; 2. Cool; 3. Bright; 4. Dark
Air Quality	How would you rate your satisfaction with the air quality at your workspace right now?	7-point scale (Very unsatisfied to very satisfied)
	Do you have access to an operable window?	Yes; No
Olfactory Comfort	Are you currently bothered by the smells, scents, or perfumes around your workspace?	1. Yes, definitely 2. Yes, somewhat 3. No
Air movement	How would you rate the air movement at your workspace?	1. Still 2. Somewhat drafty 3. Drafty 4.V ery drafty
	How would you rate your satisfaction with the air movement at your workspace?	7-point scale (Very unsatisfied to very satisfied)
Humidity	How would you rate the humidity at your workspace?	1.Very humid 2.Slightly humid 3.Neutral 4.Slightly dry 5.Very dry
	How would you rate your satisfaction with the humidity at your workspace?	7-point scale (Very unsatisfied to very satisfied)
IEQ Controllability	How would you rate your OVERALL satisfaction with the level of control you have on the indoor environment at your workspace?	7-point scale (Very unsatisfied to very satisfied)
Workplace IEQ improvements	Which one of the following conditions do you think is most important to address, in terms of improving your indoor environment?	1. Lighting comfort 2. Acoustic comfort 3. Amount of plants or greenery around your workstation 4. Indoor air quality 5. Thermal comfort 6. None, I feel super happy with everything.
Additional comments	Do you have any other comments to share about the overall environmental quality at your workstation?	(please type)

Table C.3: UBC field IEQ study- Survey questions-Part 3