

Toward Resilient Building Design in Energy Performance under Climate Change

Mirata Hosseini

A Thesis

In the Department

of

Building, Civil, and Environmental Engineering

**Presented in Partial Fulfillment of the Requirements
For the Degree of
Doctor of Philosophy (Building Engineering) at
Concordia University
Montreal, Québec, Canada**

February 2021

© Mirata Hosseini, 2021

CONCORDIA UNIVERSITY
School of Graduate Studies

This is to certify that the thesis prepared

By: Mirata Hosseini

Entitled: Toward Resilient Building Design in Energy Performance under Climate Change

and submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy (Building Engineering)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final Examining Committee:

_____ Chair
Dr. Yuhong Yan, Concordia University

_____ External Examiner
Dr. Yi Huang, McGill University

_____ External to Program
Dr. Dongyu Qiu, Concordia University

_____ Examiner
Dr. Andreas Athienitis, Concordia University

_____ Examiner
Dr. Ali Nazemi, Concordia University

_____ Thesis Supervisor
Dr. Bruno Lee, Concordia University

Approved by _____
Dr. Ashutosh Bagchi, Chair of Department

March/17/2021 _____
Dr. Mourad Debbabi, Dean of Faculty

ABSTRACT

Toward Resilient Building Design in Energy Performance under Climate Change

Mirata Hosseini, Ph.D.

Concordia University, 2021

Building energy simulation is commonly used to evaluate the energy performance of buildings to support decisions made at the design stage or to quantify potential energy savings of various strategies for retrofitting existing buildings. However, in many cases, the anticipated performance through simulation output significantly deviates from actual measured data. A major reason for such discrepancy is due to uncertainty in the simulation inputs.

One source of input uncertainty is weather data representing the climate condition. In order to predict the long-term performance of the buildings with energy simulation, modellers commonly use a single Typical Meteorological Year (TMY) weather data file which supposedly represents the climatic conditions. The single weather year file is composed of hourly resolution data from the most 12 representative calendar months of 30 years which are selected based on statistical similarity to long-term weather daily-averaged data. These weather files are synthetically constructed on historical weather data over a long period of time for an array of weather parameters, such as solar radiation, temperature, wind speed and others. The statistical procedure to construct the weather files depends on the weights assigned to these weather parameters. Under current practice, these weighting factors are universally assigned regardless of climatic locations nor the building application. This approach leads to energy performance predictions that deviate from the long-term averages.

Nevertheless, the single weather file ignores the variation in building energy performance resulted from natural weather variation. This source of uncertainty becomes even more critical when the long-term superimposed effect driven by human and anthropogenic factors are added to natural variation. Historical weather data shows that compared to other regions, higher latitudes, including Canada, have been affected more by climate change, and it is expected that this change will be even more in the years to come. Uncertainty due to weather variation and climate change is one of

the main reasons for unexpected actual energy performance. Under the changing climate, building's energy performance is expected to change significantly in the northern climates, including Canada.

The current thesis mainly aims to address the two aforementioned issues with novel approaches:

1. Machine learning were deployed to extract the feature importance of the weather parameters in order to assign non-universal weighting factors straightly proportional to their impacts on energy performance of buildings. Weather files constructed with these systematically assigned weighting factors are climatic location and building type dependent. The newly constructed typical meteorological year weather files were applied to two different climatic locations to investigate the representativeness of these new weather files as compared to existing weather files and historical weather data of actual years. The representativeness was indicated in terms of the deviation in predicted energy performance of buildings between using the typical meteorological year weather file and actual historical weather data. The results indicated that typical meteorological year weather file based on the novel approach offers better prediction (with statistical significance) on energy performance for climatic locations with wider temperature range. As a result, the suggested method avoids potential under/oversizing of equipment and promotes energy conservation.

2. General circulation model (GCM) data considering various climate change scenarios based on socio-economic, population, land use, technology, and policies are used to provide information about future climatic condition. However, there are two primary challenges in application of data for building simulation:

- i. Bias in the models: considerable deviation can be found when the historical GCM data is compared to station observed weather data.
- ii. Inadequate resolution: GCM data has daily temporal resolution rather than the hourly resolution required in building energy simulation.

In order to use this data for simulation purposes and better predict future building performance, further processing is conducted. A statistical bias-correction technique, known as the quantile-quantile method, is applied to remove the bias in the data in order to adapt GCMs to a specific location. The study then uses a hybrid classification-regression (K-Nearest Neighbour – Random

Forest) machine learning algorithm to downscale the bias-corrected GCM data to generate future weather data at an hourly resolution for building energy simulation. In this case, the hybrid model is structured as a combined model, where a classification model serves as the main model together with an auxiliary regression model for cases when data is beyond the range of observed values. The proposed workflow uses observed weather data to determine similar weather patterns from historical data and uses it to generate future weather data, contrary to previous studies, which use artificially generated data. However, in cases where the future GCM data showed temperatures ranging outside of the observed data, the study applied a trained regression model to generate hourly weather data. The current study suggests a workflow that can be applied to global and regional models data to generate future weather files year by year for building simulation under various scenarios and, consequently, extreme weather characteristics are preserved for extreme or reliability analysis and design optimization.

In addition, a novel method is introduced to find building design solutions under uncertainty of weather variation and climate change. The design options are architectural and envelop features at different levels. A full factorial design of experiment is used for large-scale simulations and training deep neural network surrogate models to assess energy performance of design alternatives under multiple future years under various climate change scenarios. The method with application of a novel performance indicator is applied to explore design space and find the design solutions that most probably contribute to meet building energy performance targets over the project's lifespan.

This workflow takes into account the effect of weather variation under various climate change scenarios and suggests several design solutions that can be offered to stakeholders, architects, engineers, and third-parties including insurance companies. This way, design alternatives can be compared, and designs with a higher probability of success can be selected as a final solution. In addition, policy-makers can use the results and the suggested workflow to adopt and update national and provincial building energy codes such as National Energy Codes of Canada for Buildings (NECB) in line with the national policies following the Paris climate change agreement.

ACKNOWLEDGMENTS

I would like to convey appreciation and sincere gratitude to those who supported and encouraged me throughout the completion of this research. I acknowledge the outstanding leadership, guidance, understanding and supports received by my supervisor, Dr. Bruno Lee. His combined expertise in Building Energy Performance and Design, Large-Scale Simulation, and System Engineering, is exemplary and without him, this thesis could not be started and completed.

I would also like to thank the committee members, Dr. Andreas Athienitis, Dr. Ali Nazemi, and Dr. Dongyu Qiu for their valuable advice and remarks which directed me through developing the research objectives and findings.

Finally, I would like to extend my heartfelt gratitude to my family, my parents, Mehdi and Farkhondeh, my elder brother Dr. Amin, and my lovely late brother, Dr. Emad's soul who left us too early, and my fiancé, Shima whose sincere love and infinite patience encouraged me during my studies.

LIST OF PUBLICATIONS

This dissertation is formed based on scientific papers written to meet the Ph.D. project objectives set during the research to advance scientific and industrial goals. Most of the details of this study are comprised of the papers below and therefore the submitted manuscript should be aimed only as a summary of the overall research. This study is the outcome of my research as a Ph.D. student at the Department of Building, Civil, and Environmental Engineering, Concordia University, Montreal, Canada, and has resulted in 6 journal papers and one conference paper listed below:

Journal papers:

Paper 1 M. Hosseini, B. Lee, S. Vakilinia, Energy performance of cool roofs under the impact of actual weather data, *Energy and Buildings*, Volume 145, 2017, Pages 284-292

Paper 2 M. Hosseini, A. Bigtashi, B. Lee, Evaluating the applicability of typical meteorological year under different building designs and climate conditions, *Journal of Urban Climate*, under review

Paper 3 M. Hosseini, A. Bigtashi, B. Lee, A systematic approach in constructing typical meteorological year weather files using machine learning, *Energy and Buildings*, Volume 226, 2020, 110375

Paper 4 M. Hosseini, F. Tardy, B. Lee, Cooling and heating energy performance of a building with a variety of roof designs; the effects of future weather data in a cold climate, *Journal of Building Engineering*, Volume 17, 2018, Pages 107-114

Paper 5 M. Hosseini, A. Bigtashi, B. Lee, Generating future weather files under climate change scenarios to support building energy simulation - a machine learning approach, *Energy and Buildings*, Volume 230, 2020, 110543

Paper 6 M. Hosseini, A. Bigtashi, B. Lee, A probabilistic method toward resiliency of building energy performance under the effect of climate change - simulation and machine learning approach, to be submitted to *Journal of Building and Environment*, 2021

Conference paper:

Paper 1 M. Hosseini, B. Lee, A probabilistic approach toward building energy performance design: a case study of roof design under uncertainty of weather, the 15th international conference of Building Simulation, San Francisco, August 2017

LIST OF CONTENTS

LIST OF PUBLICATIONS	vii
Chapter 1. Introduction.....	1
1.1 Problem statement.....	1
1.1 Objectives.....	3
1.2 Outlines of the thesis.....	4
Chapter 2. Overview on current practice in building energy simulation; consideration of climate and climate change.....	6
2.1 Building energy simulation approaches.....	6
2.2 Adopting typical meteorological year data (TMY) to consider climate condition	7
2.3 Considering climate change in building simulation.....	8
2.3.1 Dynamical downscaling.....	9
2.3.2 Statistical downscaling.....	10
2.4 Summary and research direction	11
2.4.1 Future climate change consideration	11
2.4.2 Downscaling techniques	12
2.4.3 Design solution to mitigate the effect of climate change.....	12
2.5 References	13
Chapter 3. The effect of natural weather variation on building energy performance under various envelop Design.....	17
3.1 Introduction	17
3.1.1 TMY vs. AMY.....	18
3.2 Methodology	19
3.2.1 Weather data	19
3.2.2 Weather file creator.....	20
3.2.3 Normalized root-mean-square deviation.....	20

3.2.4	Case study	21
3.3	Results and discussion.....	23
3.3.1	Effect on heating demand	26
3.3.2	Effect on cooling demand.....	26
3.3.3	Effect on total demand.....	27
3.3.4	Overall effect	28
3.4	Conclusion.....	31
3.5	References	31
Chapter 4. the Applicability of Canadian Weather Year for energy calculation (CWEC); the effect of building design and climate.....		
		33
4.1	Introduction	33
4.2	METHODOLOGY	37
4.2.1	Case Study	37
4.2.1.1	Weather condition	37
4.3	Results and discussion.....	41
4.3.1	The effect of climate (weather).....	41
4.3.2	The effect of building type.....	46
4.3.3	The effect of design	47
4.4	Conclusion.....	53
4.5	References	54
Chapter 5. A systematic approach in constructing typical meteorological year weather files using machine learning		
		58
5.1	Introduction	58
5.2	Methodology	63
5.2.1	Weighting factors.....	63
5.2.2	Application of Random Forest: feature importance extraction.....	70

5.3	Results and Discussion.....	73
5.3.1	Models performance	73
5.3.2	Features importance: systematic approach to defining weighting factors	78
5.3.3	Approaches performance	81
5.4	Conclusion.....	83
5.5	References	84
Chapter 6.	A priliminary study of climate change effect on building energy performance	86
6.1	Introduction	86
6.1.1	Climate change and its impact on building energy performance.....	87
6.1.2	Climate change prediction models and their implication on building energy simulation	88
6.1.3	Adapting building design processes to adapt to climate change.....	89
6.1.4	Effects of highly reflective roofs (cool roofs) on buildings.....	90
6.2	Methodology	93
6.2.1	Adaptation of a climate change model to a base case building	93
6.2.2	Future hourly weather generation	95
6.3	Results	95
6.4	Conclusion.....	97
6.5	References	98
Chapter 7.	Generating Future Weather Files under Climate Change Scenarios to Support Building Energy Simulation — a Machine Learning Approach	101
7.1.	Introduction	101
7.1.1	Future climate change data and challenges relating to building energy simulation application	103
7.1.2	Previous approaches to address the challenges.....	104
7.1.3	Research gap	108

7.1.4	A proposed method to use weather classification downscaling for building simulation	109
7.2	Objective and organization of the study.....	110
7.3	Methodology	111
7.3.1	Weather parameters selection	113
7.3.2	Bias-Correction.....	115
7.3.3	Spatio-temporal downscaling.....	121
7.4	Results and Discussion.....	135
7.5	Conclusion.....	141
7.6	Remark	142
7.7	References	142
Chapter 8.	Design Solutions under Weather Variation and Climate Change.....	147
8.1	Introduction:.....	147
8.1.1	System Engineering designs to mitigate climate change	149
8.1.2	Objective and organization of the study	152
8.2	Methodology	153
8.2.1	Observing energy performance of a baseline over time	153
8.2.2	Design of Experiments.....	155
8.2.3	Data preprocessing.....	156
8.2.4	Model	159
8.2.5	Performance indicator.....	164
8.2.6	Workflow summary	165
8.3	Results	166
8.3.1	Design space exploration	167
8.4	Discussion	172

8.4.1	Risk definition.....	173
8.4.2	Risk treatment.....	173
8.4.3	Risk mitigation.....	173
8.4.4	Risk reduction.....	173
8.4.5	Risk transfer.....	174
8.4.6	Risk acceptance.....	174
8.5	Contribution.....	174
8.6	References.....	174
Chapter 9.	Conclusion and Future Work.....	179
9.1	Conclusion.....	179
9.2	Future work.....	180

List of Figures

Figure 3.1: Heating, cooling, and total demand of the prototypical building, 30-years actual weather data V.S. CWEC for two designs. “ins” and “ref” represent insulation value and reflectance respectively.....	23
Figure 3.2: Heating, cooling and total demand and normalized root-mean-square deviation of CWEC for different roof design combinations.....	25
Figure 4.1: 30 years of hourly global horizontal solar Irradiance for Montreal and Vancouver from 1961 to 1989	38
Figure 4.2: Montreal, small office with small windows	42
Figure 4.3: Montreal, small office with large windows.....	43
Figure 4.4: Vancouver, small office with small windows	43
Figure 4.5: Vancouver, small office with large windows.....	44
Figure 4.6: Montreal, large office with small windows.....	44
Figure 4.7: Montreal, large office with large windows	45
Figure 4.8: Vancouver, large office with small windows.....	45
Figure 4.9: Vancouver, large office with large windows.....	46
Figure 4.10: Percent difference cumulative distribution function for Montreal and Vancouver heating.....	50
Figure 4.11: Percent difference cumulative distribution function for Montreal and Vancouver cooling.....	51
Figure 5.1: The distributions of 30 years historical weather data together with corresponding energy demands of the small office building in Montreal; the bottom figures are two samples of the above figures for better resolution.	66
Figure 5.2: The distributions of 30 years historical weather data together with corresponding energy demands of the small office building in Vancouver.	66
Figure 5.3: A tree of Random Forest; the blue scatters show the data; the vertical dash are the split points of the variables, and the horizontal dash lines show the prediction of total energy demand after splitting. The number of depths are reduced to 3 for better presentation.	68
Figure 5.4: Contrasting the steps taken under the current Sandia method with those of the proposed method and indicating where the additional steps of random forest modeling takes place and supports the systematic weighting factors assignment.	70

Figure 5.5: Results of trained models in comparison to the simulation results for cooling-based approach for Montreal; the bottom zooms into the first year (1960) of the top (approach 1).	74
Figure 5.6: Results of trained models in comparison to the simulation results for heating-based approaches for Montreal; the bottom zooms into the first year (1960) of the top (approach 2)...	75
Figure 5.7: Results of trained models in comparison to the simulation results for total-based approaches for Montreal; the bottom zooms into the first year (1960) of the top (approach 3)...	76
Figure 5.8: Feature importance extraction from Random Forest Regression (approaches 1 to 3)	79
Figure 5.9: Simulation results based on actual meteorological years and the typical meteorological year weather files made up with weighting factors based on different approaches for Montreal and Vancouver.	81
Figure 6.1: Cooling, heating and total energy for a retail building using CWEC and horizon weather data	97
Figure 7.1: The GFDL-ESM2M, near-surface daily air temperature for 2020-07-21 and 2049-07-21 under RCP 8.5, left figures show the location of Montréal–Pierre Elliott Trudeau airport...	112
Figure 7.2: Correlation coefficient for parameters; darker colors represent stronger linear correlation.	115
Figure 7.3: Quantile-Quantile bias-correction method; the future model data are statistically corrected according to historical deviation from observation.....	117
Figure 7.4: The validation period historical GCM data and historical observed data from 1976 to 2005 (left) [39] versus original GCM future data and bias-corrected GCM data from 2020 to 2049 (right).	118
Figure 7.5: The first year (1976) of the validation period historical GCM data and historical observed data (top) versus the first year (2020) of original GCM future data and bias-corrected GCM data (bottom).....	119
Figure 7.6: Yearly-averaged of original future GCM data versus the bias-corrected data for the period of 2020 to 2049.	120
Figure 7.7: FS statistics for temperature, the absolute difference between the CDFs of the predictor and predictand over n bin points (30) is calculated (step ii).	123
Figure 7.8: The feature importance (weather parameter weighting factors) extraction from Random Forest Regression to be used for KNN classification algorithm.....	125

Figure 7.9: CDF of the future weather data in comparison to the historical observed weather data for July 2020 (top) and March 2021 (bottom). 128

Figure 7.10: Train, test, and predicted result of the Random Forest regression model; hourly data (top left), daily-averaged data (bottom left), monthly-averaged data (top right), and yearly-averaged (bottom right)..... 130

Figure 7.11: The last 5 months of the training period and the first 8 months of the test period with hourly resolution (top) and daily resolution (bottom)..... 131

Figure 7.12: Temperature values selected by classification (March 1984) and daily-averaged of corrected values using regression model for March 2021. Top: daily-averaged of all days of the month; bottom: CDF of the daily values..... 132

Figure 7.13: Flowchart of the workflow to process the GCM data to be used for building energy simulation; the Random Forest algorithm is used twice; once in the feature importance sub-process (shown in pink) with single target of total demand; and once as a regression model (shown in green) with multi-target of 24 hour-temperature and input of daily-averaged temperature. 133

Figure 7.14: The daily-averaged temperature created under RCP 8.5 with three thresholds and the corresponding bias-corrected GCM temperature for the year 2020 (discontinuity at the end of February is due to the leap year). 137

Figure 7.15: Yearly-averaged temperature of GCM bias-corrected data and temperature disaggregated with the three levels of threshold under the four RCPs. 138

Figure 7.16: Annual cooling and heating energy demand of the building simulated with the weather files generated under the four RCPs and based on three thresholds. 139

Figure 8.1: Annual cooling and heating of the small office building under historical, CWEC, and future weather years. 154

Figure 8.2: Annual cooling and heating of the small office building under historical, CWEC, and future weather years; orange and blue circles indicate the weather years selected for training of heating and cooling models respectively; the pink and green circles indicate the weather year for test of the cooling and heating models respectively. 157

Figure 8.3: The architecture of the deep neural network model with two hidden layers; there are 187 neurons in the first hidden layer and 93 neurons in the second hidden layer. Note: the number of neurons in the layers are reduced in the figure for better presentation..... 160

Figure 8.4: History of models trained for heating demands over epochs. 161

Figure 8.5: History of models trained for cooling demands over epochs.....	162
Figure 8.6: Predicted annual cooling demands of the trained models with weather year 2014 (test data) in comparison to the simulation results; top: the whole 3600 designs, bottom: the first about 700 designs.....	163
Figure 8.7: Predicted annual heating demands of the trained models with weather year 2014 (test data) in comparison to the simulation results; top: the whole 3600 designs, bottom: the first about 700 designs.....	164
Figure 8.8: The workflow of the methodology.....	166
Figure 8.9: Annual cooling and heating demand under four climate change scenarios; under each scenario; 108000 points represent the demand of 3600 design combination for 30 future years; the 3600 orange points represent the same design combinations under CWEC weather condition.	167
Figure 8.10: Probability of meeting targets (functional requirement) for cooling and heating for various designs. The blue points represent the probabilities under each RCP and the purple points show the performance of designs meeting the targets with more than 60% under all the four scenarios.....	169
Figure 8.11: Probability of meeting targets (functional requirement) for cooling and heating for various designs. The blue points represent the probabilities under each RCP and the purple points show the performance of designs meeting the targets with more than 80% under all the four scenarios.....	170
Figure 8.12: The information content of designs that meet cooling demand target of 30 kWh/m ² and heating demand of 25 kWh/m ² at least in one future year in each scenario.....	171

List of Tables

Table 2.1: Explanation of each scenario in case of future socio economic, and biotic condition [20]	8
Table 3.1: Corresponding years of CWEC months for Montreal airport weather station	19
Table 3.2: Building characteristics of the case study	22
Table 3.3: Statistics of energy demand of the building in 30-years V.S. CWEC for six different roof design combinations	29
Table 4.1: Heating and Cooling Degree Days of the two studied cities	38
Table 4.2: Properties of the studied buildings	39
Table 4.3: values of design parameters used in the study	40
Table 5.1: Weightings factors of weather parameters in previous studies	60
Table 5.2: List of terms used in figures, tables, and texts	61
Table 5.3: Performance of the trained models for month-by-month cooling or heating energy based and month-by-month total energy based in Montreal and Vancouver (approaches 4 and 5)	77
Table 5.4: Weighting factors used resulted from MBM cooling or heating energy based, and MBM total energy based model for Montreal and Vancouver	80
Table 5.5: Root mean square error calculated for CWEC and the suggested approaches. RMSE reduction shows the reduction of root mean square error in comparison to CWEC method. Minus values show penalty; the numbers of percentages are rounded to zero digit. Blue and pink sections represent RMSE for cooling and heating respectively.	82
Table 6.1: Building characteristics used for simulation	94
Table 7.1: Nomenclature for phrases used in figures, tables, and texts	103
Table 7.2 Overview of the workflow	113
Table 7.3: accuracy of the regression model on test data	130
Table 7.4: The historical years from which hourly data are selected month-by-month for the future year of 2020; the months in which temperature is corrected by the values of regression model are shown with ‘R’	136
Table 8.1: Design parameters and levels selected for the case study	156
Table 8.2: The performance of the eight trained models in predicting the cooling and heating for 3600 design with the weather year 2014 (test data)	162

Table 8.3: The 12 designs meeting the targets with more than 80% probability under all the four RCP scenarios 171

Table of Acronyms

Abbreviation	Explanation
HVAC	Heating, Ventilation, and Air Conditioning
LEED	Leadership in Energy and Environmental Design
TMY	Typical Meteorological Year
RCP	Radiative Concentration Pathway
IPCC	Intergovernmental Panel on Climate Change
GCM	General Circulation Model
MIROC	Model for Interdisciplinary Research on Climate
WRF	Weather Research and Forecasting
RCM	Regional Climate Model
CDF	cumulative distribution functions
CWEC	Canadian Weather Year for Energy Calculation

Chapter 1. INTRODUCTION

1.1 Problem statement

Building energy simulation is commonly used to evaluate the energy performance of the buildings; the process is conducted to either support decision at the design stage or to quantify the possible energy savings when energy efficiency strategies are supposed to be assessed to retrofit an existing building. However, in many cases of energy-efficient buildings, the predicted performance deviated from measured data considerably. In fact, such deviation has been well-documented for quite many LEED certified building [1-4].

A major reason for such discrepancy between predicted outcome and the actual energy consumption is due to uncertainty in the simulation inputs as suggested by Hopfe et al. [5]. For building energy simulation, uncertainty is resulted from three sources of physical uncertainty, design uncertainty, and scenario uncertainty including internal heat gain and weather condition [5,6]. For instance, internal gains such as occupancy and plug load or from outside the environment and weather condition make up the major energy consumption of a building. Therefore, the natural variation of occupancy and weather condition is ignored either by simplification or assumptions during the simulations which might lead to a considerable actual performance deviation from what it was expected at the simulation stage. One source of input uncertainty is weather data representing the climate condition. In order to predict the long-term performance of the buildings with energy simulation, modellers commonly use a single Typical Meteorological Year (TMY) weather data file which supposedly represents the climatic conditions. The single weather year file is composed of hourly resolution data from the most 12 representative calendar months of 30 years which are selected based on statistical similarity to long-term weather daily-averaged data. These weather files are synthetically constructed on historical weather data over a long period of time for an array of weather parameters, such as solar radiation, temperature, wind speed and others. The statistical procedure to construct the weather files depends on the weights assigned to these weather parameters. Under current practice, these weighting factors are universally assigned regardless of climatic locations nor the building application. This approach leads to energy performance predictions that deviate from the long-term averages. Nevertheless, the single weather file ignores the variation in building energy performance resulted from natural weather variation.

The uncertainty due to natural weather condition variation might become even more critical when the long-term superimposed effect driven by human and anthropogenic factors are added to natural variation. The greenhouse gas concentrations and aerosols loadings are the two factors contributing to this long-term effect, also known as climate change. From the mid-twenties century, it is observed that human activities, majorly energy, has been the dominant cause of this change [7-8]. Over the same period, higher latitudes including Canada have been effected greater compared to other regions. Generally, historical data show that Canada experienced a warming rate as much as twice the global mean, and this has been triple in northern Canada [9]. According to the information presented by The Fifth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC) Working Group I [8], over the period of 1880 to 2012 the global temperature has increased about 0.85 °C. A Canadian study calculated an annual temperature rise for 16 major cities in Canada over the period of 1900 to 2013. For Montreal, QC, the average annual temperature rose about 2 °C which 1.4 °C occurred in summer and 2.7°C in winter [9]. The main reason for such major effect is supposedly the global system of wind transporting warm air from the highly solar-heated equator to the higher latitudes which is called the general circulation of the atmosphere [8].

Typical meteorological year data (TMY) is widely used for assessment of building energy performance; the main problem with TMY is that it cannot be used for extreme condition analysis and varying weather conditions. Knowing that building energy performance considerably influenced by weather condition, building energy performance can be significantly degraded under the effect of climate change. Under the effect of climate change, a major challenge is that the decision-makers including designers, investors, and engineers of the project cannot select the best design solution because they cannot foresee the consequence of uncertain events that affect the payoff of each design solution. In fact for the owners of the project, an insight into the success of the sustainable project is quite necessary which is ignored in typical building energy simulation. On the other hand, it is suggested to use design-builder integrated design method of delivery for high-performance energy buildings as a strategy to reduce the capital cost while to make sure the target level of energy efficiency will be met [10]. In this method, the builders collaborate with designers from the early stage and the design-builder team is legally responsible for the actual performance of the high-performance project. Therefore, every possible design solution together

with the probability of the success of the project should be included in the analysis of the design stage.

Although the climatic condition variation is not reducible, the building design can be improved such that it meets the design goals with certain level of confidence which itself requires introducing additional performance indicators to help decision-makers to properly choose a design solution.

1.1 Objectives

The objective of this research is to develop a workflow to support decision-making for the design of a sustainable building accounting for natural weather variation and the long-term effect of climate change. This workflow should be practical and sensible to decision-makers including building energy modellers, building engineers, architects, and investors. It should provide a comprehensive information about the potential design options, processing future weather data, and the outcome of each design solution together with the suggested design solution. Meanwhile, machine learning can be used to speed up the large-scale simulation.

The workflow employs a novel approach considering the variation in weather to minimize the deviation of actual building energy performance from what is expected at the design stage. This approach provides the designers a means to differentiate designs not just on the predicted performance but on the probability in achieving the performance target. The proposed workflow should have the following features:

- **Considers the effect of future climate change effect:**

The future weather condition is represented as scenarios which are the results derived from a variety of parameters such as the economy, population, technology, and corresponding emission scenarios. A range of future climate change effect with acceptable resolution should be considered, at least for the next 30 years.

- **Considers a large scale range of building design options:**

As a strategy to mitigate the effect of climate change, a large range of design options including the architectural properties such as window-to-wall ratio and building enclosure parameters such as insulation level, solar reflectance etc. should be considered in the workflow.

- **Provides probability of long-term project success:**

The probability of success associated with each design combination over the lifespan of the building should be presented.

- **Suggests the most appropriate design solutions:**

From the energy performance point of view, the most acceptable range of design options should be suggested for final decision making.

1.2 Outlines of the thesis

This thesis reports the author's effort to develop a novel workflow to properly design a sustainable building under a long-term outdoor environment condition, namely, the natural weather variation and climate change.

The second chapter is an overview of the current building simulation approaches and the methods to process climate change data for building energy simulation.

In the third and fourth chapters, buildings' energy performance with various design combinations are simulated with actual historical weather years to assess the variation of performance under natural weather variations. The simulation results are also compared with the simulations with Canadian weather year for energy calculation (CWEC) to investigate deviation of the CWEC from actual years and CWEC applicability to represent long-term energy performance.

In the fifth chapter, a novel machine learning-based method is introduced to reduce the deviation of CWEC from the actual year's energy performance. The study deploys machine learning algorithms to systematically find non-universal weather parameters weighting factors required in constructing a typical meteorological year.

In chapter six, a preliminary study of a building with multiple design combinations under an extreme climate change scenario is conducted to investigate how building enclosure design parameters can potentially mitigate the impact of weather variation and climate change. Since this is a preliminary study, a simple downscaling method is used to construct future climate change weather years.

In chapter seven, a novel machine learning-based methodology is introduced to process climate change data for building energy simulation. The study deploys machine learning algorithms to downscale climate change data to be used for building energy simulation. The method captures both the intensity and frequency of climate change data; therefore, it can potentially be also used for extreme condition analysis.

In the eighth chapter, a novel method is introduced to improve the building energy performance of building under weather variation and the effect of climate change. The study uses simulation and data-driven models to enable architects, engineers, and decision-makers to find the design solutions to meet the targets with the most probability of success under climate change.

Finally, chapter nine concludes the thesis report and explains the future work of the research.

Chapter 2. OVERVIEW ON CURRENT PRACTICE IN BUILDING ENERGY SIMULATION; CONSIDERATION OF CLIMATE AND CLIMATE CHANGE

Simulation programs use building characteristics and environment conditions data as input to predict building energy performance. These inputs include envelope and architecture properties, operation schedules and occupancy patterns, as well as outdoor climate conditions. Some inputs such as occupancy and climate conditions are stochastic but are used in a deterministic manner with assumptions and simplification; an example is using a single weather year data representing a climate condition. Depending on the level of energy modeller's confidence and whether or not he/she wants to consider the uncertain parameters, different approaches can be used to conduct simulations which are explained in section 2.1. One significant source of uncertainty in input values is the outdoor environment including climate condition and climate change which can be addressed through introducing them as weather data inputs to simulations; the latter are explained further in sections 2.2 and 2.3.

2.1 Building energy simulation approaches

Generally, in engineering numerical simulations two approaches are applied:

- 1. Deterministic approach:** Deterministic models expect single outcome where the inputs are assigned with single values to model, and the same outcome is expected with the same set of input values as the model is executed again.
- 2. Probabilistic approach:** Probabilistic models take into account the uncertainty of models by defining a random experiment in which outcome varies in an unpredictable manner when the experiment is repeated for the same design. Probabilistic models result in a stochastic outcome where inputs are assigned with probabilistic functions.

Deterministic models are simplified models which ignore the inherent variation and uncertainty of input parameters and outcome might not reflect the uncertain reality. The advantage of this approach is a faster and more straightforward performance. However, in building energy simulation, most inputs affecting energy performance are stochastic in nature and stochastic outcomes should be expected. Probabilistic approach in conducting building simulation can be used in different applications including the hygrothermal performance of building enclosure [11] and in building energy performance at design [12-14], calibration [15-16], and retrofit [17-18]

stages. The probabilistic approach is time-consuming as it is required to be repeated so many times and preprocessing and post-processing might also be required. However, when the case study is a high-performance building, the decision makers might be interested to see more reliable designs. The goal of a probabilistic approach in evaluating building energy performance is to minimize uncertainty in predicted performance by offering designers and building owners the probability of occurrence at each performance level. In other words, the probabilistic approach enables designers to consider the probability that the buildings will perform at the very least the predicted performance at the design stage if not more in actual operation. One important source of uncertainty in building energy performance is the uncertainty due to the natural weather variation and climate change that are considered in the system as weather year files as input in simulations.

2.2 Adopting typical meteorological year data (TMY) to consider climate condition

A typical meteorological year weather file, supposedly representing the climatic conditions, is commonly used as an input for building energy simulation to predict the long-term performance of the buildings. These weather files are synthetically constructed on historical weather data over a long period of time (usually 30 years) for an array of weather parameters, such as solar radiation, temperature, wind speed, and others. The weather year consists of 12 representative months of different years from historical hourly data based on statistical resemblance. The statistical procedure to construct the weather files depends on the weights assigned to these weather parameters. Under current practice, these weighting factors are assigned based on experts' judgments, regardless of climatic locations. Previous literature indicates that such an assignment approach leads to energy performance predictions that deviate from the long-term averages.

Using data analytic techniques, non-universal weighting factors can be achieved in order to reduce the resulted energy performance deviations from expected values for different applications which is addressed in chapter 5 of this study.

Other than the typical meteorological year data that represent the historical weather condition, future weather files can also be introduced to simulation programs in the format of weather files to consider future climatic conditions.

2.3 Considering climate change in building simulation

The concentration of greenhouse gases in the atmosphere reduces the amount of energy radiating back to space. In the last two centuries, it is generally believed that human activities have contributed to an increase in the production of greenhouse gases emission leading to global temperature rise which is also known as global warming or climate change [7]. Although some natural factors such as volcanic eruption partially release carbon dioxide to the atmosphere, they play only a minor role in climate change. Therefore, anthropogenic factors are expected to continue making up a major reason of this effect in the future [8].

These activities are projected by possible future scenarios which are related to social, economic, and technological states and corresponding radiative forcing (the difference between absorbed insolation by earth and energy radiated back to space in the form of infrared radiation). In the previous versions of Intergovernmental Panel on Climate Change (working group 3 and 4), scenarios were presented as A1, A2, B1, and B2 [19] which were replaced with Representative Concentration Pathways of RCP 2.6 W/m², RCP 4.5 W/m², RCP 6 W/m² and RCP 8.5 W/m² in IPCC working group 5 [8]. The numbers after RCPs come from the relative difference of radiative forcing at the end of 2100 compared to just before the industrial revolution (1750) [8]. Table 2.1 explains the socio economic, and biotic condition under each scenario [20].

Table 2.1: Explanation of each scenario in case of future socio economic, and biotic condition [20]

RCP	Explanation
2.6	Global CO ₂ emissions peaks by 2020 and then reduce to about zero by 2080. Concentrations in the atmosphere peaks at 440 ppm in mid-century and then starts reducing. World population peaks to 9 billion in the mid-century. Global economy highly grows. Oil consumption decreases whereas other fossil fuels consumption increases. Bio fuel consumption is high and renewable energy increases but still remain low. Cropping area increases faster and animal husbandry become more intensive. Deforestation continues at the current rate.
4.5	Global CO ₂ emissions peaks by 2050 with 50% higher than that it was at the year 2000 reducing continuously for 30 years and then remains constant at about half of what it was at the year 2000. The trend of concentration continues till 2070 at about 520 ppm and continues with lower slope afterwards. World population and economic moderately grows with slightly lower than that in RCP 2.6. Energy consumption would be higher than under RCP 2.6 and oil consumption remain constant till the end of century whereas use of nuclear power and renewables increase. Cropping area considerably decreases whereas reforestation increases.
6	Global CO ₂ emissions double by 2060 then significantly drops but still above the current level. The concentration continues increasing, reaching 620 ppm by 2100. World population reaches 10 billion. Compared to other three scenarios, the economy growth is lower. Energy consumption reaches a peak at 2060 then reduces to the level under RCP 2.6 until 2100. Oil consumption keeps at high level while biofuel and nuclear power will be less than other scenarios. Cropping area remains at the current trend and natural vegetation would the same as under RCP 4.5.

8.5

Global CO₂ emission continuously increases to reach 30 gigatonnes of carbon in the year 2100 in comparison to 8 gigatonnes in 2000.
The concentration rapidly increases to 950 ppm by 2100.
World population reaches 12 billion by 2100.
The economy growth would be similar to that under RCP 6 however, the incomes in developing countries will be significantly lower.
Energy consumption continuously increases to 3 times the current level.
Oil consumption significantly increases until the year 2070 before a deep decline afterwards. Coal will be highly used to provide the required inclined energy consumption.
Cropping area increases while the deforestation continues.

General circulation models mathematically simulate atmospheric, oceanic, and biotic interactions and combine them with radiative forcing scenarios to evaluate the future climates. GCM models consist of grid cells resulted from latitude and longitudinal divisions, in which, the meteorological data are calculated [21]. General circulation models outputs have a coarse resolution namely, a spatial resolution of larger than 100 km * 100 km and usually a temporal resolution of daily-average. Whereas, hourly meteorological data is required for building energy simulation. Therefore, further process, called downscaling, is required to have finer resolution data. Depending on the availability of resources and expertise, two main approaches are used for downscaling, dynamical, and statistical.

2.3.1 Dynamical downscaling

Dynamical downscaling requires an additional computationally intensive physical process which is conducted for a specific region using regional or local climate models. This method is applied in literature to project climate change on future weather condition used for building energy simulation. Kikumoto et al. [22] used GCM data as initial and boundary condition for regional climate models namely, Model for Interdisciplinary Research on Climate (MIROC) and the Weather Research and Forecasting (WRF) to construct the future weather data from 1931 to 1935 for a Japanese climate. Their energy simulation for a detached residential building showed a 15% increase in building sensible load compared to the year 2007.

Burger et al. [23] assessed the effect of climate change on cooling and heating demand of an office building attributed to three epochs namely, built: before World War I, after World War II, and from 2000 onward in Vienna, Austria. They used REMO UBA regional climate model to dynamically downscale A1B climate change scenarios with the resolution of about 10 km * 10 km. They evaluated the results for two time frames of 2011–2040 and 2036–2065. In one case,

their result showed a 41% increase in cooling and a 56% decrease in heating load compared to the years 1961-1990.

2.3.2 Statistical downscaling

This approach relies on the availability of observed weather data and can be used for all scenarios of climate change while it provides point-scale climatic variables; in addition, it can be applied to regional and /or global models. In statistical approach, a critical assumption is stationary which means that, while the climate changes, the statistical relations among the meteorological parameters data remain constant over time [24-25]. Statistical downscaling is categorized into three main groups of linear methods, weather classification, and weather generators. Linear methods including the delta method (also known as morphing) are easy to use and are widely used in previous studies. In morphing, a changing factor is directly applied to the observed data to calculate future data. This method is used to predict the future climate condition of American climate [26-29] Canadian climate [30-31], Swedish climate [33-33], Spanish climate [34], Italian climate [35], and British climate [36].

Weather classification relates a class of future weather pattern to a local observed weather and the future weather data are synoptically selected from the observed weather data. In this method, the effect of climate change is estimated by evaluation of the frequency of the change of weather pattern parameters from GCM output; moreover, it is assumed that the characteristics of each class will remain constant [37-38]. One challenging issue of weather classification may be dealing with unprecedented increasing weather variables such as temperature which is quite significant in building energy performance. This issue is addressed in chapter 7.

Tian and de Wilde [39] investigated the use of climate change projection for building energy analysis for a British climate using statistic reduction (Finklestein-Schafer) applied to the output of a weather generator for three different climate change scenarios. For the statistical method, the CDF of the temperature and radiation with equal weight was used to project climate change for the 2020s, 2050s, and 2080s. They used a linear meta-model developed with historical weather as a tool to apply energy efficiency measures to reduce the effect of climate change.

Weather generators are used for temporal downscaling; these statistical models generate numerous possible time-series weather variables using the same statistical characteristics such as correlation

coefficient of several-years historical data applied to the GCM models output. In addition, the relationship between predictor and predictand is critically assumed to be stationary from the time point of view [40]. In fact, the data generated from the generators are different from the observed data and only keep the statistical characteristics of the observed data. Among the weather generators, only a few are able to consider the relationship between the weather elements when multiple variables are predicted [25]. In addition, these weather generators are able to generate only a few weather parameters such as temperature and solar radiation [41].

2.4 Summary and research direction

Building energy simulation which has been used for design decision making of new buildings, and energy performance assessment of existing buildings requires hourly weather data; the current practice is to use a single typical meteorological data which is statistically selected from the historical data in simulations. This approach, although provides a faster process in decision making, doesn't capture the variation of energy performance resulted from natural variation of weather. Moreover, because of the effect of climate change, the actual energy performance of the building might significantly different from what is expected if historical data is used at design stage; therefore, future climate change weather data should be used for energy simulation especially for the high-performance buildings. By reviewing the previous literature related to the subject of climate change and building energy performance, it is found that most of the literature targets were quantifying the effect of climate change on building energy performance of existing buildings; therefore, their methodologies focused on quantification of long-term effect of climate change on building energy performance in the form of long time frames. Some conclusions are summarized below.

2.4.1 Future climate change consideration

Most of the studies focused on only one future climate change scenario (the most extreme global warming) while, other scenarios are ignored. This is partially due to the lack of weather data because most GCM models provide only one or two climate change scenarios. In addition, all the previous literature used a single weather year as a representative of a long period of time (usually two to three decades); although single representative year provides a holistic vision of energy performance, it doesn't capture the natural variation of that which might be quite significant.

Using GCMs providing data for all scenarios and considering natural variation can be promising for better building design solutions.

2.4.2 Downscaling techniques

Statistical downscaling is computationally inexpensive and is able to capture the station-scale climate information therefore, it seems an efficient promising way to capture the effect of climate change for building energy simulation at the design stage of sustainable projects. However, the linear method of temporal downscaling (morphing) largely used in previous studies [26-36] although is easy to use and seem practical for long time-frame weather data construction, may not be suitable to capture extreme events; as they only add the mean future temperature rise to all the observation hourly data. This might seem appropriate when long-term frames are constructed but not suitable when a year-by-year evaluation is required; because this makes all the historical data warmer, while in reality there might be some colder winters in the future.

In addition, the complicated weather generators used in the literature [39-41] generate hundreds of possible future events for a given time period under each scenario; therefore, they are computationally expensive and this makes post-processing an extremely time-consuming stage. A novel weather classification method could be an efficient way to downscale the data with a potential for year-by-year simulation weather files creation; the subject is addressed in chapter 7. Once the future weather data are processed, multiple year-by-year simulations introduce logical randomness to the approach then, the energy performance of various design options can be assessed using a probabilistic approach based on scenarios.

2.4.3 Design solution to mitigate the effect of climate change

At the design stage of the buildings, the decision must be made under uncertain future situation. Building design solutions under climate change suggested by previous studies were highly related to the method of climate change data processing and assessment. Most previous studies didn't consider all the climate change scenarios and they didn't suggest design solution mitigating the effect in a comprehensive study. In most studies, a single representative year was used to show a couple of decades' climatic conditions. Therefore, finding the design solutions methodology was limited to optimizations without considering the fluctuations and variations resulting from the weather's natural variation. A novel method to select the design variable that meets the requirements with a high probability of success under different climate change scenarios can select

those designs that meet the energy efficiency levels most of the time. The method can potentially apply multiple constructed future climate change weather files under different scenarios and predict multiple design options' energy performance under the future climate. A machine learning algorithm can be trained on simulation output to efficiently find the design solutions that reduce the building's cooling and heating demand while keeping the performance under different climate change scenarios. The method is addressed in chapter 8 of this study.

Before the effect of climate change are taken into account, preliminary studies are conducted to evaluate the natural variation of weather on energy performance of building with various design parameters in cold climate which is explained in chapter 3.

2.5 References

- [1] M. Frankel, C. Turner “ACEEE Summer Study on Energy Efficiency in Buildings”, 2008
- [2] M.S.K. Al-Zubaidy, "A Literature Evaluation of the Energy Efficiency of Leadership in Energy and Environmental Design (LEED)-Certified Buildings." *American Journal of Civil Engineering and Architecture* 3.1, 1-7, 2015
- [3] J.H. Scofield, “Efficacy of LEED-certification in reducing energy consumption and greenhouse gas emission for large New York City office buildings”, *Energy and Buildings*. *Energy and Buildings* 67, 517–524, 2013
- [4] C. Turner, M. Frankel, “Energy Performance of LEED® for New Construction Buildings”, FINAL REPORT, New Buildings Institute, Prepared for: U.S. Green Building Council, 2008
- [5] C. Hopfe, J.L. Hensen, “Uncertainty analysis in building performance simulation for design support”, *Energy and Buildings* 43, 2798–2805, 2011
- [6] W. Yao, X. Chen, W. Luo, M. Tooren, J. Guo, “Review of uncertainty-based multidisciplinary design optimization methods for aerospace vehicles”, *Progress in Aerospace Sciences* 47, 450–479, 2011
- [7] United Nations. Sustainable Development Goals. Available from: <http://www.un.org/sustainabledevelopment/energy/>
- [8] IPCC. Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge, UK; and New York, USA: Cambridge University Press, 2013

- [9] Climate data and scenarios for Canada: Synthesis of recent observations and modelling results, Environment and Climate Change Canada, ISBN: 978-0-660-04262-6
- [10] S. Pless, P. Torcellini, J. Scheib, B. Hendron, M. Leach, “How-To Guide for Energy-Performance-Based Procurement, An Integrated Approach for Whole Building High Performance Specifications in Commercial Buildings”, NREL Report Number: TPP-5500-56705 U.S Department of Energy, 2012
- [11] G.C. Rodríguez, A.C. Andrés, F.D. Muñoz, J.M. Cejudo López, Y. Zhang, “Uncertainties and sensitivity analysis in building energy simulation using macro parameters”, *Energy and Buildings* 67, 79–87, 2013
- [12] J.M. Dussault, L. Gosselin, “Office buildings with electrochromic windows: “A sensitivity analysis of design parameters on energy performance, and thermal and visual comfort” *Energy and Buildings* 153, 50–62, 2017
- [13] E. Elbeltagi, H. Wefki, S. Abdrabou, M. Dawood, A. Ramzy, “Visualized strategy for predicting buildings energy consumption during early design stage using parametric analysis” *Journal of Building Engineering* 13, 127–136, 2017
- [14] T. Østergård, R.L. Jensen, Steffen E. Maagaard, “Early Building Design: Informed decision-making by exploring multidimensional design space using sensitivity analysis” *Energy and Buildings* 142, 8–22, 2017
- [15] F.S. Westphal, R. Lamberts, “Building simulation calibration using sensitivity analysis”, IBPSA (International Building Performance Simulation Association) conference, August 2005, Montreal, Canada
- [16] J. Sun, T.A. Reddy, “Calibration of building energy simulation programs using the analytic optimization approach”, *HVAC & R Research*, 2, 177–96, 2006
- [17] F. Ascione, N. Bianco, C. De Stasio, G.M. Mauro, G.P. Vanoli, “Artificial neural networks to predict energy performance and retrofit scenarios for any member of a building category: A novel approach” *Energy* 118, 999-1017, 2017
- [18] S. Copiello, L. Gabrielli, P. Bonifaci, “Evaluation of energy retrofit in buildings under conditions of uncertainty: The prominence of the discount rate”, *Energy* 137, 104-117, 2017
- [19] IPCC special report working group III, emission scenarios for policy makers, ISBN: 92-9169-113-5
- [20] D.P. van Vuuren, J. Edmonds, M. Kainuma, K. Riahi, A. Thomson, K. Hibbard, G.C. Hurtt, T. Kram, V. Krey, J.F. Lamarque, T. Masui, M. Meinshausen, N. Nakicenovic, S.J. Smith, S.K. Rose, The

representative concentration pathways: an overview, The representative concentration pathways: an overview. *Climatic Change*, Vol 109, Issue 1-2, 5-31

[21] R. L. Wilby, J. Troni, Y. Biot, L. Tedd, B.C. Hewitson, D.M. Smith, R.T. Sutton, “A review of climate risk information for adaptation and development planning.” *International Journal of Climatology* 29, 1193-1215, 2009

[22] H. Kikumoto, R. Ookaa, Y. Arima, T. Yamanaka, “Study on the future weather data considering the global and local climate change for building energy simulation”, *Sustainable Cities and Society* 14, 404–413, 2014

[23] T. Berger, C. Amann, H. Formayer, A. Korjenic, B. Pospischal, C. Neururer, R. Smutny, “Impacts of climate change upon cooling and heating energy demand of office buildings in Vienna, Austria”, *Energy and Buildings* 80, 517–530, 2014

[24] E. Zorita, H. von Storch, “The analog method as a simple statistical downscaling technique: comparison with more complicated methods”. *Journal of Climate* 12 (8), 2474-2489, 1999

[25] S. Trzaska, E. Schnarr, “A Review of Downscaling Methods for Climate Change Projections”, U.S. Agency for International Development (USAID), September 2014

[26] P. Shen, “Impacts of climate change on U.S. building energy use by using downscaled hourly future weather data”, *Energy and Buildings* 134, 61–70, 2017

[27] J. Huang, K.R. Gurney, “The variation of climate change impact on building energy consumption to building type and spatiotemporal scale”, *Energy* 111, 137-153, 2016

[28] A. Jiang, Y. Zhu, A. Elsafty, M. Tumeo, “Effects of Global Climate Change on Building Energy Consumption and Its Implications in Florida”, *International Journal of Construction Education and Research*, 14(1) 22-45, 2018

[29] L. Wang, X. Liu, H. Brown, “Prediction of the impacts of climate change on energy consumption for a medium-size office building with two climate models”, *Energy and Buildings* 157, 218–226, 2017

[30] A. Robert, M. Kummert, “Designing net-zero energy buildings for the future climate, not for the past”, *Building and Environment* 55, 150-158, 2012

[31] M. Hosseini, F. Tardy, B. Lee, “Cooling and heating energy performance of a building with a variety of roof designs; the effects of future weather data in a cold climate”, *Journal of Building Engineering*, 17 107-114, 2018

- [32] U. Yao, A. Tettey, A. Dadoo, L. Gustavsson, “Energy use implications of different design strategies for multi-storey residential buildings under future climates”, *Energy* 138, 846-860, 2017
- [33] V. M. Nik, “Making energy simulation easier for future climate – Synthesizing typical and extreme weather data sets out of regional climate models (RCMs)”, *Applied Energy* 177, 204–226, 2016
- [34] J.M. Rey-Hernández , C. Yousif , D. Gatt , E. Velasco-Gómez , J. San José-Alonso, F.J. Rey-Martínez, “Modelling the long-term effect of climate change on a zero energy and carbon dioxide building through energy efficiency and renewable”, *Energy & Buildings* 174, 85–96, 2018
- [35] M. Cellura, F. Guarino, S. Longo, G. Tumminia, “Climate change and the building sector: Modelling and energy implications to an office building in southern Europe”, *Energy for Sustainable Development* 45 46–65, 2018
- [36] M.F. Jentsch, A.S. Bahaj, P.A.B. James, “Climate change future proofing of buildings—Generation and assessment of building simulation weather files”, *Energy and Buildings* 40, 2148–2168, 2008
- [37] W.A.R. Brinkmann, “Modification of a correlation-based circulation pattern classification to reduce within-type variability of temperature and precipitation”. *International Journal of Climatology*, 20, 839–852, 2000
- [38] H. J. Fowler, S. Blenkinsop, and C. Tebaldi, “Linking climate change modelling to impacts studies: Recent advances in downscaling techniques for hydrological modelling”, *International Journal of Climatology*, 27, 1547–1578, 2007
- [39] W. Tian, P. de Wilde, “Thermal building simulation using the UKCP09 probabilistic climate projections”, *Journal of Building Performance Simulation* 4(2), 105-124, 2011
- [40] B.D. Lee, Y. Sun, H. Hu, G. Augenbroe, C.J.J. Paredis, “A Framework for generating stochastic meteorological years for risk-conscious design of buildings”, National Conference of IBPSA-USA, At Madison, August 2012
- [41] M. Semenov, LARS-WG stochastic weather generator. United Kingdom, Department of Computational and Systems Biology (2012), available at:
<http://resources.rothamsted.ac.uk/sites/default/files/groups/mas-models/download/LARS-WG-Manual.pdf>

Chapter 3. THE EFFECT OF NATURAL WEATHER VARIATION ON BUILDING ENERGY PERFORMANCE UNDER VARIOUS ENVELOP DESIGN.

Weather conditions account for a major source of deviation between simulation results and actual energy performance of buildings. Typical meteorological year (TMY) weather data are commonly used to evaluate building energy demand at design stage. However, such data might overestimate or underestimate the energy demand of buildings considerably depending on the building designs. Also, TMY does not capture the yearly weather variations which is important for evaluating the potential energy savings and penalties for specific energy efficient measures in the long run. This chapter aims to provide a better understanding of the natural weather variation on building energy performance and the effect of building design to reduce the corresponding performance variation of climate conditions. The study also investigates the deviation of results of simulations with Canadian weather year for energy calculation (CWEC) from the simulations with multiple actual weather years for multiple design combinations in Montreal.

This chapter was published in Energy and Buildings, Volume 145, Pages 284-292, M. Hosseini, B. Lee, S. Vakilinia, “Energy performance of cool roofs under the impact of actual weather data”, © Elsevier Ltd. 2017

3.1 Introduction

The actual weather condition variation such as variation in solar radiation, air temperature, and wind speed affect the performance of the building considerably. Hence, the idea of using a single representative weather year was formed to generalize long term weather condition for a specific location. Typical meteorological year (TMY) and weather year for energy calculation (WYEC) weather data are commonly used to evaluate building energy consumption at design stage. However, these data, while typical as they are intended to be, for certain years, they might overestimate or underestimate the energy consumption of the building and this can have a great impact on predicting the amount of energy saving obtained from building energy efficiency measures. The impact of 30-year actual weather data on HVAC source energy use, total source energy use, peak electric demand, peak electric demand reduction and energy savings is

investigated in a previous study [1]. The impact of weather on the long-term performance of buildings is also evaluated from a life cycle perspective [2]. Since TMY or WYEC are not taken into account extreme weather conditions [3]; actual data are suggested for building design, especially when the building solutions are strongly influenced by weather conditions [4].

3.1.1 TMY vs. AMY

Other literatures compared the long term mean energy consumptions of buildings according to building codes using actual years with what estimated by simulation using typical meteorological data. Hong et al. [2] compared long-term actual meteorological years (AMY) simulation from 1980 to 2009 with that of TMY3 for three types of office buildings according to two energy efficiency codes namely, ASHRAE Standard 90.1, 2004 and ASHRAE Standard 90.1, 2010, across 17 ASHRAE climate zones; Their results showed that in most cases TMY3 underestimated long term annual energy consumption of the buildings where in one case the deviation amounted to -9%. They concluded that energy performance prediction based on TMY3 weather data is not necessarily representative of the average energy use over a long period, and it can be significantly higher or lower than that based on the AMY data. In addition, they suggested that the impact of weather is greater for buildings in colder climates than warmer climates. In another study, Yang et al. [5] conducted building energy simulation for office buildings in the five Chinese climate zones namely severe cold winter, cold winter, warm winter, hot summer and cold winter, mild and hot summer, using AMY from 1971 to 2000 to compare the long term energy consumption mean with that simulated using TMY. They found the deviation came with a mean bias errors ranged from -4% in Guangzhou to 0% in Beijing and root-mean-square errors from 3% in Harbin to 5% in Guangzhou.

This study investigates the long term performance of various roof design combinations and compares the mean long term energy performance simulation from the year 1960 to 1989 with the predicted energy performance of a typical meteorological year (CWEC). It demonstrates the impact due to the variation of weather on energy performance of buildings based on historical data and identifies optimal roof designs.

3.2 Methodology

A large scale simulation considering various roof designs in the case of solar reflectance and thermal resistance are conducted with actual historical weather data and a typical meteorological year. For each design combination, the deviation of the results with the typical meteorological year from the actual years are then calculated. The following sections describe how historical weather data are collected and imported to simulation program.

3.2.1 Weather data

Canadian Weather Energy and Engineering Datasets (CWEEDS) is a set of hourly weather data for different locations in Canada [6]. Depending on the location (station), datasets include the weather data necessary for urban planning and energy efficient buildings between the years 1953 and 2005. Canadian Weather Year for Energy Calculation (CWEC), a variant of TMY and Weather Year for Energy Calculations (WYEC) [7], is a single typical meteorological year including twelve statistically selected months from CWEEDS. The selection is carried out by a comparison of cumulative density function (CDF) of the monthly meteorological data such as solar radiation, outdoor air temperature, and wind speed for long term (usually 30 years) dataset. In this study, the station of Montreal Trudeau international airport with station identification number 94792 was selected to investigate. For this station, the months of CWEC are selected from the years 1960 to 1989 [8]. Table 3.1 shows the corresponding years of each month of CWEC for the weather station.

Table 3.1: Corresponding years of CWEC months for Montreal airport weather station

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Year	1966	1970	1961	1979	1971	1970	1977	1978	1979	1986	1984	1978

There are a few hours in some of the years of CWEEDS dataset that solar radiation data is missed, especially the first few daylight hours of the morning; to fill in the blank points, interpolation is conducted to estimate the missing data. The weather year files are then constructed with a weather file creator tool for the simulation purpose.

3.2.2 Weather file creator

Elements [9] a free open-source software tool is used to create weather file for energy simulation. Elements enable users to create weather file in the formats of DOE-2 (.bin, .fmt) and EnergyPlus (.epw). A special feature of Elements is that some of the weather parameters will be calculated automatically according to psychrometric chart as there are rational relations among some of the meteorological data; for example, wet bulb and dew point temperature can be calculated with dry bulb temperature, atmospheric pressure and relative humidity. In addition, global solar radiation is calculated with normal and diffuse solar radiation into Elements; moreover, wind speed and wind direction are also included in the weather data (overall 10 weather parameter). After constructing the historical weather years, the large scale simulations are conducted with multiple historical actual weather years and the CWEC. For each design combination, the deviation of result simulated with CWEC from the actual years are calculated which is explained in the following section.

3.2.3 Normalized root-mean-square deviation

The root-mean-square deviation (RMSD) also known as root-mean-square error (RMSE) is frequently used to show the sample standard deviation of the differences between values predicted by a model and the actual measured values. The RMSD sums up the errors in predictions for various times into a single measure therefore it is a good measure of accuracy, to compare estimating errors of different models for a variable [10]. By analogy, for each design combination, RMSD of annual energy demand resulted from CWEC compared to annual energy demand of actual-year is defined as the square root of the mean square deviation:

RMSD of CWEC can be calculated based on equation 3.1.

$$\text{RMSD}(E_{\text{CWEC}}) = \sqrt{\text{MSE}(E_{\text{CWEC}})} \quad (3.1)$$

Where: $\text{MSE}(E_{\text{CWEC}})$ is the mean of the square of the errors and can be calculated from equation 3.2.

$$\text{MSE}(E_{\text{CWEC}}) = \frac{1}{n} \sum_{i=1}^n (E_i - E_{\text{CWEC}})^2 \quad (3.2)$$

Where: E_i is the annual heating, cooling, and total energy of each actual year, E_{CWEC} is the annual heating, energy, and total energy resulted from simulation with CWEC, and n is equal to 30 representing the number of actual years.

Normalizing the RMSD facilitates the comparison between datasets with different scales [11]. Common choice is the range defined as the difference between the maximum and the minimum annual energy of the 30 actual-years simulations as equation 3.3.

$$NRMSD = \frac{RMSD}{E_{max} - E_{min}} \quad (3.3)$$

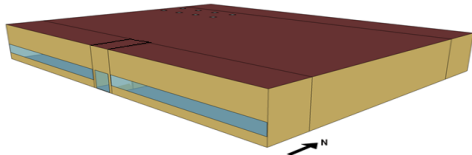
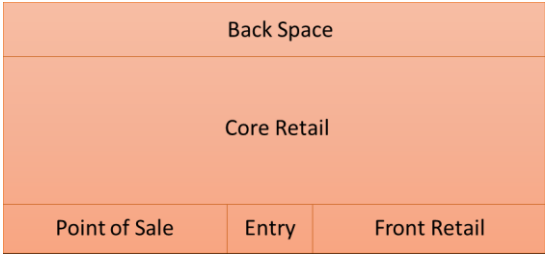
3.2.4 Case study

A full factorial simulation study is conducted to study the long term performance for a variety of different roof designs and the associated deviation of CWEC prediction from actual weather years in Montreal. The study is mostly focused on the deviation of results simulated with typical meteorological data from long term energy performance of cool roofs considering the effect of design variation namely, roof solar reflectance and roof thermal insulation level.

Since a big-box building incurs a relatively large roof surface (as compared to wall surfaces), a 2299 m² floor area retail store prototypical building offered by U.S. Department of Energy [12] is selected as the case study. The building enclosure is modified according to the National Energy Code of Canada for Buildings (NECB) 2011 [13]. The design parameters are roof insulation and solar reflectance of the roof. A parametric study of energy performance of the buildings of the different design variations has been performed using JePlus, an EnergyPlus simulation manager for parametric studies [14]. For parametric study of each design variation, the CWEC weather data are applied together with the actual weather data from 1960 to 1989. The results are 31 sets (based on CWEC and the 30 actual years) of cooling, heating and total annual energy demand data. Overall, 3906 simulations (14 roof insulation levels from 2.4 m²-K/W to 15.4 m²-K/W with an interval of 1 m²-K/W, 9 roof solar reflectance levels from 0.1 to 0.9 with interval of 0.1, and 31 sets of weather data) are run using JePlus.

Since the focus is on roof performance rather than on HVAC system design, the energy demand (heating and cooling) rather than the energy consumption (considering the effect of HVAC system efficiency) is investigated. Table 3.2 shows the characteristics of the case study building.

Table 3.2: Building characteristics of the case study

Item	Descriptions
General	
Location	Montreal
Building Prototype	standalone retail
Form	
Total Floor Area (square meter)	2299 (54.2 m x 42.3 m)
Building shape	
Number of Floors	1
Window-to-Wall Ratio (WWR)	25.4% on the south facing facade
Thermal Zoning	
	
Envelope	
Exterior walls	
Construction	Steel-Frame Walls: Wood Siding + wall Insulation+1/2 in Gypsum Board
R-value (m ² K /W)	4.1
Roof	
R-value (m ² K /W)	Built-up Roof: roof membrane + roof insulation + metal decking 2.4 to 15.4
Window	
R-value (m ² K /W)	0.5
SHGC	0.3
Foundation	
R-value(m ² K/W)	5.9
Air Barrier System	
Infiltration	0.001024 m ³ /s/m ² of above ground envelope area
Lighting	
Average power density (W/m ²)	12.5, Core zone 36.2
Plug load	
Average power density (W/m ²)	3.2, Point of sale 21.5 and back space 8
Occupancy	
Area (m ²)/Person	6.2, back space 27.8

3.3 Results and discussion

The results of the building’s annual heating, cooling, and total (heating + cooling) energy demand for two different roof designs, one with solar reflectance of 0.9 and thermal resistance of 2.4 m²-K/W (solid lines and curves) and the other with solar reflectance of 0.1 and thermal resistance of 15.4 m²-K/W (dotted and dashed lines and curves) are shown in Figure 3.1.

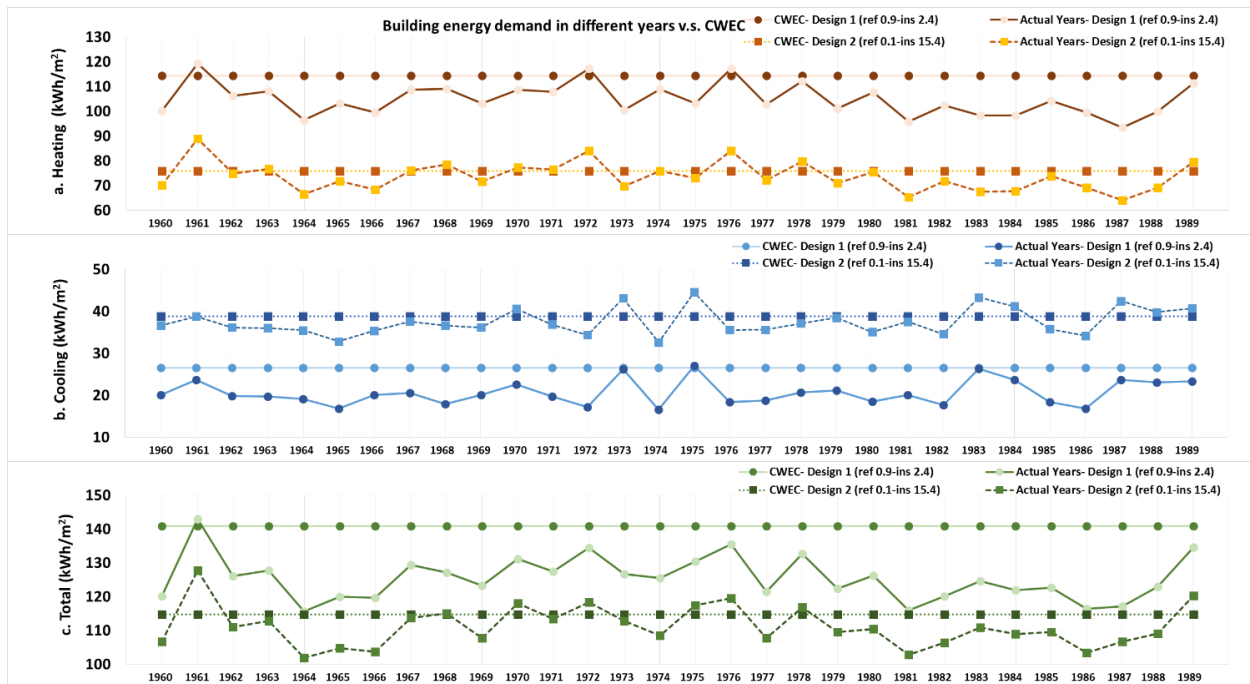


Figure 3.1: Heating, cooling, and total demand of the prototypical building, 30-years actual weather data V.S. CWEC for two designs. “ins” and “ref” represent insulation value and reflectance respectively.

From Figures 3.1.a, 3.1.b, and 3.1.c, generally there are larger difference between energy demand simulated with CWEC and actual years for the first design with high solar reflectance-low thermal insulation compared to that for second design with low solar reflectance-high thermal insulation. However, this is not always true; for instance, for heating, Figure 3.1.a shows that in 1987 which has the minimum energy demand among all the 30 years, the difference is about 21 kWh/m² for the first design and 16 kWh/m² for the second design respectively; in contrast, in 1961 which has the maximum heating energy consumption among the thirty years, the difference is only around 3 kWh/m² for the first design and 13 kWh/m² for the second design. Furthermore, in a few years such as 1968, 1978, and 1989, CWEC overestimate the heating energy for the first design while it

underestimates the heating energy for the second design although this inaccuracy is quite small (around 1 kWh/m²)

As for cooling, generally, the Figure 1.b shows that the more insulation level and low solar reflectance increases the cooling energy significantly. Moreover, with the first design, the deviation between cooling energy simulated by CWEC and actual years is almost zero for the years with the most cooling energy namely, 1973, 1975, and 1983; whereas with the second design the difference increases up to round 6 kWh/m² in 1975. In addition, in a few years such as 1970, 1984, 1987, 1988, and 1989 the cooling energy.

Total energy is the summation of cooling and heating energy demand. As Figure 3.1.c illustrates, in some years such as 1960, 1964, and 1965, CWEC overestimates the total energy dramatically as a result of overestimation in both heating and cooling energy for both designs. In other years such as 1972 and 1976, CWEC underestimates heating energy as much as it overestimates the cooling energy consumption for both designs so that the discrepancies cancel out each other and the total energy of CWEC would be very close to the actual year results. In another years including 1979, for the second design CWEC estimates the heating and cooling energy accurately therefore, the total energy consumption simulated with CWEC is very close to what simulated with actual years. Unlike the second design, for the first design CWEC overestimates both cooling and heating energy, consequently it overestimates the total energy in 1979. However, generally it seems that, the simulation with CWEC is much closer to the extreme events; namely, 1961, 1972, 1976, 1978 for heating, 1973, 1975, 1983 for cooling, and 1961, 1972, 1976 for total energy. Unlikely, for the second design, energy consumptions simulated with CWEC stands in the middle of the actual-years variation. Since different years represent various behavior compared to CWEC according to design, deviation of energy simulated with CWEC from actual years for different design combinations helps to have a better understanding.

Figures 3.2.a, 3.2.b, and 3.2c show the energy demand and normalized root-mean-square deviation of CWEC for each design combinations of roof solar reflectance and roof thermal resistance.

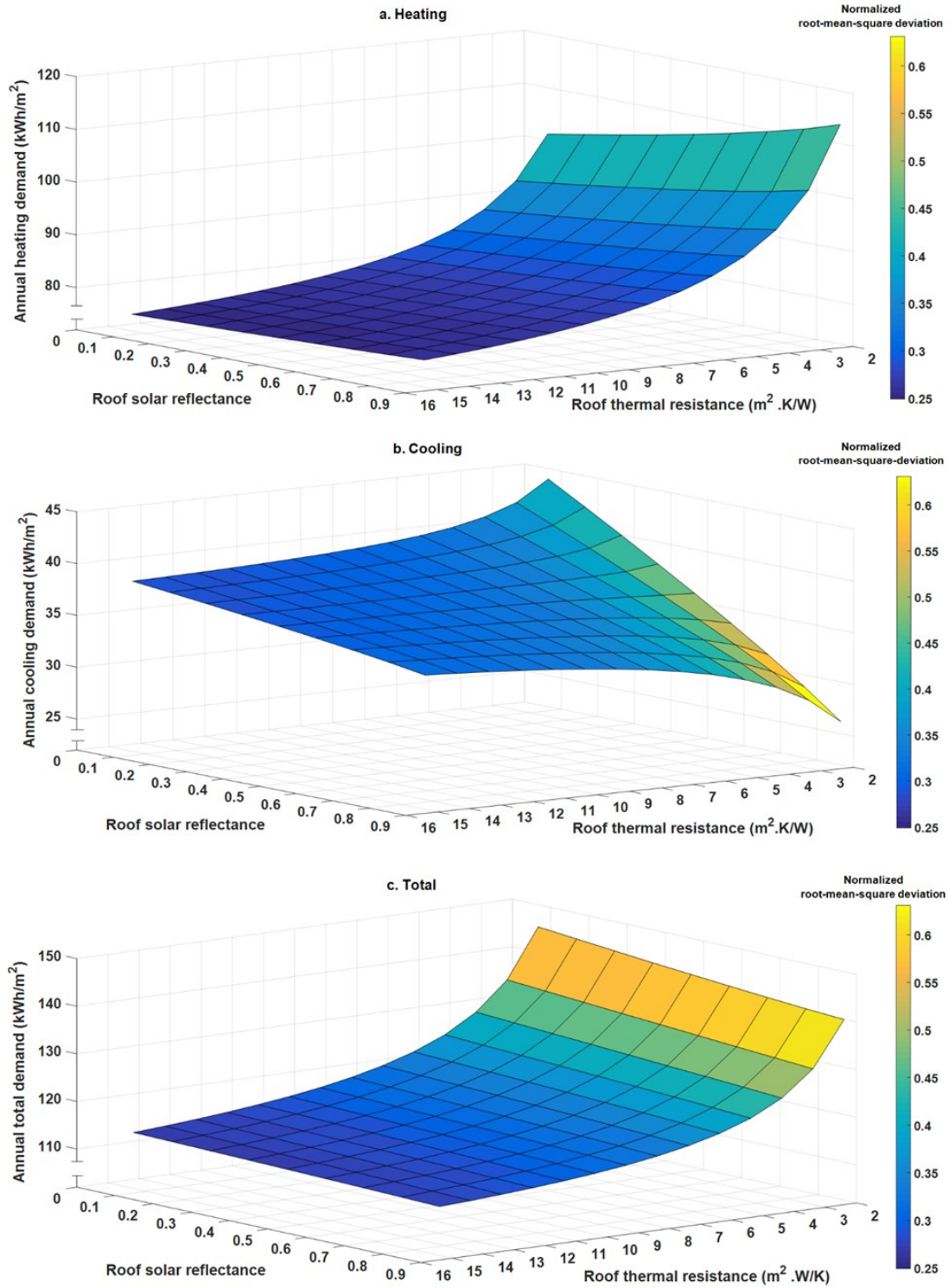


Figure 3.2: Heating, cooling and total demand and normalized root-mean-square deviation of CWEC for different roof design combinations.

3.3.1 Effect on heating demand

As can be seen from Figure 3.2.a, heating energy demand of the building is more influenced by the level of insulation rather than solar reflectance. With low level of insulation, changing the solar reflectance from 0.9 to 0.1 leads to a 12 kWh/m² reduction in heating demand from 114 to almost 102 kWh/m² and this difference would be even much smaller (about 1kWh/m²) when high level of insulation is used for roof. Increasing the level of insulation however, dramatically reduces the heating demand of the building. For example, changing the level of insulation from 2.4 to 15.4 m².K/W results in a considerable heating demand reduction of 36 kWh/m² and 27 kWh/m² for high reflectance (cool roof) and low reflectance (dark roof) respectively.

Meanwhile, Figure 3.2.a shows that deviation of CWEC from actual years, for higher reflectance is just slightly bigger than that for lower reflectance. For example, for thermal resistance of 2.4 m²-K/W the heating normalized root-mean-square deviation of CWEC is about 0.44 for reflectance of 0.9 and 0.42 for reflectance of 0.1. Whereas, with roof thermal resistance of 15.4 m²-K/W the root-mean-square deviation of CWEC from actual years is 0.25 for both reflectance of 0.9 and 0.1.

For a constant level of solar reflectance, normalized root-mean-square deviation of CWEC of heating decreases significantly, as roof thermal resistance increases; however, this reduction will not be influenced by changing the solar reflectance considerably. For instance, with solar reflectance of 0.9 increasing level of insulation from 2.4 to 15.4 m²-K/W corresponds to reduction of normalize root-mean-square deviation from 0.45 to 0.25 (0.2 difference) and similarly, for solar reflectance of 0.1 these numbers change from 0.42 to 0.25 (0.17 difference).

3.3.2 Effect on cooling demand

As Figure 3.2.b shows, for low solar reflectance (dark roof), the cooling energy decreases as the level of insulation increases; lower solar reflectance increases the amount of solar radiation absorbed by roof and makes the roof surface hot and this considerably increases the heat transfer from outside to inside; therefore, extra level of insulation helps to prevent the heat transfer from outside to inside of the building in warm days of the year. However, increasing the level of roof insulation doesn't always reduce the energy demand. For high solar reflectance (cool roof) as level of insulation increases, the cooling energy increases as well. This is mostly happens in commercial buildings including retail stores where, because of high internal heat gain such as lighting, solar heat gain through the windows, occupant, and miscellaneous load, inside temperature tends to get

hot; because of low roof surface temperature, the most heat transfer is from inside to outside. In these condition, more level of insulation traps heat inside the building and increases the cooling load.

Moreover, solar reflectance plays the dominant role in cooling demand of the building. As Figure 3.2.b illustrates, with low level of insulation ($2.4 \text{ m}^2\text{-K/W}$), changing solar reflectance from 0.9 to 0.1 drastically increases the cooling demand from 26 to 44 kWh/m^2 (18 kWh/m^2 difference); even with very high level of insulation of $15.4 \text{ m}^2\text{-K/W}$, still the difference is considerable (5.5 kWh/m^2).

By comparing Figure 3.2.b and 3.2.a, the normalized root-mean-square deviation of CWEC from actual years for cooling demand is larger than that for heating demand. In addition, for cooling, design variation has a larger effect on deviation of CWEC from actual years. Solar reflectance plays the major role also in cooling normalized root-mean-square deviation of CWEC. It can be observed that with a constant level of insulation, the cooling demand deviation of CWEC from actual years is larger for higher reflectance designs (reflective roofs) compared with low reflectance designs (dark roofs). For thermal resistance of $2.4 \text{ m}^2\text{-K/W}$, the cooling normalized root-mean-square deviation of CWEC is about 0.62 for reflectance of 0.9 decreasing to about 0.38 for reflectance of 0.1 (0.24 difference). Meanwhile, with roof thermal resistance of $15.4 \text{ m}^2\text{-K/W}$ the normalized root-mean-square deviation of CWEC remains at around 0.3 as solar reflectance decreases from 0.9 to 0.1. Moreover, for a constant level of solar reflectance, cooling normalized root-mean-square deviation of CWEC considerably decreases as thermal insulation level increases from 2.4 to $15.4 \text{ m}^2\text{-K/W}$ for high reflectance (0.35 reduction) whereas it just slightly decreases for low reflectance designs (0.15 reduction).

3.3.3 Effect on total demand

By using high solar reflectance design, reduction in cooling energy demand is larger than increase in heating energy demand. In other words, cooling energy saving of cool roof is greater than the heating penalty of that; this means a net annual energy saving of cool roof. For example, for low level of insulation of $2.4 \text{ m}^2\text{-K/W}$, changing solar reflectance from 0.1 to 0.9 decreases the cooling energy from 44 to 26 kWh/m^2 (18 kWh/m^2) whereas, it increases the heating energy by 12 kWh/m^2 , which implies 6 kWh/m^2 of net annual total energy saving for cool roof (Figure 3.2.c). This is because during winter in cold climates, the days are shorter, sun angle is lower, the sky is cloudier;

which leads to less solar radiation available hitting the roof in winter. As the result shows, for heating demand the insulation level plays the dominant role in roof design as less solar radiation hits the flat roof in heating season, whereas in cooling season, solar reflectance plays the major role in roof design. By comparing the results from cooling and heating, the importance of solar reflectance overweighs the importance of insulation level in roof design. However, in terms of the total annual demand, the best design combination is the low solar reflectance with high level of insulation in cold climate; such that it reduces the cooling demand in cooling season while decreases the heating demand in heating season. As Figure 3.2.c shows, the maximum total demand belongs to the design with low reflectance (dark roof) with low level of insulation, and the minimum total demand belongs to the design with high solar reflectance (cool roof) and high level of insulation.

On the other hand, Figure 3.2.c shows that even though increasing level of insulation considerably decreases root-mean-square deviation of CWEC from actual years by 0.3, changing solar reflectance doesn't have a great effect on that; such that the root mean square deviation of CWEC from actual years in worst case remains in the order of 0.55 with only a slight variation of 0.05 by changing solar reflectance from 0.1 to 0.9.

3.3.4 Overall effect

Normalized root-mean-square error only shows the amount of deviation. Since it is a positive number, it doesn't indicate overestimation or underestimation considering long term duration of time. Hence, by comparing the mean of the thirty actual years with CWEC and also the difference of CWEC and mean with standard deviation of the thirty years, one can determine the appropriateness of applying CWEC for different designs. Among all the 135 designs the results of 6 designs are summarized in Table 3. The six designs consist of: Design 1 with thermal resistance of $2.4 \text{ m}^2\text{-K/W}$ and solar reflectance of 0.9, design 2 with thermal resistance of $2.4 \text{ m}^2\text{-K/W}$ and solar reflectance of 0.1, design 3 with thermal resistance of $5.4 \text{ m}^2\text{-K/W}$ and solar reflectance of 0.9, design 4 with thermal resistance of $5.4 \text{ m}^2\text{-K/W}$ and solar reflectance of 0.1, design 5 with thermal resistance of $15.4 \text{ m}^2\text{-K/W}$ and solar reflectance of 0.9, and design 6 with thermal resistance of $15.4 \text{ m}^2\text{-K/W}$ and solar reflectance of 0.1. Note that the level of insulation for design 3 and 4 is based on national energy code of Canada for buildings (NECB, 2011).

Table 3.3 presents the maximum, minimum, mean, variance, standard deviation, energy results from CWEC and overestimation of CWEC, $(\frac{CWEC-Mean}{Mean})$, for each six roof design.

Table 3.3: Statistics of energy demand of the building in 30-years V.S. CWEC for six different roof design combinations

Design Number	Thermal resistance (m ² -K/W)	Solar reflectance	Max	Min	Max-Min	Mean	Variance	Standard deviation	CWEC	CWEC-Mean	Overestimation (%)
Heating			(kWh/m²)								
1	2.4	0.9	119	93	26	105	42	6	114	10	9
2	2.4	0.1	108	82	27	93	40	6	102	9	10
3	5.4	0.9	101	76	25	86	37	6	91	5	6
4	5.4	0.1	96	71	25	81	36	6	86	5	6
5	15.4	0.9	91	66	25	76	34	6	78	2	3
6	15.4	0.1	89	64	25	74	34	6	76	2	3
Cooling											
1	2.4	0.9	27	17	10	21	9	3	27	6	29
2	2.4	0.1	48	35	14	40	12	3	44	4	10
3	5.4	0.9	35	24	11	28	9	3	32	3	12
4	5.4	0.1	46	33	12	38	11	3	41	2	6
5	15.4	0.9	40	29	11	34	9	3	35	2	5
6	15.4	0.1	45	33	12	38	10	3	39	1	3
Total											
1	2.4	0.9	146	110	36	125	42	7	141	15	12
2	2.4	0.1	156	116	40	133	39	6	146	13	10
3	5.4	0.9	136	100	36	114	39	6	123	8	7
4	5.4	0.1	142	104	38	119	37	6	126	7	6
5	15.4	0.9	131	95	36	109	37	6	113	4	4
6	15.4	0.1	133	97	37	111	36	6	115	4	3

From Table 3.3, for different designs, the difference between maximum and minimum heating demand among the 30 actual years remains almost constant at around 25 kWh/m² no matter what roof design it is; whereas, for cooling, solar reflectance has a considerable effect on this difference. With low level of insulation (designs 1 and 2), the difference between maximum and minimum is smaller for cool roof compared to that for dark roof; although, increasing the level of insulation reduces this difference.

As can be seen from Table 3.3, for heating, with low thermal resistance (designs 1 and 2), the simulation conducted with CWEC deviates from the mean of thirty actual years considerably (10 kWh/m² or approximately 10 % overestimation), such that it is located out of the first standard

deviation. On the contrary, with high level of insulation (designs 5 and 6), heating energy of the building simulated with CWEC is close to the mean of the thirty-year simulations (2 kWh/m² or 3% deviation). With a standard constant level of insulation (5.4 m²-K/W as suggested by NECB, 2011, designs 3 and 4), changing the solar reflectance from 0.9 to 0.1 doesn't have a significant impact on deviation of CWEC from the mean of thirty actual years for heating demand (6% overestimation for both designs), and both designs are within the standard deviation.

In addition, although more solar reflectance and thermal insulation reduces the cooling energy demand of the building, it increases the overestimation of CWEC for cooling demand. As can be seen from table 1, for design 1, the simulation result with CWEC is far beyond the standard deviation (by 6 kWh/m² or 29% overestimation), In other words, CWEC can underestimate the cooling energy saving of a reflective low insulation roof by 29 %. Design 2 however, makes the CWEC closer to the mean of thirty-year simulations yet, out of the standard deviation with 4 kWh/m² or 10 % overestimation.

Moreover, as it is illustrated, with standard level of insulation (designs 3 and 4), although lower solar absorption reduces cooling demand of the building, it increases overestimation of CWEC such that the difference between the simulation run with CWEC and the mean of thirty years is larger than the standard deviation. Whereas, lower solar reflectance decreases the overestimation of CWEC yet, the simulation with CWEC falls within the standard deviation. This means that with standard level of insulation, CWEC overestimates the cooling energy of the building with reflective roof by 12 % whereas it overestimates that for dark roof by around 6 %.

Overall, Table 3.3 shows that, for heating, with a constant level of standard insulation (designs 3 and 4), the difference between the results simulated with CWEC and the mean of thirty years are almost the same for cool roof and dark roof; while for the same condition, cooling energy of building with cool roof (design 3) is lower than what is expected from simulation with CWEC, in comparison to what is expected from dark roof (design 4). Consequently, a higher energy saving can be obtained in real life by an application of cool roof in the long term

3.4 Conclusion

A parametric study of energy performance of the buildings of the different design variations has been performed. For each design variation, the CWEC weather data is applied together with the long term actual weather data from 1960 to 1989. The mean performance of the 30 years is also compared to that of CWEC for the different design variations to fully assess the appropriateness of applying a TMY weather file to evaluate different roof designs with building energy simulation.

The results show that, the best design combination is the low solar reflectance with high level of insulation in cold climate. Moreover, the appropriateness in applying TMY weather data to building energy simulation is highly dependent on the roof designs, which can result in significant underestimation or overestimation of the energy savings. Although using cool roofs instead of a regular dark roof decreases the predicted cooling and total energy demand of the building with TMY, the deviation between such prediction and the mean energy demand as evaluated with AMY weather data is also increased. However, such deviation is fairly small compared to the absolute value of the total energy demand of a dark roof. It is only through this comparison of energy performance evaluated under TMY and AMY weather data, insight on the long term energy performance of different roof configurations (notably cool roof and dark roof) can be obtained. The result also indicates that a higher than predicted energy saving can be obtained in the long term by deploying cool roof based on a conventional TMY weather data based simulation.

3.5 References

- [1] T. Hong, W.K., Chang, H. W, Lin. “A sensitivity study of building performance using 30-year actual weather data”, In Proceedings of 13th Conference of International Building Performance Simulation Association, August 2013, Chambéry, France
- [2] T. Hong, W.K. Chang, H.W. Lin, “A fresh look at weather impact on peak electricity demand and energy use of buildings using 30-year actual weather data”, Applied Energy, 111, 333–350, 2013
- [3] R. S. McLeod, C. J. Hopfe, Y. Rezgui, “A proposed method for generating high resolution current and future climate data for Passivhaus design”, Energy and Buildings 55, 481–493, 2012
- [4] A. A. Kasam, B. D. Lee, C. J. Paredis, “Statistical methods for interpolating missing meteorological data for use in building simulation”, Building Simulation, 7: 455, 2014
- [5] L. Yang, J.C. Lam, J. Liu, C.L. Tsang, “Building energy simulation using multi-years and typical meteorological years in different climates”, Energy Conversion and Management 49, 113–124, 2008
- [6] Environment Canada, historical weather data:
http://climate.weather.gc.ca/prods_servs/engineering_e.html, retrieved September, 2016

- [7] WYEC2 user's manual, <http://search.proquest.com/docview/192572875?pq-origsite=gscholar>, retrieved September, 2016
- [8] CWEEDS documentation, ftp://ftp.tor.ec.gc.ca/Pub/Engineering_Climate_Dataset/Canadian_Weather_Energy_Engineering_Dataset_CWEEDS_2005/ZIPPED%20FILES/ENGLISH/CWEEDS%20documentation_Release9.txt
- [9] Elements, custom weather file creator, <http://bigladdersoftware.com/projects/elements/>, retrieved September, 2016
- [10] R.J. Hyndman, A.B. Koehler, "Another look at measures of forecast accuracy International Journal of Forecasting 22, 679 – 688, 2006
- [11] "Coastal Inlets Research Program (CIRP) Wiki - Statistics", Retrieved 1 September 2016
- [12] U.S. Department of Energy, prototypical buildings, https://www.energycodes.gov/development/commercial/prototype_models, retrieved September, 2016
- [13] National Energy Code of Canada for Buildings 2011, http://www.nrc-cnrc.gc.ca/eng/publications/codes_centre/2011_national_energy_code_buildings.html
- [14] JEPlus, an EnergyPlus simulation manager for parametrics, <http://www.jeplus.org/wiki/doku.php>, retrieved September, 2016

Chapter 4. THE APPLICABILITY OF CANADIAN WEATHER YEAR FOR ENERGY CALCULATION (CWEC); THE EFFECT OF BUILDING DESIGN AND CLIMATE

The intended purpose of a single-year typical meteorological weather file is to assess the long-term average performance, while disregarding year-to-year fluctuations and extreme conditions. This chapter aims to provide a better understanding of the applicability of current Canadian typical meteorological year (CWEC) for building different building design and under different cold climate conditions. The study explores design parameters that are seldom prescribed in energy efficiency standards and codes but have great impact on cooling and heating energy demands. The studied design parameters include window-to-wall ratio (WWR), window solar heat gain coefficient (SHGC), floor construction (concrete or wood), and solar reflectance of exterior walls and roof. These design parameters are applied to two different building types — small and large office buildings, and two different climatic locations — Montreal and Vancouver, Canada.

This chapter is submitted to the journal of Urban Climate, M. Hosseini, A. Bigtashi, B. Lee, “Evaluating the applicability of Typical Meteorological Year under different building designs and climate conditions”, 2021

4.1 Introduction

Typical meteorological year (TMY) data, is developed to estimate the long-term buildings energy performance or solar energy conversion systems performance through computer simulations. Although the file is created to estimate the long-term average weather data, many studies have found TMY data to be an inadequate representation when evaluating specific energy system applications or building criteria [1, 2] . The methodology behind the TMY file generation has also been found to play a significant role in results [3, 4] putting into question the reliability of these files.

The artificial single year is composed of 12 typical months which contain actual hourly weather data. Each month is selected based on the similarity to median weather conditions as a means to best represent long-term weather data. A variety of different file formats including International

Weather year for Energy Calculations (IWEC)[5], Typical Meteorological Year (TMY) [6], TMY2 [7], TMY3 [8], Test Reference Year [9], and Canadian Weather Year for Energy Calculation (CWEC) [10] are readily available.

There are a few methods to generate the typical meteorological data including, but not limited to, Sandia National Laboratory method [6, 11], the Danish method [9] and the Festa and Ratto method [11]. A study conducted by Janjai and Deeyai [4] evaluated the three above-mentioned methods; the statistical test results showed a minor difference between the long-term average of the weather parameters and the corresponding TMY generated from the different methods. The TRNSYS energy simulation conducted for two solar water heating systems and two photovoltaic systems with varying sizes in four tropical climates showed a good agreement between the solar fraction and electricity resulted from the TMYs and the average long-term data. The Sandia method is the most well-known and commonly applied method for generating TMY files. Individual months are selected from the pool of available meteorological data, where the Cumulative Distribution Function (CDF) is applied to each parameter considered for both short-term and long-term daily mean values [12]. The parameters considered are the dry bulb temperature, relative humidity, wind velocity and global solar Irradiance, where the min, max and range are considered when applicable.

There have been many studies conducted on the effects of different typical weather year generation methods on simulated building performance [1, 3, 13]; as well as the impact for applications in different climates [14–21].

In an earlier study, Crawley [22] found an annual energy consumption variation between the long-term average and the typical weather year ranging between 7.0% and 11.0% for a simulated prototypical office building while using different weather files for varying climates, namely; TRY, TMY, TMY2, WYEC, and WYEC2. The authors stated that an identical building model was kept for energy simulation of all climates; furthermore, the older and/or lower energy efficiency standards could contribute to higher amount of deviation. Other studies have expanded their research to include the effects of different typical weather years on different building design aspects, such as daylighting, and energy system applications [2, 23].

Another study was conducted by Ebrahimpour & Maerefat [1] to assess the difference between TMY files generated using different methods. The Sandia method, as well as the Meteornorm and Weathergenerator software were used to generate TMY files which were subsequently compared

to the long-term average measured data. The goal was to determine which generation method would provide minimal difference to the long-term average measured data. The results show that the Sandia method, as well as the data generated from Meteonorm, best met the above-mentioned criteria.

Questioning whether a single typical meteorological year can properly represent the range of climate conditions over a period, Crawley & Lawrie [24] suggested generating files by selecting extreme climatic months to create an eXtreme Meteorological Year (XMY). Guo et al. [25] applied a similar concept, which they refer to as typical hot-year (THY) in order to assess the indoor environment during extreme heat waves.

Yang et al. [26] conducted a study where a typical office building was simulated for five different climate zones in China using actual meteorological year (AMY) data from 1971 to 2000 to compare the long-term mean to typical meteorological year results. The results showed root-mean-square errors ranging from 3% to 5% between the AMY average and TMY. Furthermore, Hong et al. [27] evaluated the impact of using 30-year actual weather data for simulation in terms of HVAC source energy use, total source energy use, peak electric demand, peak electric demand reduction and energy saving. They also analyzed the impact of weather on the long-term performance of buildings considering three types of office buildings according to two energy efficiency codes across 17 ASHRAE climate zones. The study highlighted that the TMY's capability of considering extreme weather conditions or regular events was dependent on climate. However, the intended use for TMY files are to represent typical weather conditions rather than extreme events. Moreover, building configuration, purpose and other design combinations were ignored in this assessment. Their results showed that TMY3 underestimated long-term annual energy consumption of the buildings by 9%.

Cui et al. [28] assessed the long-term representation of a typical year by comparing it to a 55-year actual weather data set for an office building simulated for 10 large cities covering all climate zones in China. Their results indicated that the typical year simulations overestimated or underestimated the energy use and peak load in many cases, including 2.7 kWh/m² (17.8 %) difference in cooling load in one case.

A study by Hosseini et al. [29] evaluated various roof designs for a Canadian cold climate, where a variety of parameters such as roof solar reflectance and thermal insulation were considered. The

results from the case study conducted on a single storey retail store office building indicated a smaller deviation between the CWEC and long-term average results for design with high solar reflectance (cool roofs) and higher thermal insulation.

A majority of commonly used TMY files focus on the thermal aspects of the climatic environment. Other studies have expanded their research to include the effects of different typical weather years on different building design aspects, such as daylighting and energy system applications [23]. Sun et al. [2] conducted an analysis on the impact of using different TMY data files for daylighting simulations. The study considered three different TMY data files (CSWD, SWERA and IWEC) for four cities in China with varying climates. The study found little difference in results between the different TMY data files when static metrics, such as Daylighting Factor (DF), were considered. However, when considering dynamic metrics, such as Daylight Autonomy (DA) and Useful Daylight Index (UDI), a significant divergence was noted, indicating that the TMY generation method must be taken into account when assessing results of daylighting studies. Similarly, Realpe et al. [23] evaluated the application of various TMY datasets for Concentrated-PV (CPV) systems. In the study, five different TMY data files were generated using the Sandia method, two simplified drivers which solely considered Direct Normal Irradiance (DNI) and two filtered drivers which took into account the characteristics of the CPV system. The study determined that the Sandia method was not suitable for simulating CPV systems, while driver-based files better represented the long-term average for both monthly and annual cases.

As previously mentioned, most studies conducted on typical meteorological year data and the impact of their generation methods utilize a single prototypical building set in varying climates. In most cases, only the climate based building envelope thermal properties are considered; with the exception of two studies conducted by Hosseini et al. [29, 30] which evaluated the deviation of results of CWEC from long-term average results of actual years considering a variety of roof designs. However, variations in other building properties such as WWR and window solar heat gain coefficient (SHGC) are neglected. This paper evaluates the applicability of typical meteorological year by considering building design properties, including WWR and SHGC, for two office buildings located in two cold climate cities.

4.2 METHODOLOGY

Building energy simulation is used to evaluate the applicability of the CWEC file in representing long-term average energy performance. The large-scale simulation involves a variety of building designs run once with 30-year actual historical data and once with CWEC. Design options include WWR, window solar heat gain coefficient (SHGC), floor type, and wall and roof solar reflectance, which equates overall to 16 design combinations, where each design parameter has two possible options.

The historical weather data of Montreal and Vancouver is extracted from the Canadian Weather Energy and Engineering Datasets (CWEEDS). The dataset includes the weather data necessary for urban planning and energy efficient building application for a period between the years of 1953 and 2014. The historical data is taken for a period ranging from 1960 to 1989 in order to compare to the acquired Canadian Weather Year for Energy Calculation (CWEC) file. CWEC is a single typical meteorological year which includes twelve statistically selected months. The selection is carried out by comparing the cumulative density function (CDF) of monthly meteorological parameters such as solar radiation, outdoor air temperature, and wind speed to the long-term average (usually 30 years) dataset.

The probability density function for the heating and cooling energy consumption over the 30-year period is calculated for each design option. The averages of each distribution are then compared to the results from the single CWEC simulation for each design combination.

4.2.1 Case Study

4.2.1.1 Weather condition

Two cold climate Canadian cities, Montreal (QC) and Vancouver (BC), are considered in this study. Generally, Vancouver has a milder climate and is classified in ASHRAE as zone 5, whereas Montreal, which exhibits colder winters, is classified as zone 6. The ASHRAE cooling and heating degree days of these two cities are shown in Table 4.1. Vancouver is overall cloudier and, generally, has less solar radiation available throughout the year than Montreal. Figure 4.1 compares the availability of solar radiation for both cities.

Table 4.1: Heating and Cooling Degree Days of the two studied cities

City	HDD18	CDD10
Montreal	4603	1192
Vancouver	3157	853

Figure 4.1 illustrates a histogram plot of 30 years of daily-averaged horizontal solar irradiance for Montreal and Vancouver. There is a greater frequency of values exceeding 200 W/m² for Montreal, indicating sunnier skies for the region, in comparison to Vancouver.

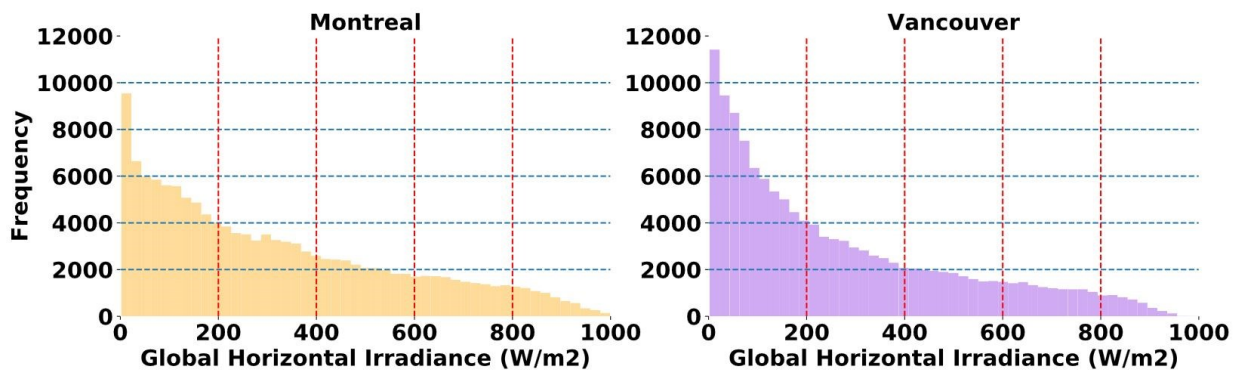


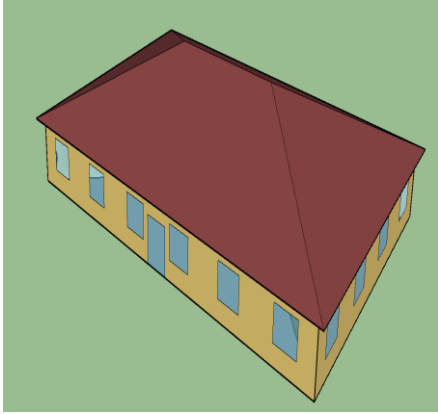
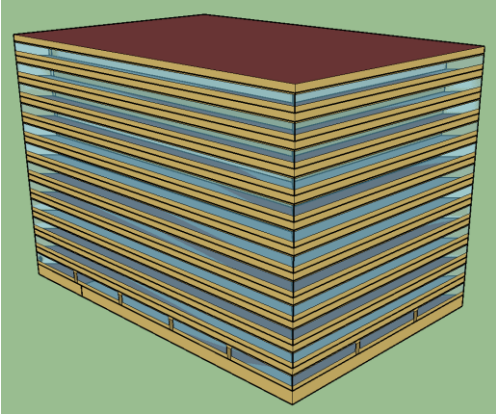


Figure 0.1: 30 years of hourly global horizontal solar Irradiance for Montreal and Vancouver from 1961 to 1989

In this study, the stations selected for investigation are the Montreal-Trudeau International airport, identification number 94792, and the Vancouver International airport, identification number 24287.

4.2.1.2 Building Properties

To evaluate the effect of varying designs on the overall applicability of CWEC, a single-storey small office building and a 12-storey large-office building are simulated. The two buildings are based on the US Department of Energy (DOE) prototypical office buildings (DOE) with associated EnergyPlus input files, which are publicly available [31]. The building envelopes are based on the level of efficiency measures set by the National Energy Code of Canada for Buildings (NECB) [32]. The original building specifications are available at the DOE website, however, the updated NECB specifications are summarized in Table 4.2.

Table 0.2: Properties of the studied buildings

Building Prototype	Small Office	Large Office
Floor Area (m²)	511	3567 (46356 total conditioned)
Building Shape		
Number of Floors	1	12 (plus basement)
Thermal Zoning	 Perimeter zone depth: 5m. Four perimeter zones, one core zone and an attic zone. Percentages of floor area: perimeter 70%, core 30%	 Perimeter zone depth: 4.5m. Each floor has four perimeter zones, one core zone and one IT closet zone. Percentages of floor area: Perimeter 29%, Core 70%, IT Closet 1% The basement has a datacenter zone occupying 28% of the basement floor area
Floor to floor height (m)	3	3.96
Floor to ceiling height (m)	3	2.74
Exterior Wall Construction	Wood-frame walls (2X4 16 in): 1 in. Stucco + 5/8 in. gypsum board + wall insulation+ 5/8 in. gypsum board	Mass (pre-cast concrete panel): 8 in. heavy-weight concrete + wall insulation + 0.5 in. gypsum board
Exterior Wall U-factor (W/m²-K)	0.246 (Montreal), 0.316 (Vancouver)	0.246 (Montreal), 0.316 (Vancouver)
Roof Construction	Attic roof with wood joist: Roof insulation + 5/8 in. gypsum board	Built-up roof: Roof membrane + roof insulation + metal decking
Roof U-factor (W/m²-K)	0.183 (Montreal), 0.227 (Vancouver)	0.183 (Montreal), 0.227 (Vancouver)
Window U-factor (W/m²-K)	0.39 (Montreal), 0.42 (Vancouver)	0.39 (Montreal), 0.42 (Vancouver)

In this study, four primary design parameters were selected; window-to-wall ratio (WWR), solar heat gain coefficient (SHGC), flooring and roof and exterior wall solar reflectance. The WWR is the ratio between the total glazed area (window) and the exterior envelope wall area. The value ranges between 0 and 1, where full window coverage is considered as being 1 or 100%. In this study, the two considered WWR values are 0.8 and 0.2. The solar heat gain coefficient refers to the characteristic of a window to transmit heat directly or indirectly into the building. The value for the parameter is a fraction ranging from 0 to 1; the closer the value is to 0, the less heat the window transmits into the building. In this case, the selected values for SHGC are 0.4 and 0.32 for Montreal and 0.45 and 0.35 for Vancouver. The flooring types selected in this study are concrete and wood. In this study, heavy weight 8-inch concrete slab or wood were selected for the small office building cases, while 100mm normal weight concrete or wood were used for the large office building cases. The exterior wall and roof solar reflectance represents the fraction of solar radiation which is reflected back from the wall and roof. This parameter allows to reduce the overall absorbed solar energy. Thus, a roof or wall with a high solar reflectance will absorb less energy and will be cooler than a regular wall or roof. The values range between 0 and 1, where light colored surfaces which have high solar reflectance approach 1 in value, while darker surfaces approach 0. In this study, the walls and roofs are assumed to have a solar reflectance of 0.8 or 0.4. Each parameter assessed in this study has two possible values as summarized in Table 4.3.

Table 4.3: values of design parameters used in the study

	WWR		Windows SHGC		Floor		Roof and wall solar reflectance	
	High	Low	High	Low	High	Low	High	Low
Small office Montreal	0.8	0.2	0.4	0.32	Concrete	Wood	0.8	0.4
Small office Vancouver	0.8	0.2	0.45	0.35	Concrete	Wood	0.8	0.4
Large office Montreal	0.8	0.2	0.4	0.32	Concrete	Wood	0.8	0.4
Large office Vancouver	0.8	0.2	0.45	0.35	Concrete	Wood	0.8	0.4

The focus of this case study is to evaluate the applicability of the CWEC file for a variety of building designs rather than for HVAC system design, therefore, the efficiency of the HVAC system is kept constant among different building designs.

Python-EnergyPlus (EPPY) package is used for the large-scale simulation. Each of the 16 design configurations is run with 30 actual historical year and the CWEC weather files. Overall 496

simulations were run for each building type - location. The results for the CWEEDS actual years are then compared to the CWEC results.

4.3 Results and discussion

After conducting the simulations, the outputs are processed in order to be analyzed. For each design option, the probability density function (PDF) of the 30-years cooling and heating energy consumptions are calculated, and the mean is compared to the corresponding results from CWEC. Figures 4.2 to 4.5 show the results for each building located in each studied location. SHGC represents the solar heat gain coefficient of the window, F stands for the floor type which can take the characteristic of concrete (C) or wood (W), and REF stands for solar reflectance of the exterior walls and roof. The dark blue shows the distribution of the long-term cooling, the light blue line shows the mean of the distribution, and the green line shows the cooling resulted from CWEC. By the same token, the red color illustrates the distribution of the long-term heating, the orange line stands for the mean of the distribution, and the gold color represents the heating resulted from CWEC. In each set of figures, a) presents the eight scenarios with small windows (WWR of 0.2) and b) presents the eight scenarios with larger windows (WWR of 0.8). In addition, the subplot located on the right shows the deviation of results of CWEC from the long-term mean, in which a positive value indicates underestimation by CWEC whereas a negative value signifies an overestimation.

4.3.1 The effect of climate (weather)

The cooling and heating energy consumption in Montreal for all simulated cases are larger than in Vancouver, as seen in Figures 4.2 to 4.9. The wider distribution for the Montreal cases indicates harsher weather conditions in contrast to the cooling and heating distributions exhibited in the Vancouver cases, which are much closer together, indicating a milder climate.

A greater deviation of the CWEC results from the long-term average results was noted for the Vancouver cases. An overall assessment for all small office building cases show cooling energy consumption ranging from 10 to 60 kWh/m² for Montreal and 7 to 50 kWh/m² for Vancouver. For heating energy consumption, the Montreal cases range between 50 to 100 kWh/m², while Vancouver cases range between 25 to 70 kWh/m².

Moreover, for the simulated large office buildings, the cooling energy consumption for Montreal and Vancouver fall between the ranges of 10 kWh/m² and 50 kWh/m² and between 5 kWh/m² and 45 kWh/m² respectively. The heating energy consumption ranges between 20 kWh/m² and 60 kWh/m² for Montreal and 3 kWh/m² and 35 kWh/m² for Vancouver.

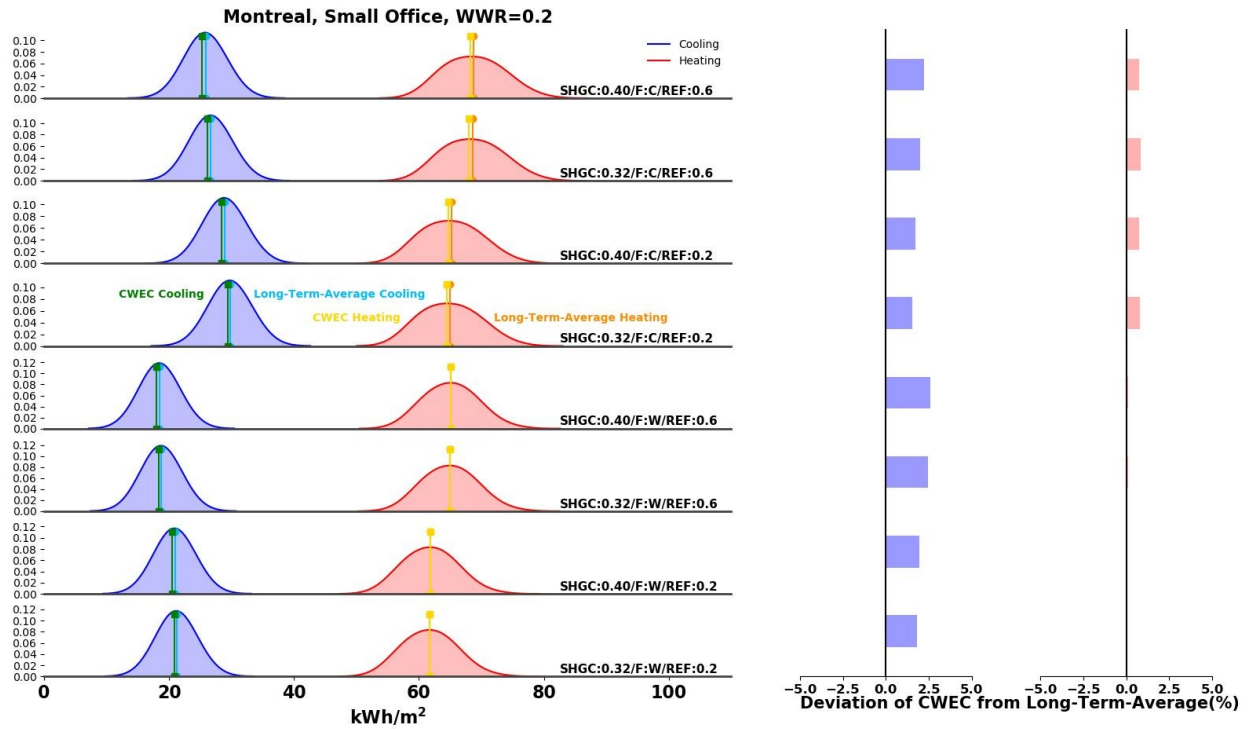


Figure 4.2: Montreal, small office with small windows

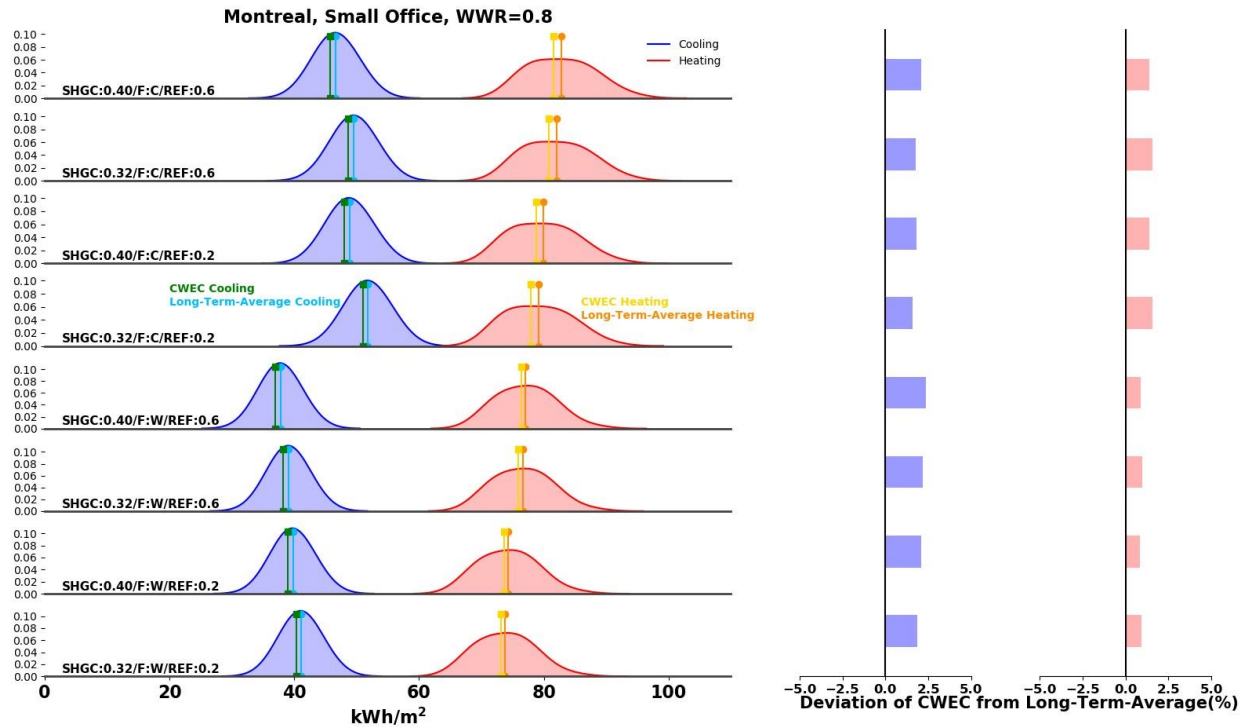


Figure 4.3: Montreal, small office with large windows

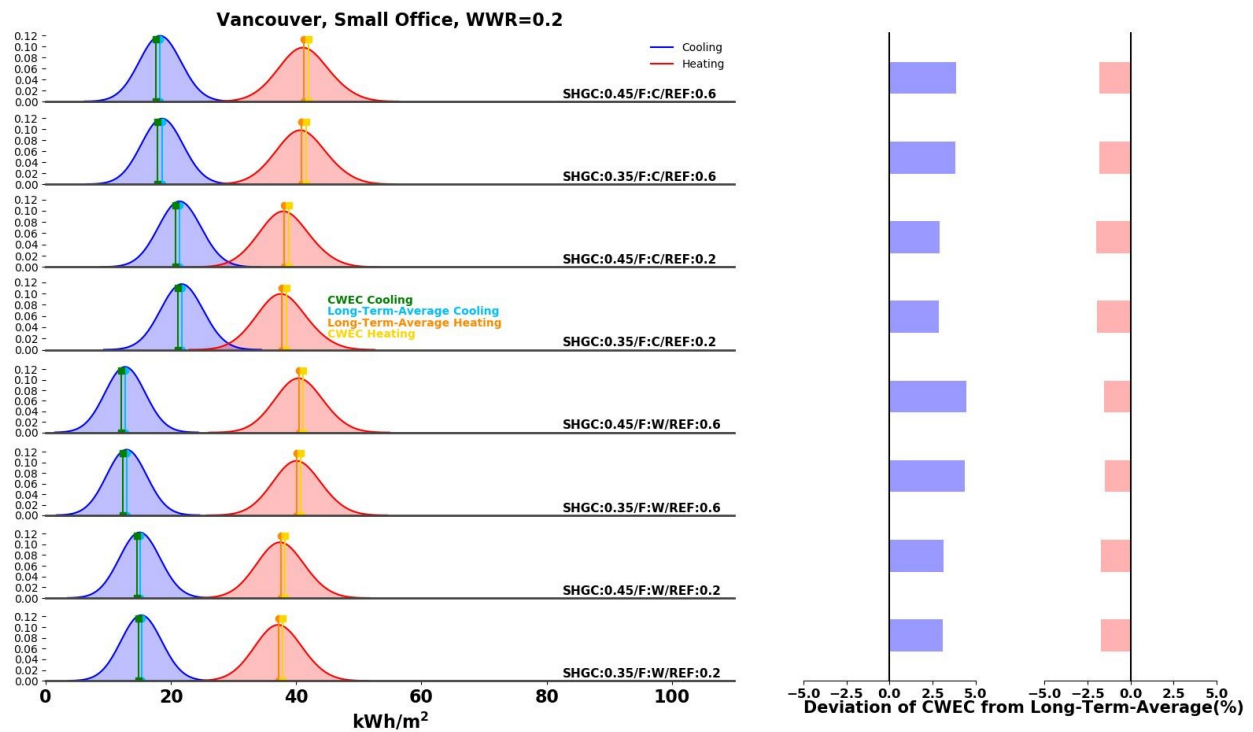


Figure 4.4: Vancouver, small office with small windows

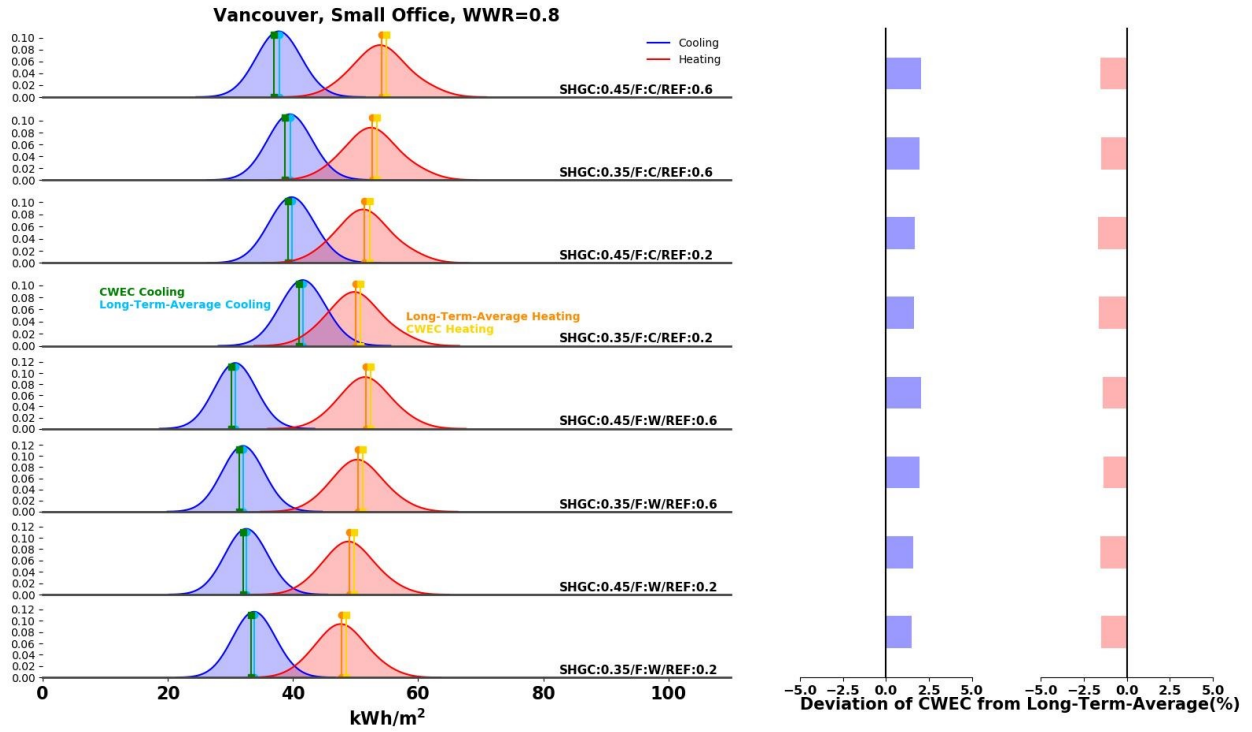


Figure 4.05: Vancouver, small office with large windows

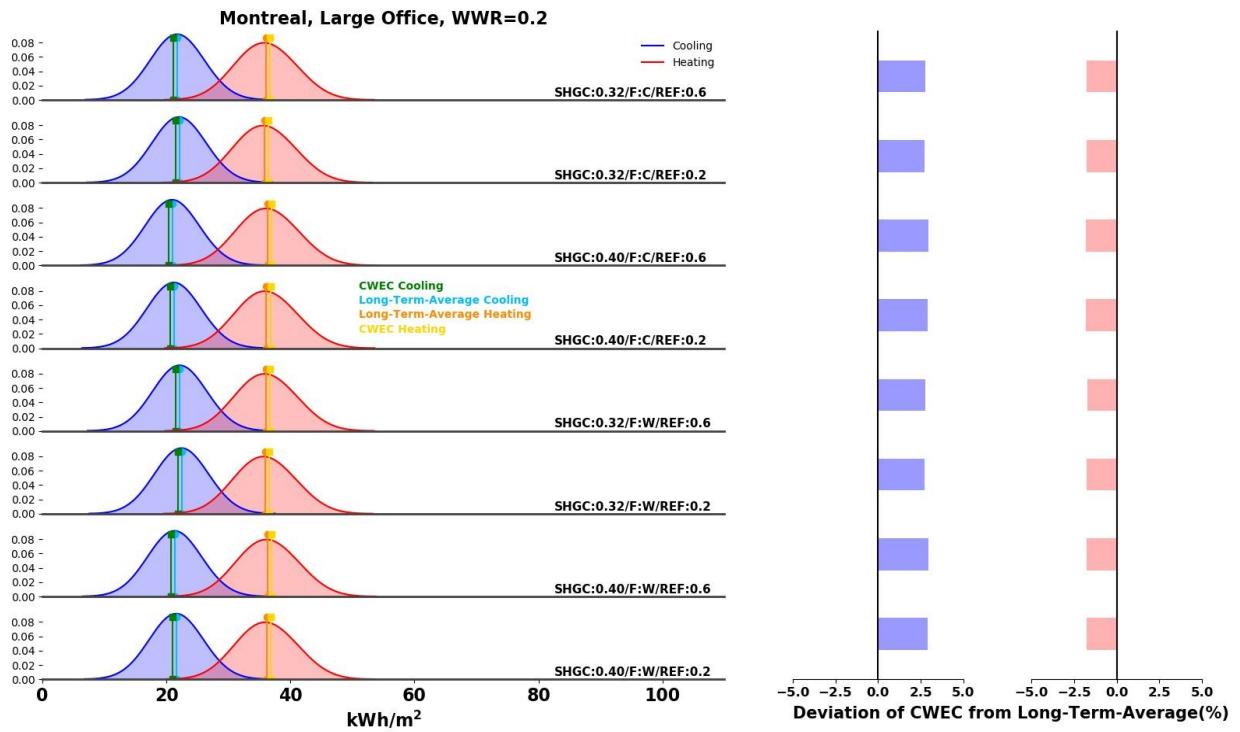


Figure 4.06: Montreal, large office with small windows

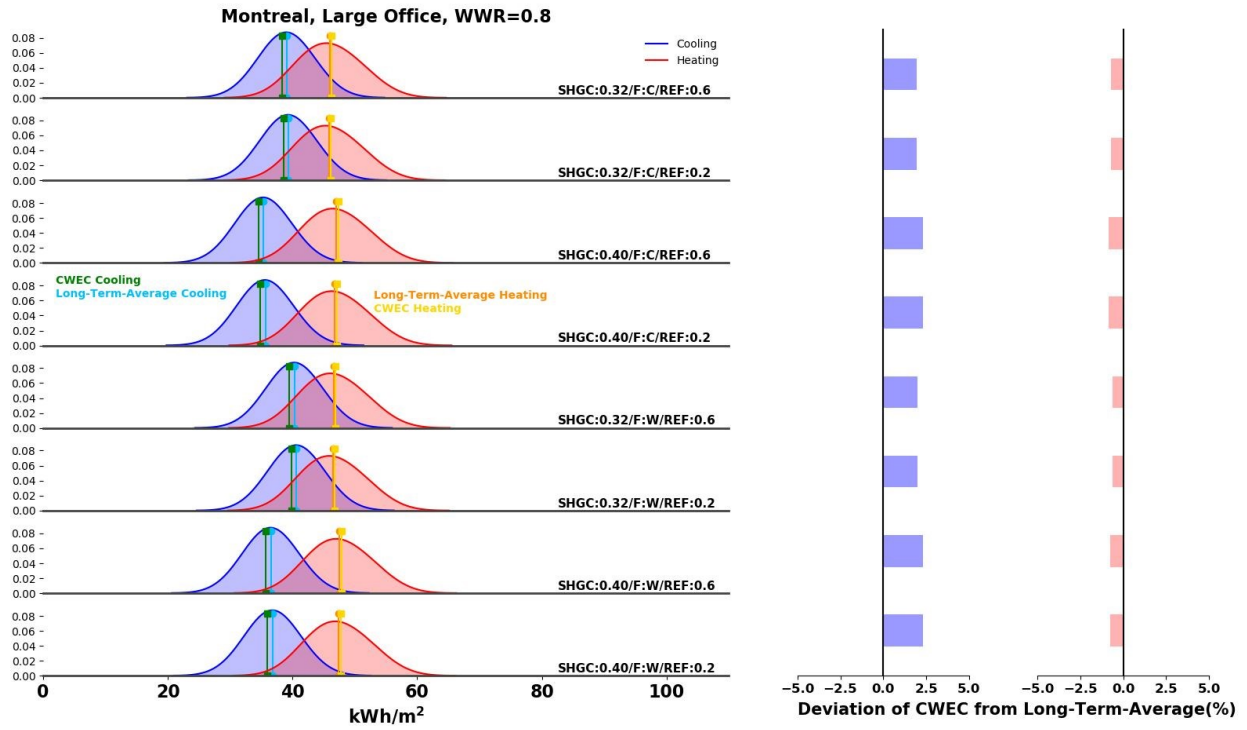


Figure 4.7: Montreal, large office with large windows

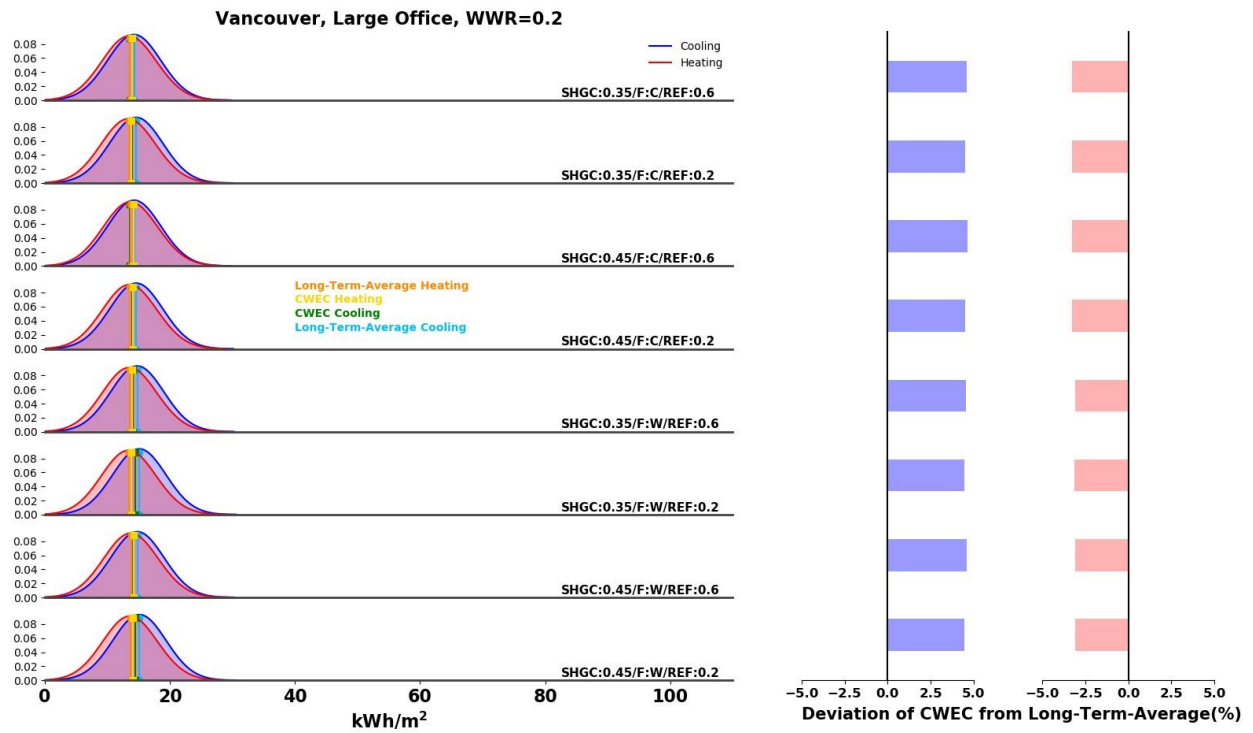


Figure 4.8: Vancouver, large office with small windows

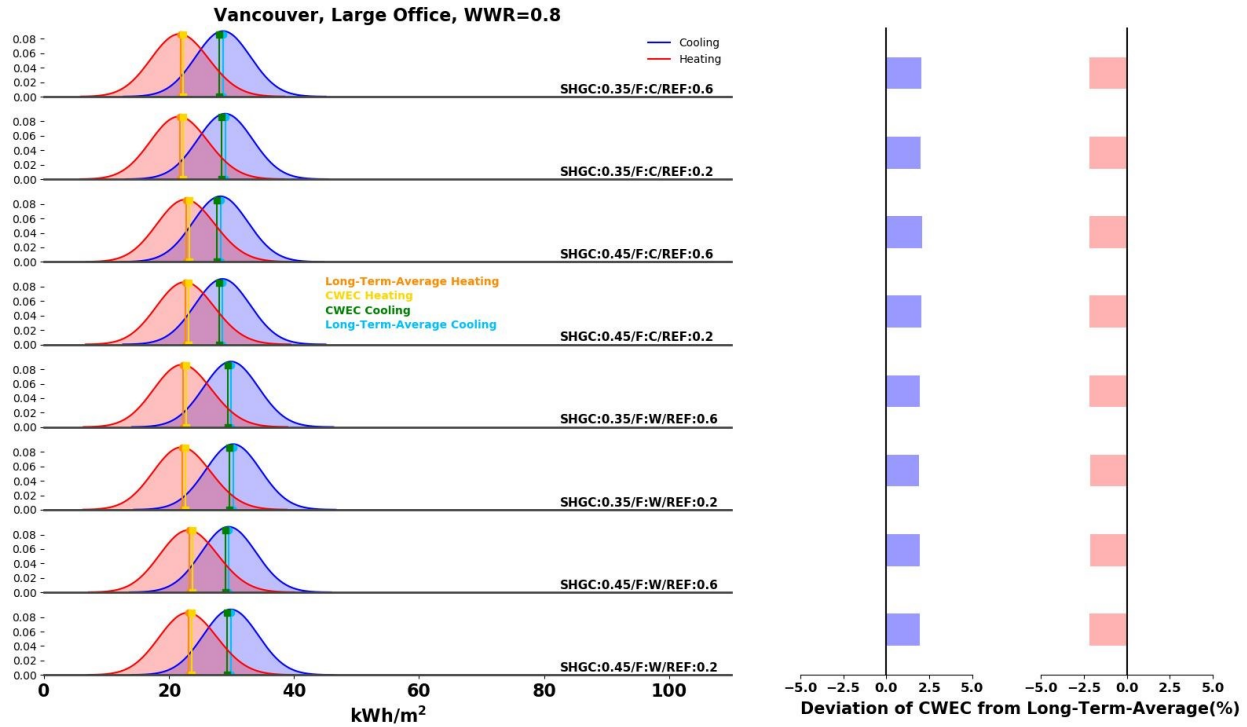


Figure 4.09: Vancouver, large office with large windows

Although the buildings in Vancouver exhibit a smaller range for both annual heating and cooling energy consumption, the deviation noted between the CWEC results and that of the long-term average is greater than for Montreal. The deviation of the CWEC results for cooling in Montreal ranges from slightly below 2% to 3%, while heating ranges from 0% to 2%. Vancouver exhibits slightly higher deviation values, which range from 2% to more than 4.5% for cooling and slightly below 2% to 3% for heating. This slight increase in deviation can be explained by the greater number of cloudy days and the smaller amount of solar radiation available in the region.

4.3.2 The effect of building type

As shown in Figures 4.2 to 4.5, the cooling energy consumption for both Montreal and Vancouver is smaller than the heating energy consumption for the typical use of a small office building in cold climate. However, in the Vancouver small office cases, the heating and cooling distributions are much closer, which is due to Vancouver's milder climate. Similarly, Figures 4.6 and 4.7 show larger heating energy consumption in comparison to the required cooling for a large office building in Montreal. In this case, the heating and cooling are closer in value compared to

the small office buildings. However, Figures 4.8 and 4.9 illustrate that for a large office building in Vancouver, the heating energy consumption is less than the cooling energy consumption, despite Vancouver being a cold climate region. This is a common phenomenon which occurs in many cold climates for large office buildings with large internal heat gain. The large office building contains datacenters on each floor, which contributes to a considerable amount of internal heat gain; the increase in internal heat gain leads to an increase in cooling consumption while reducing heating energy consumption. Considering the milder climate of Vancouver, the high internal heat gain makes the cooling larger than the annual heating.

The type of building also influences the deviation of the CWEC from long-term average results. Comparing Figures 4.2 to 4.5 with 4.6 and 4.9, a greater deviation is noted for the large office building cases. The largest deviation noted is for the large office building in Vancouver, with more than 4.5% deviation. In larger office buildings with a relatively large amount of internal heat gain, the total cooling or heating energy consumption can be more sensitive to fluctuations in internal gain, infiltration, and fresh air required in thermal zones.

4.3.3 The effect of design

Building design can play an important role in cooling, heating, and deviation of the CWEC results from that of the long-term average. In this study, the selected design options are parameters which have been considered as being mostly neglected in previous literature, building codes and standards. These parameters include WWR, window solar heat gain coefficient, floor type and roof and exterior wall solar reflectance, which can have a direct relationship with solar radiation in building energy performance. Design, building type and climatic conditions are related to each other considerably. In other words, a design's overall performance is dependent on climatic condition and building type. Figures 4.2 to 4.9 illustrate larger variations in performance for the small office buildings in comparison to the large office building cases. This variation in performance is due firstly, to the fact that small office buildings typically possess lower occupancy and lower internal heat gain and can therefore be more sensitive to the outdoor environment condition. Secondly, the total roof and exterior wall area in contrast to the conditioned area is greater than for large office buildings, which cause a larger heat transfer to occur through the building enclosure.

Figures 4.2 to 4.5 show that for the small office building, increasing the WWR from 0.2 to 0.8 increases the cooling energy consumption considerably; for Montreal and Vancouver an increase of about 20 kWh/m² is noted, while the reduction in heating energy consumption is about 10 kWh/m². The reasoning behind the heating energy saving being only half of the cooling increase is that the heating occurs during the Fall, Winter and part of Spring, when the sun angle is lower. On the other hand, cooling occurs at the end of spring and during the summer months, when the sun angle is higher, and therefore, allows for greater amount of solar gains. In other words, a higher amount of solar radiation is allowed into the thermal zones of the building, leading to a higher amount of cooling demand during this period. Moreover, a closer inspection of Figures 4.2 to 4.5 show that making use of solar gains has the potential to bring cooling and heating distributions closer. In this case, a lower roof and wall solar reflectance paired with high thermal capacity flooring and increasing the SHGC of the windows to allow for more solar radiation to enter in the thermal zones would allow to bring the distributions closer.

In addition, increasing the use of solar gains in the design reduces the deviation between the CWEC and long-term average results for cooling. A larger WWR makes the deviation of the CWEC results smaller. An example of this can be seen in Figure 4.3a, where the cooling deviation for the “WWR=0.2, window SHGC=0.45, concrete floor, solar reflectance=0.8” design is slightly less than 4%, while the same design with a larger WWR applied (“WWR=0.8, window SHGC=0.45, concrete floor, solar reflectance=0.8”) exhibits a deviation of 2% (Figure 4.5).

In another example, replacing the concrete floor with wood causes the deviation between the CWEC and long-term average results to increase for cooling and decrease for heating. This can be seen in Figure 4.4, where “WWR=0.2, window SHGC=0.45, concrete floor, solar reflectance=0.8” design has a cooling and heating deviation of about 4% and 2% respectively, while “WWR=0.2, window SHGC=0.45, wooden floor, solar reflectance=0.8” design increases the cooling value to 4.5% while decreasing the heating deviation to 1.5%. This change can be attributed to the fact that there is a large availability of solar radiation which becomes stored in the concrete floor, which behaves as a thermal mass. In cold climates, the solar radiation availability is quite different between seasons, where the solar gain potential is much greater during the summer in comparison to winter which has limited solar radiation. The concrete behaves as a thermal mass by storing the solar gains and releasing them later in the day. The application of a

thermal mass is useful during the cold winter months as a means to reduce the overall heating consumption, since the gains are stored and released later in the evening when the outdoor temperature begins to drop. However, for office buildings with typical operating hours, this benefit is negligible given that most occupants have vacated the building during this time. Furthermore, the limited solar radiation available during the winter months presents an issue for the concrete floor which has a high thermal capacity. In this given case, the deviation variation is most likely due to the high weighting factor (40%) attributed to solar radiation in the generation of the CWEC file. The large weighting factor indicates that solar radiation is supposed to play the most significant role in the thermal performance of a building. Therefore, if a building is less sensitive to the solar radiation according to its design, climate and use, there is an increase in uncertainty and the deviation of CWEC results from the long-term average results increases.

Furthermore, as Figures 4.6 to 4.9 show, when the type of building changes from a small office building to a large one, the influence of the selected design properties decreases, especially in the case of Vancouver. As described in section 4.3.2, for larger office buildings with larger internal gains, the cooling and heating energy consumption are less sensitive to the design of the building. The roof and exterior walls represent a smaller percentage of the total building area, therefore, energy efficiency measures such as solar reflectance or window SHGC would not be as effective. Even in the case of WWR, which is the most effective considered design parameter, there is little effect. Increasing the window area does allow for an increase in solar gains, however, the affected area represents only a small portion of the total floor area, reducing its effectiveness. In this case, other weather parameters such as temperature and wind might play more important roles, as they deal with infiltration, required fresh air and HVAC systems. As Figures 4.6 to 4.9 show, for large office buildings, no matter the design, cooling and heating deviation of the CWEC results for WWR=0.2 is about 3% and 2% in Montreal and 4.5% and 3% in Vancouver respectively (Figures 4.6 and 4.8). These numbers change to around 2% and 1% for Montreal and 2% for both cooling and heating in Vancouver if WWR = 0.8 is used (Figures 4.7 and 4.9).

Previously in the study, the historical weather data simulation results were compared to the CWEC simulation results. In order to further assess the applicability of CWEC in simulation, the peak demand for 98%, 99% and 99.6% were evaluated for small office buildings. These values were

selected following the ASHRAE Design Day Fundamentals, which uses these percentiles when evaluating temperature for HVAC design. Figure 4.10 and 4.11 show the percent difference cumulative distribution function, for select designs, between the peak load for each year of the historical weather data and CWEC simulation results for all above-mentioned percentiles. The x-axis represents the percent difference between the CWEC and historical results, where a negative percentage indicates that the CWEC results are below the historical weather data results. In other words, data points located to the left of the y-axis indicate underestimation of loading by the CWEC results, where the y-axis shows the proportion of years (a total of 30 years are represented).

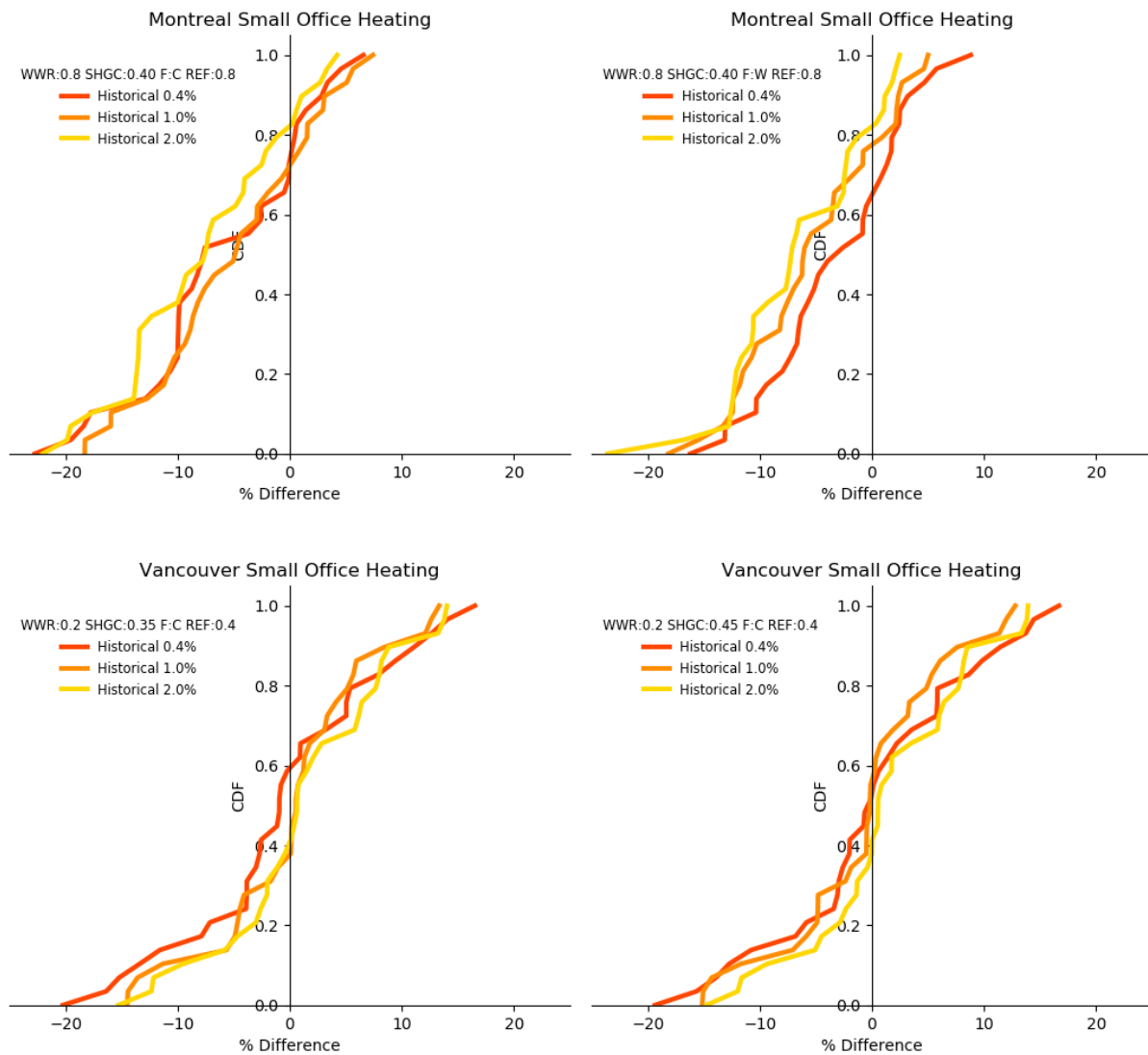


Figure 4.10: Percent difference cumulative distribution function for Montreal and Vancouver heating.

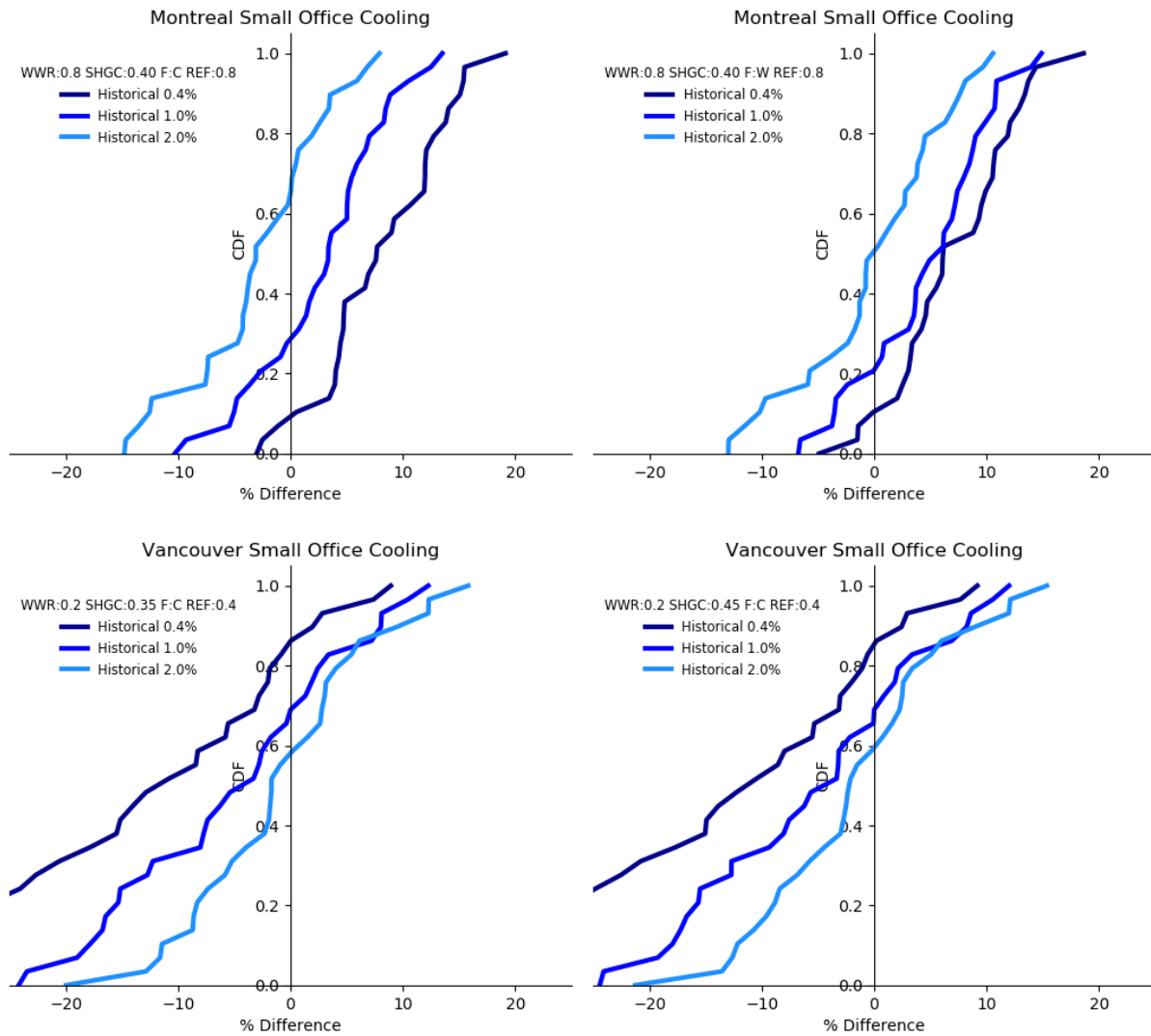


Figure 4.11: Percent difference cumulative distribution function for Montreal and Vancouver cooling.

The selected designs were among the cases which had the greatest overall deviation from the CWEC simulation results for heating (Montreal) and cooling (Vancouver). The corresponding heating and cooling figures for the same designs were then selected.

The graphs in Figures 4.10 and 4.11 illustrate the ratio of number of historical years where the CWEC simulation results underestimate the load. The Montreal cases ranged between 62.7% and 82.1% of the years having underestimated loads during the heating season for the 98th percentile. This represents 18 to 25 years, depending on the design, where the CWEC results are lower than

that of the historical weather data results. Vancouver designs exhibit a smaller number of years where the results are higher than that of the CWEC results, with the percentage ranging between 39.9% to 60.7%, namely, 12 to 18 years for the 98th percentile. During the heating season, designs which are considered sensitive to solar radiation fluctuations are more susceptible to be “underestimated” by the CWEC results. Designs sensitive to solar radiation have a greater number of years where the simulation results are higher than the CWEC simulation results. The discrepancy between the historical weather data results and those of CWEC may be attributed to the weighting factor applied during the CWEC file generation process. In the case where a greater importance is attributed to solar radiation in contrast to another parameter, for example temperature, the expected amount of information provided by the specified weather parameter is expected to be greater. In this case, when a design is sensitive to solar radiation fluctuations, the influence of solar radiation is increased, and a notable increase in “underestimated” years occurs.

During the cooling season, a larger percentage of years have results which are higher than CWEC for the Vancouver cases, with the exception of the 98th percentile which lies below Montreal; however, it exhibits a larger range in years (5 to 18 years). In all cases for the Vancouver cooling season, designs which are sensitive to solar radiation are subjected to a drop in percentage of years where the simulated results are above CWEC results. Thus, designs less sensitive to solar radiation cause an increase in the number of years which are “underestimated”. Montreal exhibits the same behavior as Vancouver, where designs with a small window-to-wall ratio and large solar reflectance exhibit results which fall above CWEC for a greater number of years. These observations align with those previously made for the heating season. For the cooling season, less sensitive designs restrict the quantity of information provided by the file, increasing unknowns. The reason that Vancouver experiences a greater number of “underestimated” years may be due to the amount of solar radiation the region receives, which is lower than Montreal. It appears that during the cooling season, when the information provided by solar radiation is greater, the frequency of “underestimated” years is decreased. The patterns shown during the heating a cooling season have led the authors to suspect that there may be an issue in the attributed importance of solar radiation (greater) and/or temperature (lower).

Although the deviation of the CWEC results from that of the long-term average results changes due to the climatic condition, building type and design, it's quite fairly negligible, such that it remains under 5%. However, the deviation can be reduced with updated weighting factors depending on the application.

4.4 Conclusion

A large-scale simulation was conducted in order to assess the impact on energy performance prediction between deploying a typical meteorological year (CWEC) file and actual meteorological year (AMY) files. The study was conducted using data for two cold climate cities, where two building types with varying design parameters were simulated. The effect of varying design parameters was further investigated in order to fill the void in existing studies which have ignored the selected parameters. These design parameters were among those neglected by standards and codes.

The results of the large-scale simulation are summarized below:

- The deviation of the results simulated with CWEC from the long-term average results for all the building types and designs were noted to be larger in Vancouver, where lower solar radiation is available.
- The large office buildings generally showed a higher deviation of the CWEC results from that of the long-term average.
- The designs considered less sensitive to solar radiation fluctuations generally showed higher deviation of results when comparing to the long-term average. For example, for a small office building in Vancouver, the design “WWR=0.2, window SHGC=0.35, wood floor, solar reflectance=0.6” showed a deviation of about 4.5% whereas, the design “WWR=0.8, window SHGC=0.45, concrete floor, solar reflectance=0.2” contributed to a deviation of less than 2%.
- A year-by-year evaluation of the 98th, 99th and 99.6th percentile of the heating season found that CWEC simulation results were lower than the historical data simulation results for a greater number of years for designs sensitive to solar radiation. The cooling season showed a greater discrepancy for designs less sensitive to solar radiation.

The building design and configuration may significantly impact the deviation of the CWEC results from that of the long-term average. This impact is especially significant when smaller buildings are considered, since they are often heavily influenced by minor changes in design parameters. In one case, a change in the WWR was able to reduce the deviation of the CWEC results by more than half. The study also noted a decrease in deviation between the CWEC results from the long-term average results for designs exhibiting higher sensitivity to solar radiation fluctuation. A high weighting factor of 40% for solar radiation during the generation of the CWEC file might exacerbate over/underestimation of the CWEC results. In addition, considering the lower amount of solar availability in cold climate and the fact that larger buildings are less susceptible to conduction and radiation, the solar radiation weighting factor of 40% might lead to an overestimation in energy prediction. Although, all the aforementioned factors play role in the deviation between the CWEC and long-term average results, all the deviations remained below 5%. This indicates that the CWEC may still be reliable to be used for estimation of cooling and heating energy performance. Furthermore, when comparing the percentile results on a year-by-year, the CWEC results were found to be below a large proportion of the years for the historical data simulation. However, modifications of the weighting factors might reduce the deviation of results. By doing so, the uncertainty can be reduced for cases with less solar radiation effect on building.

4.5 References

- [1] A. Ebrahimpour and M. Maerefat, "A method for generation of typical meteorological year," *Energy Convers. Manag.*, vol. 51, no. 3, pp. 410–417, 2010.
- [2] J. Sun, Z. Li, and F. Xiao, "Analysis of Typical Meteorological Year selection for energy simulation of building with daylight utilization," *Procedia Eng.*, vol. 205, pp. 3080–3087, 2017.
- [3] G. Chiesa and M. Grosso, "The influence of different hourly typical meteorological years on dynamic simulation of buildings," *Energy Procedia*, vol. 78, pp. 2560–2565, 2015.
- [4] S. Janjai and P. Deeyai, "Comparison of methods for generating typical meteorological year using meteorological data from a tropical environment," *Appl. Energy*, vol. 86, no. 4, pp. 528–537, 2009.
- [5] D. J. Thevenard and A. P. Brunger, "The development of typical weather years for international locations: Part I, algorithms," *ASHRAE Trans.*, vol. 108 PART 2, pp. 376–383, 2002.

- [6] I. J. Hall, R. R. Prairie, H. E. Anderson, and E. C. Boes, "Generation of a typical meteorological year," *Gener. a Typ. Meteorol. year*, 1978.
- [7] W. Marion and K. Urban, "Users manual for TMY2s: Derived from the 1961--1990 National Solar Radiation Data Base," *Users Man. TMY2s Deriv. from 1961--1990 Natl. Sol. Radiat. Data Base*, 1995.
- [8] S. Wilcox and W. Marion, *Users manual for TMY3 data sets*, no. May. National Renewable Energy Laboratory Golden, CO, 2008.
- [9] H. Lund, *The Design Reference Year Users Manual*. Thermal Insulation Laboratory, Technical University of Denmark, 1995.
- [10] D. L. Siurna, L. J. D'Andrea, and K. G. T. Hollands, "Canadian Representative Meteorological Year for Solar System Simulation.," in *Proceedings of the 10th annual conference of the Solar Energy Society of Canada (SESCI '84), Calgary*, 1984, pp. 85–88.
- [11] R. Festa and C. F. Ratto, "Proposal of a numerical procedure to select Reference Years," *Sol. Energy*, vol. 50, no. 1, pp. 9–17, 1993.
- [12] N. M. Sawaqed, Y. H. Zurigat, and H. Al-Hinai, "A step-by-step application of Sandia method in developing typical meteorological years for different locations in Oman," *Int. J. Energy Res.*, vol. 29, no. 8, pp. 723–737, 2005.
- [13] S. Tsoka, K. Tolika, T. Theodosiou, and K. Tsikaloudaki, "Evaluation of stochastically generated weather datasets for building energy simulation," *Energy Procedia*, vol. 122, pp. 853–858, 2017.
- [14] S. A. Kalogirou, "Generation of typical meteorological year (TMY-2) for Nicosia, Cyprus," *Renew. Energy*, vol. 28, no. 15, pp. 2317–2334, 2003.
- [15] F. Bre and V. D. Fachinotti, "Generation of typical meteorological years for the Argentine Littoral Region," *Energy Build.*, vol. 129, pp. 432–444, 2016.
- [16] A. L. S. Chan, "Generation of typical meteorological years using genetic algorithm for different energy systems," *Renew. Energy*, vol. 90, pp. 1–13, 2016.
- [17] A. L. S. Chan, T. T. Chow, S. K. F. Fong, and J. Z. Lin, "Generation of a typical meteorological year for Hong Kong," *Energy Convers. Manag.*, vol. 47, no. 1, pp. 87–96, 2006.
- [18] Y. Jiang, "Generation of typical meteorological year for different climates of China," *Energy*, vol. 35, pp. 1946–1953, 2010.

- [19] O. S. Ohunakin, M. S. Adaramola, O. M. Oyewola, and R. O. Fagbenle, "Generation of a typical meteorological year for north-east, Nigeria," *Appl. Energy*, vol. 112, pp. 152–159, 2013.
- [20] M. Petrakis *et al.*, "Generation of a 'typical meteorological year' for Nicosia, Cyprus," *Renew. Energy*, vol. 13, no. 3, pp. 381–388, 1998.
- [21] S. Pusat, I. Ekmekçi, and M. T. Akkoyunlu, "Generation of typical meteorological year for different climates of Turkey," *Renew. Energy*, vol. 75, pp. 144–151, 2015.
- [22] D. B. Crawley, "Which weather data should you use for energy simulations of commercial buildings?," *ASHRAE Trans.*, vol. 104, no. 2, pp. 498–515, 1998.
- [23] A. M. Realpe, C. Vernay, S. Pitaval, C. Lenoir, and P. Blanc, "Benchmarking of Five Typical Meteorological Year Datasets Dedicated to Concentrated-PV Systems," *Energy Procedia*, vol. 97, pp. 108–115, 2016.
- [24] D. B. Crawley and L. K. Lawrie, "Rethinking the tmy: Is the 'typical' meteorological year best for building performance simulation?," in *14th International Conference of IBPSA - Building Simulation 2015, BS 2015, Conference Proceedings*, 2015, pp. 2655–2662.
- [25] S. Guo, D. Yan, T. Hong, C. Xiao, and Y. Cui, "A novel approach for selecting typical hot-year (THY) weather data," *Appl. Energy*, vol. 242, no. November 2018, pp. 1634–1648, 2019.
- [26] L. Yang, J. C. Lam, J. Liu, and C. L. Tsang, "Building energy simulation using multi-years and typical meteorological years in different climates," *Energy Convers. Manag.*, vol. 49, no. 1, pp. 113–124, 2008.
- [27] T. Hong, W. K. Chang, and H. W. Lin, "A fresh look at weather impact on peak electricity demand and energy use of buildings using 30-year actual weather data," *Appl. Energy*, vol. 111, pp. 333–350, 2013.
- [28] Y. Cui, D. Yan, T. Hong, C. Xiao, X. Luo, and Q. Zhang, "Comparison of typical year and multiyear building simulations using a 55-year actual weather data set from China," *Appl. Energy*, vol. 195, pp. 890–904, 2017.
- [29] M. Hosseini, B. Lee, and S. Vakilinia, "Energy performance of cool roofs under the impact of actual weather data," *Energy Build.*, vol. 145, pp. 284–292, 2017.
- [30] M. Hosseini, F. Tardy, and B. Lee, "Cooling and heating energy performance of a building with a variety of roof designs; the effects of future weather data in a cold climate," *J. Build. Eng.*, vol. 17,

no. January, pp. 107–114, 2018.

- [31] United States Department of Energy. (DOE), “Commercial Prototype Building Models.” Accessed April 4, 2019: https://www.energycodes.gov/development/commercial/prototype_models.
- [32] National Research Council of Canada. (NRC), “National Energy Code of Canada for Buildings (NECB),” 2011.

Chapter 5. A SYSTEMATIC APPROACH IN CONSTRUCTING TYPICAL METEOROLOGICAL YEAR WEATHER FILES USING MACHINE LEARNING

The single typical meteorological year file is composed of hourly resolution data from the most 12 representative calendar months of 30 years which are selected based on statistical similarity to long-term weather daily-averaged data. These weather files are synthetically constructed on historical weather data over a long period of time for an array of weather parameters, such as solar radiation, temperature, wind speed and others. The statistical procedure to construct the weather files depends on the weights assigned to these weather parameters. Under current practice, these weighting factors are universally assigned regardless of climatic locations nor the building application. This approach leads to energy performance predictions that deviate from the long-term averages. This chapter introduces a novel machine learning algorithm to extract the feature importance of the weather parameters in order to assign weighting factors straightly proportional to their impacts on energy performance of buildings. Weather files were constructed with these systematically assigned weighting factors which are climatic location dependent. The typical meteorological year weather file based on the novel approach offers better prediction (with statistical significance) on energy performance for certain climatic As a result, the suggested method avoids potential under/oversizing of equipment and promotes energy conservation.

This chapter is published in Energy and Buildings, Volume 226, 110375, M. Hosseini, A. Bigtashi, B. Lee, “A systematic approach in constructing typical meteorological year weather files using machine learning”, © Elsevier Ltd, 2020

5.1 Introduction

Although typical meteorological year weather files are meant to be a representation of long-term weather, some studies have noted deviations between energy performance simulated with the typical meteorological weather file and that through measurement under long-term actual weather. A study by Hosseini et al. [1] (previous chapters) demonstrated that the deviation could be quite considerable depending on the designs of buildings. On the other hand, another study conducted by Janjai and Deeyai [2] showed minor differences between the long-term average energy

performance and the corresponding energy performance simulated with typical meteorological year weather files. However, the study focused on the simulated energy performance of solar related technologies, such as water heating systems and photovoltaic systems, without considering the energy performance of the whole building. The limited scope might explain the negligible deviation between the long-term average energy performance and that predicted with typical meteorological weather files.

In building energy simulation, typical meteorological year weather data is usually a synthesized single year of weather data that represents multiple years of historical weather data over a period of thirty years, if not more, from which typical meteorological weather files are created. There are a few existing methods used to generate the typical meteorological data including, but not limited to, the Sandia National Laboratory method [3], the Danish method [4], the Festa and Ratto method [5] and the ISO 15927-4 standard [6].

The Sandia method is one of the most frequently used methods, where the Finkelstein-Schafer (FS) [3] statistic is calculated to determine typical weather months of a year. Here is a brief description of a generic generation procedure to generate typical meteorological year weather file in CWEC format. The FS statistic is used to find the absolute value of the difference between long-term data and each of the historical candidate months. In other words, for each weather data month, all historical months are evaluated and the month which matches most statistically to the long-term weather pattern is selected. The procedure is summarized below.

For each month of a year:

- i. The cumulative distribution function (CDF) of each of the 9 weather parameters, namely, horizontal solar radiation, average temperature, maximum temperature, minimum temperature, average dew point temperature, maximum dew point temperature, minimum dew point temperature, average wind speed, and maximum wind speed are calculated.
- ii. The FS statistic is calculated for each parameter according to equation 1:

$$FS = (1/n) \sum_{i=1}^n \delta_i \quad (1)$$

Where: δ is absolute difference between the long-term data CDF and the historical candidate month data CDF, and n is the number of readings in a month (usually 30, representing a 30-year period).

- iii. Weighting factors of the weather parameters are commonly assigned based on experts' judgments according to their importance on building energy demand. The weighting factors from previous studies are summarized in Table 5.1 for different weather file types. The month with minimum weighted sum of FS displays the most similar weather pattern to the long-term historical weather that is selected according to equation 2, where the weighted sum is calculated for each candidate month:

$$WS = \sum w_i FS_i \quad (2)$$

Where: w_i is the weighting factor attributed to each parameter (shown in Table 1), and FS_i is index FS statistic.

- iv. The hourly weather data belonging to the minimum WS is used to fill up the corresponding month of the twelve-month weather file.

The same procedure will be repeated for each month of the representative year until all the 12 months of the year are selected.

Table 5.1: Weightings factors of weather parameters in previous studies

	TMY[3]	TMY2 & 3 [7,8]	IWEC & CWEC[9]	(Kalogirou, 2003)[10]	(Chan, 2016) [11]
Maximum dry-bulb temperature	0.042	0.050	0.050	0.031	0.061
Minimum dry-bulb temperature	0.042	0.050	0.050	0.031	0.003
Mean dry-bulb temperature	0.083	0.100	0.300	0.063	0.258
Standard deviation of dry-bulb temperature	-	-	-	0.031	-
Maximum dew point temperature	0.042	0.050	0.025	-	0.106
Minimum dew point temperature	0.042	0.050	0.025	-	0.008
Mean dew point temperature	0.083	0.100	0.050	-	0.017
Maximum relative humidity	-	-	-	0.031	-
Minimum relative humidity	-	-	-	0.031	-
Mean relative humidity	-	-	-	0.063	-
Standard deviation of relative humidity	-	-	-	0.031	-
Maximum wind speed	0.083	0.050	0.050	0.031	0.146
Mean wind speed	0.083	0.050	0.050	0.063	0.082
Standard deviation of wind speed	-	-	-	0.031	-
Air pressure	-	-	-	-	-
Global horizontal solar radiation	0.500	0.250	0.400	0.250	0.319
Direct normal solar radiation	-	0.250	-	0.250	-

As presented in Table 5.1, depending on weather file formats, the weighting factor attributed to each weather parameter may vary slightly. Global horizontal solar radiation always carries the most weight regardless of the file formats. Table 5.2 provides list of terms used in this study.

Table 5.2: List of terms used in figures, tables, and texts

Terms	Abbreviation	Explanation
Maximum dry-bulb temperature	max_temp	Maximum dry-bulb temperature in a day (°C)
Minimum dry-bulb temperature	min_temp	Minimum dry-bulb temperature in a day (°C)
Mean dry-bulb temperature	temp	Mean daily dry-bulb temperature (°C)
Maximum dew point temperature	max_dew	Maximum dew point temperature in a day (°C)
Minimum dew point temperature	min_dew	Minimum dew point temperature in a day (°C)
Mean dew point temperature	dew_temp	Mean daily dew point temperature (°C)
Maximum wind speed	max_wind	Maximum wind speed in a day (m/s)
Mean wind speed	wind	Mean daily wind speed (m/s)
Solar radiation	rad	Mean daily global horizontal solar irradiance (W/m ²)
Cooling demand	cooling	Mean daily cooling energy demand (kWh)
Heating demand	heating	Mean daily heating energy demand (kWh)
Total demand	total	Total of daily average cooling and heating energy demand (kWh)

An important issue with the current typical meteorological weather file generation methods is the subjectivity behind attributing weighting factors. This process relies on experts' judgments in order to determine the importance of each weather parameter on the performance of the system. Furthermore, the level of importance attributed to each weather parameter is designated regardless of the climatic location or the application. Differences in weighting factors not only affect building energy performance as shown by the previous studies but may also lead to changes in the composition of typical meteorological weather files. In the study conducted by Georgiou et al. [12] a change in the weighting factors showed a considerable change in the selected months for the typical year, which led to a smaller difference between long-term average and actual years. The study conducted a simulation for cooling and heating energy consumption of a building, as well as that for the solar collector, and wind turbine. The study used the arbitrary weighting factors and concluded that weighting factors should be optimized based on intended use of the weather file to enhance "*precision and accuracy*" of simulation results.

A few studies applied optimization and sensitivity analysis to find the importance of each weather parameter with respect to a specific performance. Kalamees et al. [13] conducted a sensitivity analysis to weigh the weather parameters by changing the monthly average value of the parameters

for each month and observing the change in monthly cooling and heating demand of a building through simulations for climate zones 1 to 4. Chan [11] used genetic algorithms to optimize the weighting factors according to the application. The study focused on four primary applications, an air-conditioned building and three renewable energy systems, including a building integrated (attached) photovoltaic (BIPV or BaPV) system, a wind turbine and a concentrated solar power generation (CSP) system. Their results showed a very good agreement between the long-term average and the simulation results with the optimal weighted typical meteorological year weather file. However, the genetic algorithm used in this study requires a coupled optimization algorithm with building energy simulation (EnergyPlus in this case) for new mutation. In other words, the procedure would be based on trial and error which requires multiple time generations of new sets of weighting factors that make a new weather file. This is followed by energy simulation to minimize the uncertainty between long-term average of the output and the output of the corresponding weather file in each mutation stage. Therefore, this procedure can be fairly time-consuming as it requires numerous simulations. With the demonstrated uncertainty associated with experts' judgments based weighting factors, it is therefore paramount to develop a systematic and yet faster approach in properly assigning weighting factors to weather parameters according to their importance to the performance of specific applications (different sets of weighting factors for investigations of different types of performance).

In data mining and machine learning, a similar concept is applied to find how useful a variable is at predicting a target variable (feature importance). The Sandia method is a K-nearest neighbor machine learning algorithm for classification which is currently using experts' judgments for defining weighting factors. Applying another supervised regression machine learning algorithm can fit the data and extract the features importance, allowing for systematic selection of weighting factors, which can then be introduced to the FS. Applying the machine learning algorithm requires the simulation to run only once and won't require repetition of simulations, which saves considerable processing time.

Depending on the data and type of problem, various machine learning algorithms can be used to find the feature importance. In its simplest form in linear models, such as linear regression or logistic regression, the importance of each variable is simply the weight or coefficient for each input variable [14]. In fact, these are the weights that are used for prediction. Non-linear, individual

models such as decision tree and its ensemble model, Random Forest, find the importance of variables based on the variable's role in reducing the variance in the output [14, 15]. The Random Forest algorithm is particularly effective in attending to regression models [16] since the algorithm considers a small but random subset of features at each level and avoids the potential bias that could be introduced if the dataset contains a few predominant features.

The current study introduces a novel method based on machine learning to systematically determine weighting factors in proportion to feature importance of weather parameters according to the climate and application in which the typical meteorological year weather file is applied. In the case of this study, the Canadian Weather year for Energy Calculation (CWEC) typical meteorological year weather file was selected, with the period ranging between 1960 and 1989. The CWEC file applies the same set of weighting factors used in the original International Weather year for Energy Calculation (IWEC) [9]. The proposed novel method was applied to a small prototypical office building to improve the quality of CWEC and reduce the deviation from the long-term average. The reduction in deviation is evaluated using the root mean square error (RMSE). For the purpose of this study, a small office building was chosen due to its higher surface to volume ratio; that is, it is more sensitive to weather conditions in comparison to a large building. It must be noted that the methodology can be applied to other types of buildings and applications as well.

5.2 Methodology

5.2.1 Weighting factors

The purpose of weighting factors in the generation of typical meteorological year weather files as suggested by Wilcox [8] is to reduce the uncertainty, in this case deviation from the average, caused by representing multiple years by one singular year. The value attributed to each weighting factor depends on the importance of the selected weather parameter. The experts' judgment in assigning values to weighting factors is to be replaced by machine learning algorithm through a systematic approach proposed in this study. Weather data are pre-processed to determine the feature importance. Feature importance in this case is the importance of each weather parameter on the building energy demand; that is, the amount of energy demand that can be attributed to each weather parameter. The identification of such importance will help reduce the uncertainty in building energy performance as it makes the FS procedure (in Sandia method) more adaptive to

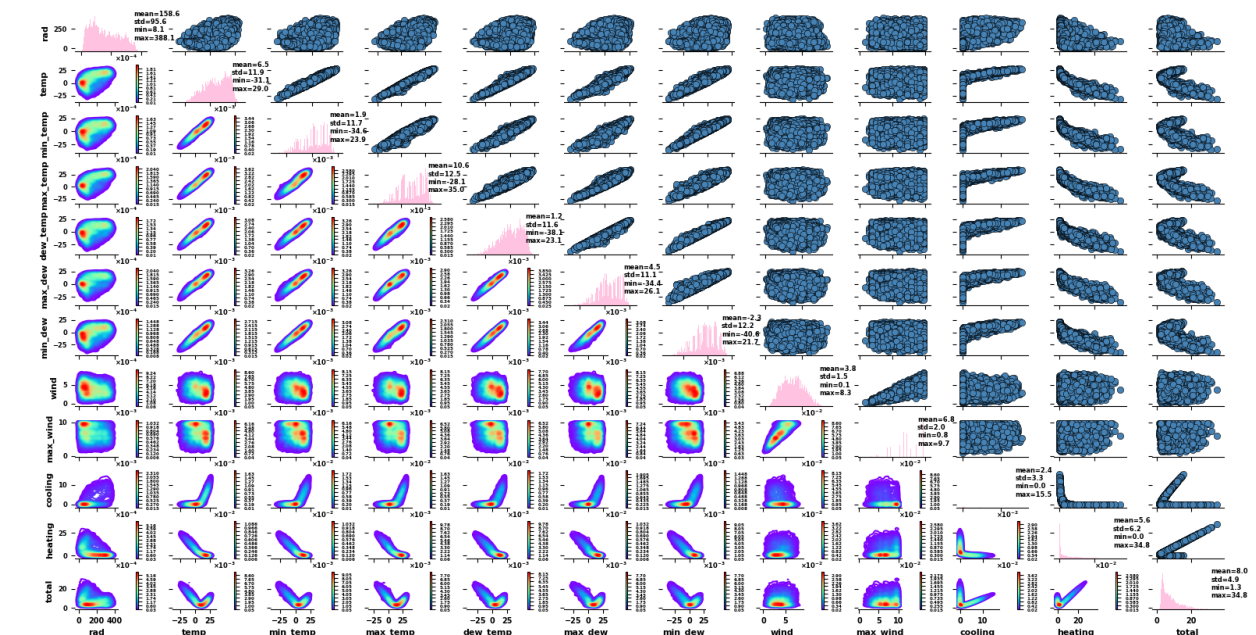
the problem [17] which will become apparent later in the discussion. The building energy demand is climate and application specific. To demonstrate the applicability of the proposed approach, the case study building was based upon DOE prototypical small office building [18] with the building envelope adapted to the specification of the National Building Energy Code of Canada [19] and was simulated with EnergyPlus (version 8.9) simulation program. The case study building was studied for two locations— Montreal and Vancouver. Thirty-year actual historical weather data at the respective airports, ranging from 1960 to 1989, were used. Energy performance was evaluated in terms of hourly cooling, heating, and total energy demand for the thirty years. The methodology is explained in the following sections.

The datasets are resampled to daily average data as the FS (equation 1) are calculated over daily-averaged resolution data. Figures 5.1 and 5.2 present the scatter plot, histogram and the projection of distribution of the 30-years historical daily-averaged weather parameters and the corresponding total energy demand for Montreal and Vancouver respectively. The upper diagonal section shows the scatter pair plot of the variables; the diagonal midsection shows the histogram of data for each variable together with their corresponding mean, standard deviation, maximum, and the minimum. Finally, the lower diagonal section shows the projection of distribution of pair data. There are three groups of correlations that can be observed in Figures 5.1 and 5.2. The first group relates the 9 weather parameters among themselves. For example, from the data of Montreal, it can be observed that the temperature parameters are highly correlated among themselves, mildly correlated with radiation, and not quite correlated with wind. The second group relates the 3 energy performance parameters with the 9 weather parameters. There exists a high linear correlation between total energy demand and temperature, in contrast to wind speed, which exhibits a non-linear relationship with the total energy demand. The third group relates the energy performance parameters between themselves. In this case, a higher correlation is noted between total energy and heating energy for Montreal, which is explained by the cold climate.

From Figure 5.1, it can be observed that dry-bulb temperatures (max, mean, min) and the dew point temperature are linearly correlated in Montreal, whereas from Figure 5.2, the correlation is not clear in Vancouver. The more frequent humid weather in Vancouver helps explain the situation. Only heating energy demand is more or less linearly correlated to temperatures. All other forms of energy demand and weather parameters are non-linearly correlated. Although both

locations are considered cold climates, there are different correlations and relative importance among the weather parameters and between weather parameters and the energy demands.

The bottom of Figure 5.1 displays two samples of the figures in a larger size for better resolution. The two high resolution kernel density estimate (KDE) diagrams present the probability density function of a continuous variable, in this case solar radiation and temperature respectively, in relation to total energy demand. The warmer colors indicate a higher density function. From the figure, in Montreal, dry-bulb temperatures within the range of 10 to 15 °C leads to minimum total energy demand whereas the values larger or smaller than this range increases the total energy demand as cooling and heating energy demand increases, respectively. The figure shows that the data is denser (more frequent) within the range of about 7 to 20 °C and 5 to 10 kWh (red region). During the heating season, a total energy demand of up to 30 kWh/m² is noted in instances where the solar irradiance is less than about 130 W/m². A solar irradiance higher than 200 W/m² leads to a total energy demand of less than 15 kWh/m² which occurs mostly during the cooling season. The red region in the density plot suggests that the most frequent solar irradiance is within the range of around 50 to 200 W/m² which is associated with a total energy demand of about 5 kWh/m².



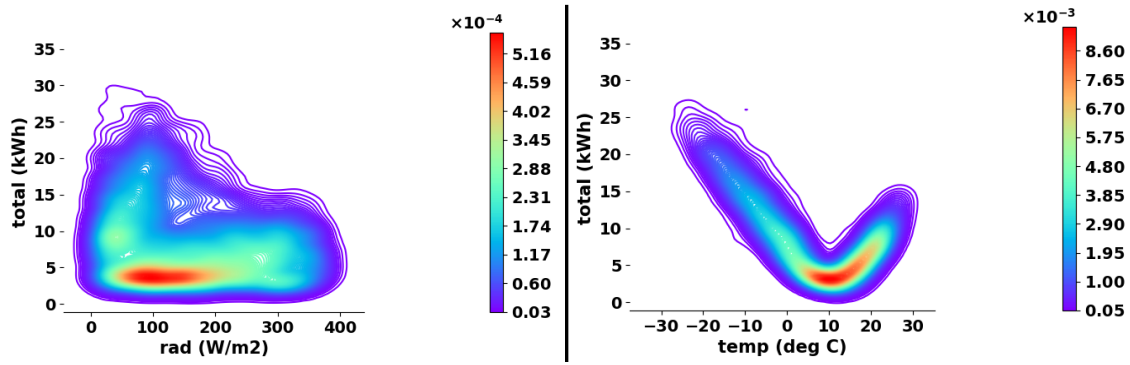


Figure 5.1: The distributions of 30 years historical weather data together with corresponding energy demands of the small office building in Montreal; the bottom figures are two samples of the above figures for better resolution.

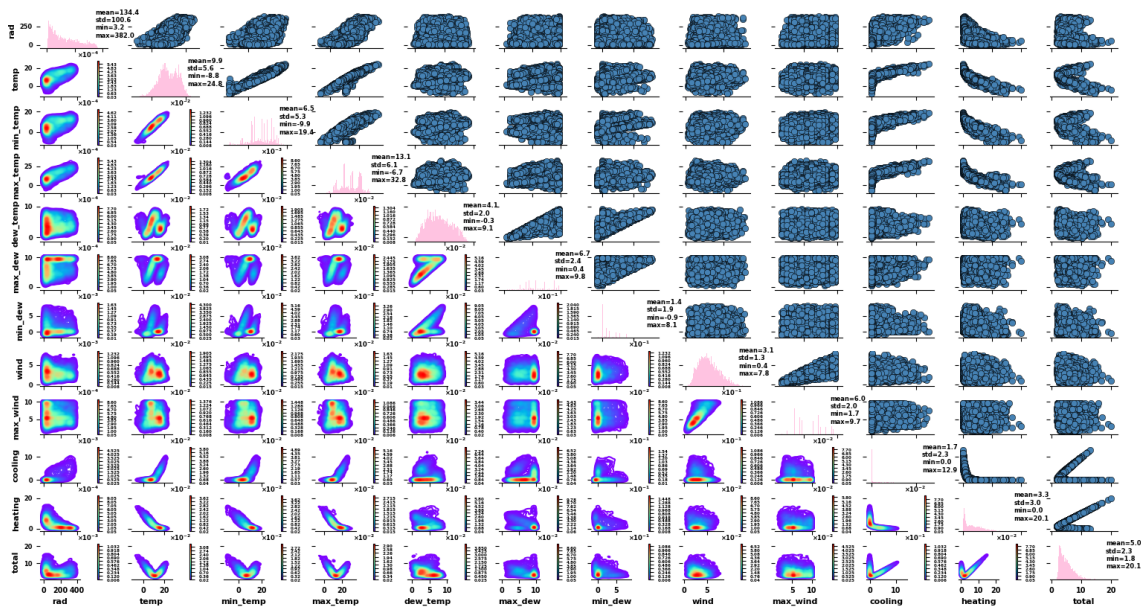


Figure 5.2: The distributions of 30 years historical weather data together with corresponding energy demands of the small office building in Vancouver.

The correlation between the dry-bulb and dew point temperatures are stronger in Montreal than in Vancouver which is explained by the difference in relative humidity levels in the two cities, as mentioned above. Some variables exhibit almost linear relationships, such is the case for temperature, min/max daily temperature and the total energy demand, as well as wind speed and maximum wind speed in a day, while others have non-linear relationships. The relationship

between variables is important to consider in the process of selecting an appropriate model to capture non-linearity of the relationship among the independent and dependent variables.

5.2.1.1 Random Forest algorithm and corresponding feature importance

Due to its simplicity and lower hyperparameters (parameters in machine learning algorithms for more accurate and robust results; for example number of trees in the forest) involved in the model, the Random Forest regression algorithm is used to statistically find relationships between the attributes and the dependent variables, as it can capture the non-linear relationship among the variables. The energy demands are the dependent variables, while the nine weather parameters are considered as continuous independent variables. A Random Forest model includes many regression trees splitting a dataset into smaller and smaller datasets on decision nodes while the tree develops. The classification and regression tree (CART) model [20] used in this study is a binary tree, meaning that each decision node representing an input variable splits into two branches and the splitting continues until it reaches the leaf nodes, where the cost function reaches the minimum error between the prediction and the actual result. The leaf nodes have output values (this can be cooling, heating, or total energy demands) that can be used for prediction. In a simple explanation, the whole algorithm recursively finds the variable and the corresponding value which has the greatest potential to reduce the variance of the dataset and splits it at that value [20]. This way the searching will have the lowest cost.

In the case study, there are nine weather parameters which are considered as variables. The algorithm evaluated all the variables iteratively. The algorithm offered by Breiman [14] is used in this study. Figure 5.3 shows how a regression tree of the Random Forest model works for the multivariate problem. At every node, all the weather parameters are evaluated with their values to find the best split point where the corresponding prediction shows minimum error. The mean square error is defined as equation 3:

$$\text{MSE} = \frac{\sum_{i=1}^n (y_i - \bar{x})^2}{n} \quad (3)$$

Where: y_i is the corresponding total energy demand, cooling or heating of each level of a weather parameter in the sample; \bar{x} is the mean of all the corresponding total energy demands in the sample (predicted energy demand of the sample), and n is the number of samples.

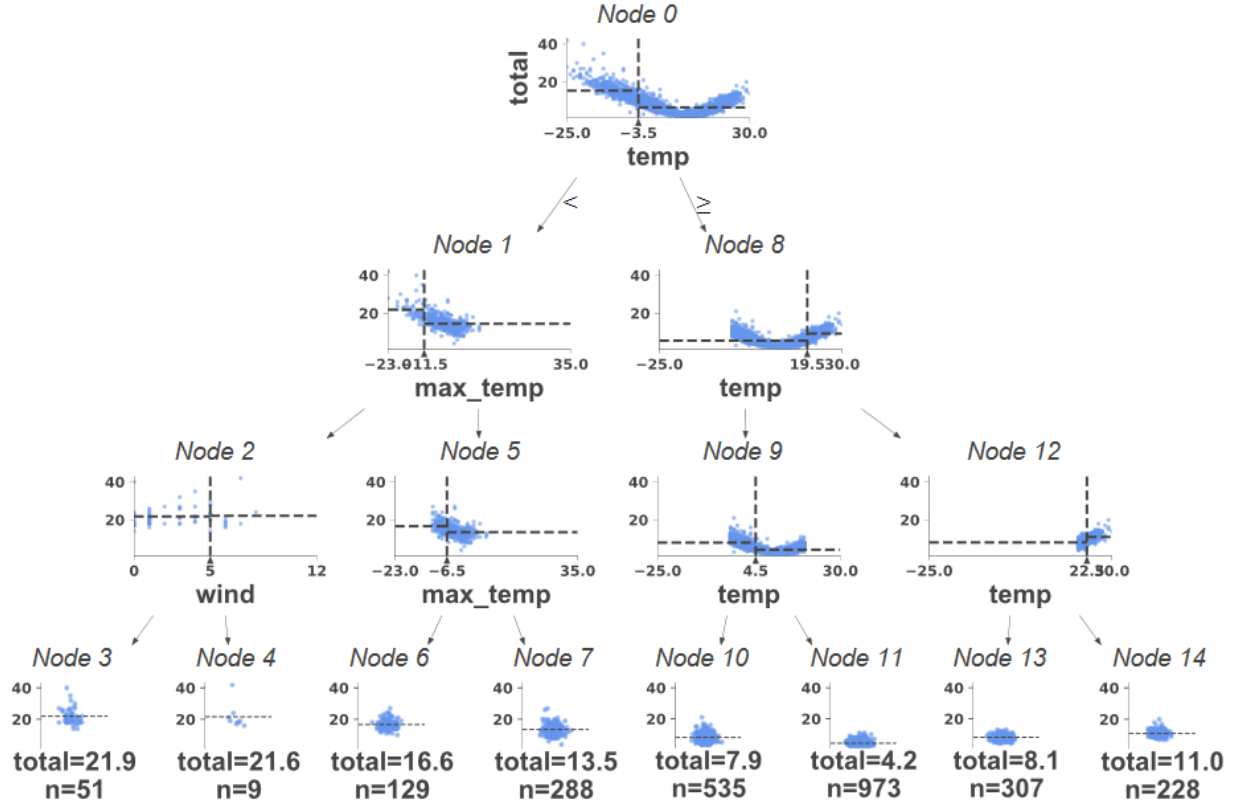


Figure 5.3: A tree of Random Forest; the blue scatters show the data; the vertical dash are the split points of the variables, and the horizontal dash lines show the prediction of total energy demand after splitting. The number of depths are reduced to 3 for better presentation.

Since each node must be divided into two other nodes, MSEs must be weighted according to the number of samples. Therefore equation 4 is used to weight the MSE of feature levels as follow:

$$MSE_w = \frac{n_l * MSE_l + n_r * MSE_r}{n} \quad (4)$$

In each node, MSE_w is calculated for all the values of each weather parameter (features) and the minimum value is selected as the split point of the node; the split point divides the parent node into two child nodes on the left and the right. n_l and n_r of equation 4 are the numbers of left and right samples (child nodes), MSE_l and MSE_r are the MSE of left and right samples (child nodes), and n is the number of parent node samples. This process persists until the variance is reduced to approximately zero.

Figure 5.3 shows a single tree of the Random Forest displaying only three levels of depth for better illustration; for the first node of the original dataset, average temperature is the best feature to split, at -3.5 °C. If the temperature is below -3.5 °C, the predicted total energy demand was about 16 kWh, as defined by the dashed horizontal line from Figure 5.3 (Node 0). If the temperature is above -3.5 °C, the predicted total energy demand was about 8 kWh. For each split, two branches are created with opposing ranges, where the units are defined by the defined parameter. The range is defined by the split point, and each branch has an associated predicted energy demand. This procedure continues till the terminal node (leaves) are reached, where the MSE will be almost zero. When the tree is constructed, an unseen data traverses the entire tree till it reaches the leaves where it can find the predicted value. The details are further explained below and with equations 5 – 8.

An interesting feature of the tree is the calculation of the importance of each of the weather parameters in predicting the total energy demand. At every node j of a decision tree, the importance of the node, denoted by ni_j , is evaluated to split the dataset into two child nodes:

$$ni_j = MSE_j - MSE_{wlj} - MSE_{wrj} \quad (5)$$

where $MSE_j = MSE$ at node j , $MSE_{wlj} = MSE_w$ of child node from left split on node j , $MSE_{wrj} = MSE_w$ of child node from right split on node j .

Then the importance for each feature is calculated as:

$$fi_i = \frac{\sum_{j: \text{node } j \text{ splits on feature } i} ni_j}{\sum_{k \in \text{all nodes}} ni_k} \quad (6)$$

Where: fi_i is the importance of feature i , and ni_j is the importance of node j for a single tree.

The feature importance can be divided by the sum of all feature importance values in order to normalize to a value between 0 and 1.

$$normfi_i = \frac{fi_i}{\sum_{j \in \text{all features}} fi_j} \quad (7)$$

Finally, the average feature importance of each feature is calculated based on the number of trees in the forest.

$$RFfi_i = \frac{\sum_{j \in \text{all trees}} normfi_{ij}}{T} \quad (8)$$

Where: $RFfi_i$ is the average feature importance i from all the trees of Random Forest, $normfi_{ij}$ is the normalized feature importance of feature i in tree j , and T is the number of trees in the Random Forest model.

As mentioned above, Random Forest is an ensemble model which contains multiple decision tree models. As a data-driven model, trees have randomness nature and are trained on random subsets of the dataset. Therefore, the output of the model and the feature importance can slightly change when new models are trained. As the number of the trees in the Random Forest model increases to a high number, the feature importance converges to a constant value as they are averaged from multiple trees feature importance. Therefore, a larger number of trees, in this case 1000, is considered in this study in order to have a stabilized result.

5.2.2 Application of Random Forest: feature importance extraction

The algorithm of Random Forest and the method of feature importance calculations are described in the previous section. As a summary of the methodology, Figure 5.4 shows the flowchart of the proposed method in comparison to the current Sandia method.

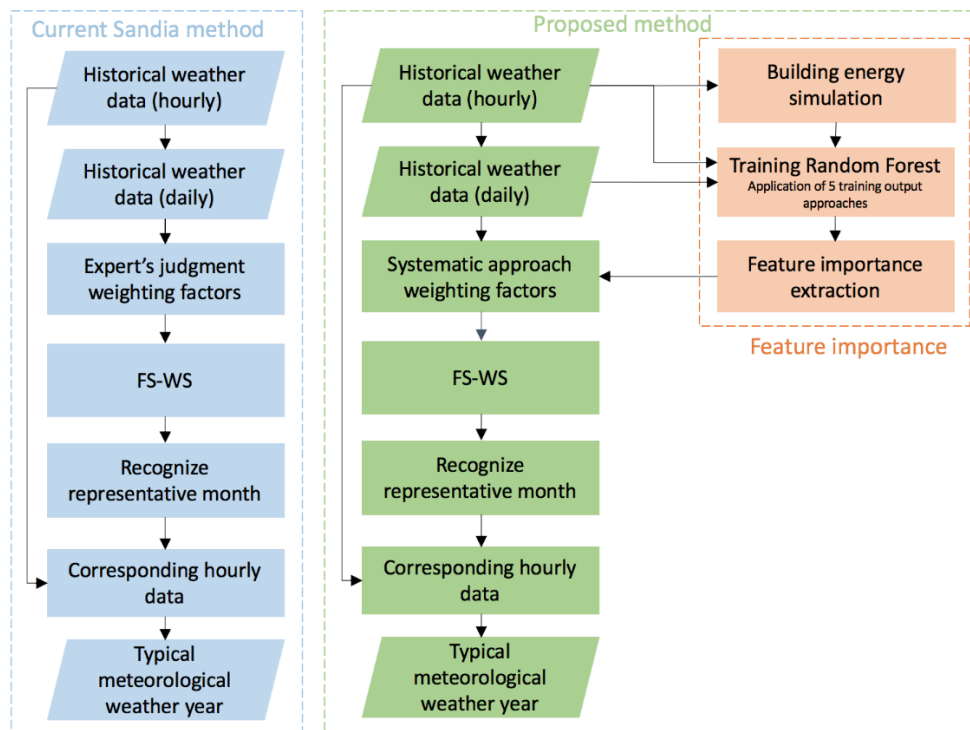


Figure 5.4: Contrasting the steps taken under the current Sandia method with those of the proposed method and indicating where the additional steps of random forest modeling takes place and supports the systematic weighting factors assignment.

The left side of the flowchart shown in blue is the current Sandia method which is explained in the introduction. The right side shown in orange shows the process in identifying the feature importance of the weather parameters. The section in green presents how the process is incorporated into the Sandia method. Thirty years of hourly weather data (from 1960 to 1989) are fed into the building energy simulation program (EnergyPlus, in this study, as described under section 5.2.1) to generate the hourly cooling and heating energy demands. The hourly outputs are converted to daily average values. The daily data, together with the daily-average weather data (from daily averaged dataset), maximum and minimum temperature in a day (from hourly dataset), maximum and minimum dew point in a day (from hourly dataset), and maximum wind speed in a day (from hourly dataset) and hourly solar irradiance are then used to train the Random Forest model. The trained data is taken from both the hourly dataset and the daily-averaged dataset; some parameters such as average temperature is selected from the daily averaged dataset whereas, other parameters such as maximum temperature (max temperature in a day at an hour) is taken from hourly dataset. Since here, the goal of using Random Forest is to extract the features (feature importance) from the dataset and not for prediction, 100% of the dataset is considered as training data. The model can be trained with different approaches which is explained in section 5.2.2.1. Once the model is trained, the importance of each weather parameter is extracted from the trained model using equations 5 to 8 in section 5.2.1. The extracted importance from the feature importance process are used as systematic weighting factors instead of the current Sandia weighting factors, which rely on experts' judgments. The new systematically assigned weighting factors, together with the daily-averaged data from the original Sandia method, is introduced to the FS-WS (equations 1 and 2) in order to find the representative months. The 12 representative months of different years are identified and the corresponding hourly weather data are extracted from the original dataset and joined to form a typical meteorological year.

Random Forest is used in this case solely to extract the feature importance, and therefore, the entirety of the dataset is used to train the model. The testing phase is not necessary in this application since Random Forest is not used for prediction, which is also applicable for cases where the goal is to evaluate the fit of the data. Once the feature importance is extracted, the weights are applied to the FS-WS equations, which are a form of the k-nearest neighbor algorithm. The k-nearest algorithm has no explicit training phase, incorporates the whole dataset, and simply generates values from the dataset. Once these values are generated, the CDF is taken and compared

to find the most similar candidate month. The similarity is cross-validated by minimizing the error by adopting different training approaches to be explained in the next section.

5.2.2.1 Training approaches

Thirty years of historical weather data of the nine weather parameters is used as training input, while the corresponding energy simulation results are used as training output. As mentioned in the previous section, the corresponding cooling, heating, and total energy demands of the buildings, which are converted to daily averaged data in order to apply the FS-WS procedure. Different approaches can be used to train the Random Forest model and extract the respective features importance. The proposed systematic approach in assigning weighting factors attends to the problem that, under certain applications and for certain climatic conditions, a universal set of weighting factors might cause discrepancy in long-term energy prediction. In most building projects, total energy demand is the main concern. There are also cases that cooling and heating energy demands would prefer to be identified separately, as cooling and heating might be provided by different energy sources (and thus at different prices). There are also cases that, for some climatic locations, a single set of weighting factors might not fully reflect the variations in influence of each weather parameters throughout the year. To properly train the random forest model, five different training approaches, investigating potentially different applications and climatic characteristics, are suggested. Details of each of these approaches are presented as follow:

- 1) Cooling Energy Based: the original dataset of 30 years energy demand is reduced to a subset of daily cooling having values larger than zero; the cooling energy demand serves as the output for training.
- 2) Heating Energy Based: the same approach as above where heating energy demand serves as the output for training.
- 3) Total Energy Based: the same approach as above where the total energy (summation of cooling and heating) is used as the output for training.
- 4) Month-by-Month (MBM) Cooling or Heating Energy Based: the original dataset are rearranged based on the months, which means 12 subsets of the original dataset are generated such that each subset represents a particular month of all 30 years of data. Afterwards, two subsets are generated, one with cooling energy demand greater than zero

and another with heating energy demand greater than zero in order to train the random forest model for cooling and heating respectively. The feature importance is then extracted from either the cooling or heating model with larger R^2 value. Consequently, 12 monthly feature importance are extracted for each weather parameter.

5) Month-by-Month (MBM) Total Energy Based: the original dataset is rearranged based on the months (overall 12 subsets) and each subset is reduced to the sets having a total energy demand larger than zero; for each subset a model is trained with the target of total energy demand which leads to 12 monthly feature importance for each weather parameter. This approach is considered as a strategy to use the model that is trained with greater numbers of data.

The accuracy of the models in different approaches are shown in section 5.3.1. Once the models are accurately fit to the output, the features importance are extracted. Once each model is trained, the feature importance is extracted from the model following the five approaches. For each approach with new feature importance, a typical meteorological file is created. The output of the building energy simulation for each of the five approaches are then compared to the results from actual years. The root mean square error (RMSE) between the simulations from actual years and the new weather years is calculated; this is shown in section 5.3.2.

5.3 Results and Discussion

5.3.1 Models performance

The comparison of the simulation results and the trained model results for the first three approaches for Montreal are presented in Figures 5.5 to 5.7. The accuracy of the models is confirmed by a very high R^2 value. The bottom of the figures zoom into the first year (1960) and shows that there is a good agreement between the simulation results and the model results. Vancouver data display similar R^2 value and are therefore omitted from presenting.

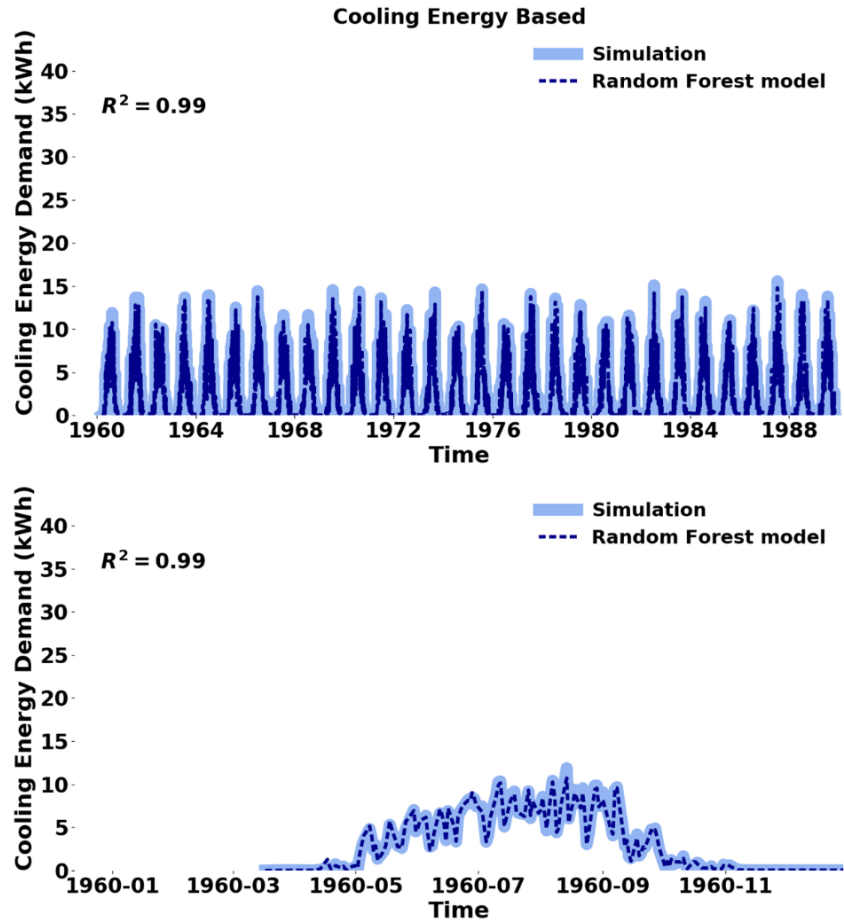


Figure 5.5: Results of trained models in comparison to the simulation results for cooling-based approach for Montreal; the bottom zooms into the first year (1960) of the top (approach 1).

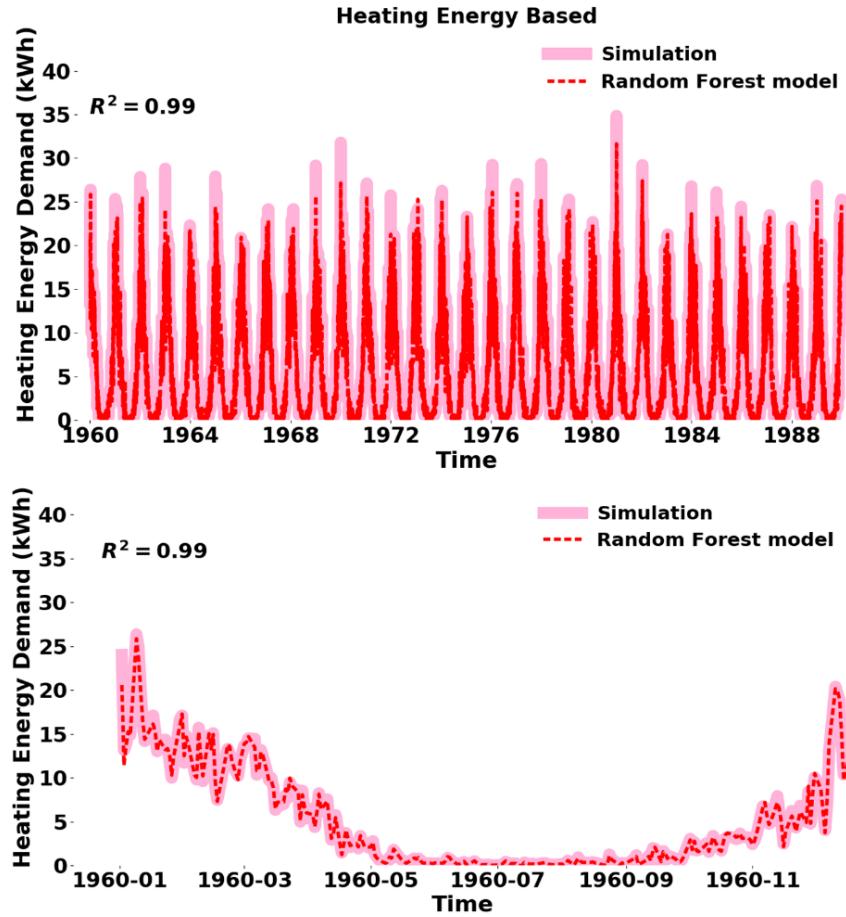


Figure 5.6: Results of trained models in comparison to the simulation results for heating-based approaches for Montreal; the bottom zooms into the first year (1960) of the top (approach 2).

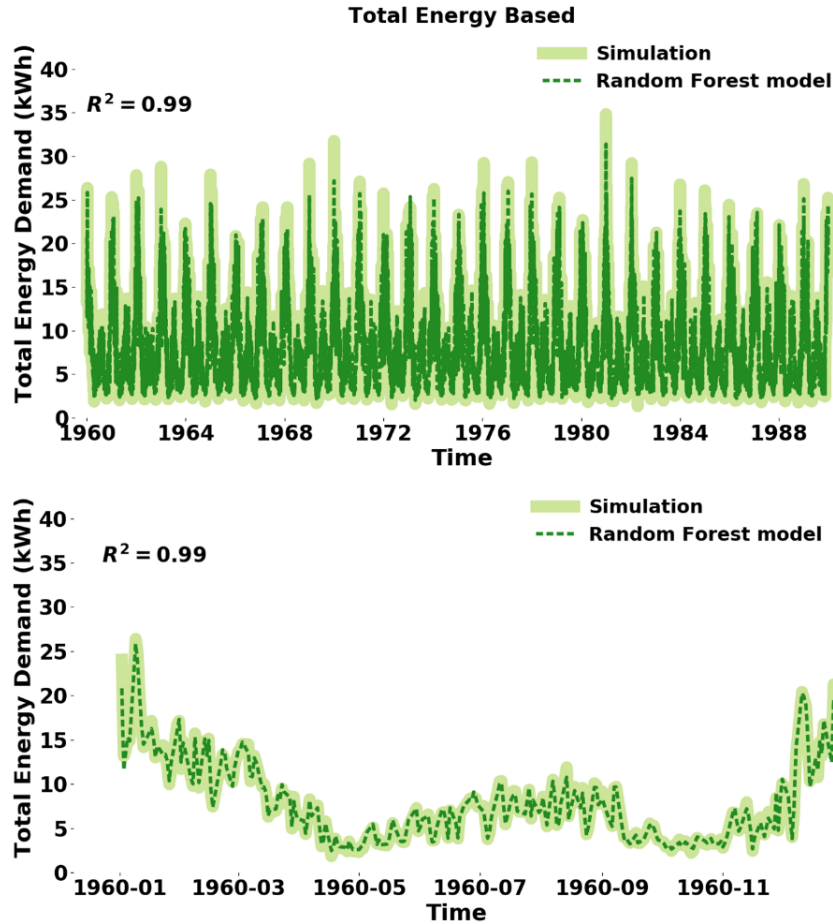


Figure 5.7: Results of trained models in comparison to the simulation results for total-based approaches for Montreal; the bottom zooms into the first year (1960) of the top (approach 3).

Table 5.3 shows the performance of the trained models based on the month-by-month (MBM) cooling or heating energy based approach (approach 4) and month-by-month total energy based in Montreal and Vancouver. For MBM cooling or heating, the dataset is rearranged to the 12 separate months and for each month a cooling (with cooling data) and a heating model (with heating data) are trained, then the model with a larger R^2 value is selected for further analysis. Due to the fact that during the cooling seasons, the number of data having positive cooling demand is larger than for the heating demand in the datasets, the models with cooling energy demand show larger R^2 values and are therefore selected for those particular months (May to August for Montreal and May to September for Vancouver); for the same reason, heating energy models are used for those heating predominant months (January to April and September to December for Montreal). Aside from the number of data, the variability (larger variance) can be a reason for the lower R^2 value

for some months. For example, the January model in both approach 4 and 5 has lower R^2 in comparison to other months for Montreal.

Table 5.3: Performance of the trained models for month-by-month cooling or heating energy based and month-by-month total energy based in Montreal and Vancouver (approaches 4 and 5)

	Montreal			Vancouver		
	MBM Cooling or Heating Based					
	Cooling Based	Heating Based	R2	Cooling Based	Heating Based	R2
Jan		✓	0.95		✓	0.96
Feb		✓	0.98		✓	0.97
Mar		✓	0.98		✓	0.95
Apr		✓	0.98		✓	0.94
May	✓		0.98	✓		0.96
Jun	✓		0.98	✓		0.96
Jul	✓		0.97	✓		0.95
Aug	✓		0.98	✓		0.97
Sep		✓	0.98	✓		0.95
Oct		✓	0.99		✓	0.97
Nov		✓	0.98		✓	0.97
Dec		✓	0.98		✓	0.98
	MBM Total Based					
Jan			0.95			0.97
Feb			0.98			0.97
Mar			0.98			0.95
Apr			0.97			0.92
May			0.98			0.94
Jun			0.98			0.96
Jul			0.97			0.95
Aug			0.98			0.97
Sep			0.97			0.95
Oct			0.96			0.96
Nov			0.98			0.97
Dec			0.98			0.98

5.3.2 Features importance: systematic approach to defining weighting factors

As defined in the methodology section, the feature importance for each parameter was generated using Random Forest for all five approaches. Once the models from different approaches are trained, the feature importance of each model is extracted before they are used in FS statistic. Figure 5.8 shows the feature importance calculated for the weather parameters using cooling, heating, and total energy (approaches 1 to 3). The bars show the importance of each weather parameter on cooling, heating, and total energy demand for the two locations. In each of the diagrams, the green line represents the cumulative importance while the grey line shows the 95% threshold. The cumulative threshold can be inferred as the level where 95% of the information can be achieved. For example, for the cooling energy based approach, temperature, solar irradiance, and wind speed contribute 95% of the energy demand uncertainty in Montreal whereas, in Vancouver, temperature, solar radiation, maximum temperature, and wind speed contribute 95% of the total uncertainty. For the heating energy based approach in Montreal, temperature alone can almost account for 95% of total uncertainty. In the case of the total energy based in Vancouver, the five weather parameters of average temperature, maximum temperature, solar radiation, wind, and dew temperature lead to the threshold.

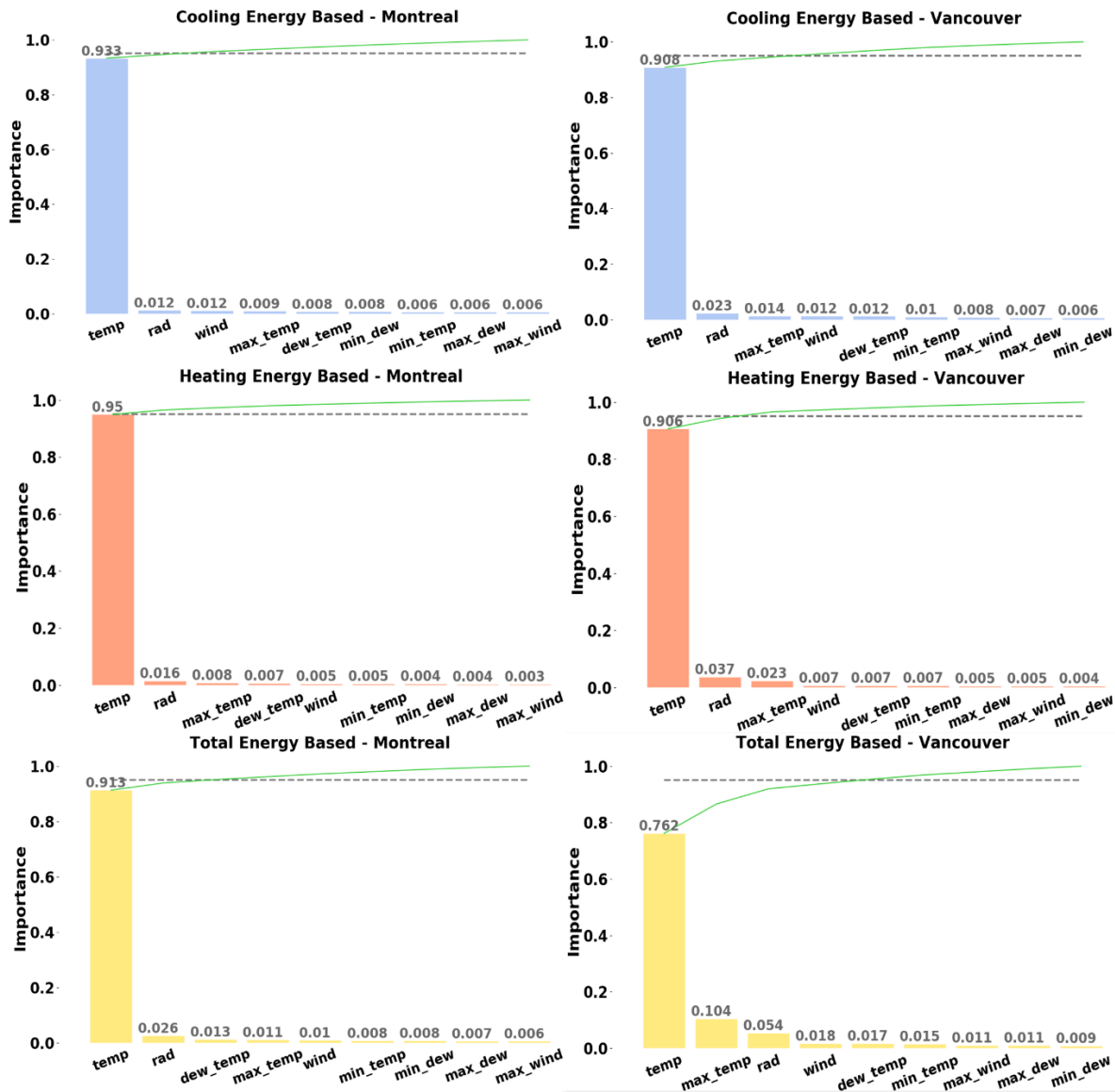


Figure 5.8: Feature importance extraction from Random Forest Regression (approaches 1 to 3)

Table 5.4 summarizes the results from MBM cooling or heating energy based model and MBM total energy based model for Montreal and Vancouver (approach 4 and 5). In both cases, the weather parameters have various values in different months. Depending on the climate, month-by-month variation of solar angle, temperature, humidity, and wind can lead to various effect on cooling or heating energy demand which is noted in Table 5.4. Although these values have a relative importance on energy demand, some interesting points are noted. For example, in Montreal, in the month of July, using the MBM cooling or heating based approach, the average

temperature and the maximum temperature show an importance of 0.544 and 0.307 respectively, whereas these values for the month of February are about 0.875 and 0.027. The summer season in Montreal has warm days and cool nights which can lead to a more stable average temperature, while the maximum temperature experiences more variation which is most likely the reason for the higher maximum temperature importance in July. Solar irradiance shows rather similar importance in July which can be explained by the fact that the case study building has a sloped roof that makes the building absorb less solar radiation compared to a flat roof. For the same reason, compared to a flat roof, the sloped roof can absorb higher solar radiation when the sun angle is lower in winter. Moreover, it should be noted that the overall lower weighting factor applied to solar irradiance does not necessarily indicate that the parameter has minimal effect on total energy demand, but rather that regardless of the variation in solar radiation, the imposed changes on total energy demand are small.

Table 5.4: Weighting factors used resulted from MBM cooling or heating energy based, and MBM total energy based model for Montreal and Vancouver.

	Montreal									Vancouver								
	MBM Cooling or Heating Based																	
	rad	temp	max temp	min temp	dew temp	max dew	min dew	wind	max wind	rad	temp	max temp	min temp	dew temp	max dew	min dew	wind	max wind
Jan	0.046	0.628	0.070	0.042	0.090	0.026	0.050	0.032	0.018	0.046	0.784	0.038	0.024	0.020	0.021	0.017	0.029	0.019
Feb	0.024	0.875	0.027	0.015	0.014	0.009	0.013	0.013	0.009	0.029	0.819	0.028	0.022	0.023	0.017	0.012	0.032	0.017
Mar	0.021	0.864	0.016	0.020	0.022	0.012	0.016	0.020	0.008	0.072	0.630	0.087	0.031	0.043	0.031	0.028	0.050	0.028
Apr	0.016	0.849	0.046	0.013	0.018	0.015	0.016	0.017	0.011	0.056	0.676	0.042	0.030	0.042	0.028	0.039	0.055	0.032
May	0.016	0.883	0.030	0.012	0.010	0.011	0.009	0.018	0.012	0.039	0.768	0.035	0.027	0.034	0.020	0.017	0.036	0.023
Jun	0.016	0.801	0.104	0.011	0.010	0.015	0.013	0.020	0.009	0.037	0.596	0.221	0.026	0.031	0.019	0.015	0.035	0.020
Jul	0.026	0.544	0.307	0.019	0.022	0.017	0.019	0.029	0.018	0.037	0.524	0.269	0.027	0.033	0.023	0.021	0.041	0.026
Aug	0.027	0.805	0.083	0.012	0.010	0.013	0.011	0.026	0.013	0.037	0.668	0.183	0.020	0.020	0.017	0.013	0.023	0.018
Sep	0.018	0.871	0.019	0.023	0.016	0.014	0.014	0.016	0.009	0.064	0.638	0.109	0.034	0.042	0.019	0.022	0.047	0.025
Oct	0.011	0.920	0.013	0.013	0.012	0.008	0.008	0.008	0.007	0.039	0.781	0.074	0.024	0.017	0.019	0.010	0.022	0.013
Nov	0.022	0.834	0.057	0.022	0.017	0.011	0.015	0.012	0.009	0.050	0.804	0.043	0.029	0.019	0.013	0.012	0.020	0.010
Dec	0.016	0.867	0.016	0.019	0.027	0.011	0.021	0.014	0.009	0.022	0.862	0.033	0.015	0.018	0.011	0.011	0.018	0.011
	MBM Total Based																	
Jan	0.046	0.632	0.069	0.042	0.087	0.026	0.050	0.031	0.018	0.046	0.786	0.036	0.026	0.020	0.021	0.017	0.030	0.019
Feb	0.025	0.877	0.026	0.015	0.014	0.009	0.013	0.014	0.009	0.028	0.825	0.026	0.023	0.023	0.017	0.012	0.030	0.016
Mar	0.021	0.862	0.016	0.021	0.022	0.012	0.017	0.020	0.008	0.069	0.641	0.088	0.029	0.042	0.030	0.025	0.047	0.027
Apr	0.018	0.822	0.052	0.017	0.020	0.019	0.018	0.020	0.012	0.073	0.595	0.047	0.043	0.052	0.037	0.040	0.078	0.036
May	0.026	0.793	0.075	0.017	0.015	0.016	0.013	0.029	0.018	0.056	0.694	0.050	0.033	0.050	0.024	0.020	0.044	0.029
Jun	0.019	0.501	0.395	0.012	0.012	0.012	0.015	0.024	0.010	0.035	0.535	0.288	0.025	0.029	0.019	0.014	0.036	0.020
Jul	0.028	0.400	0.442	0.019	0.023	0.017	0.021	0.029	0.019	0.039	0.429	0.357	0.027	0.034	0.023	0.022	0.042	0.027
Aug	0.031	0.710	0.170	0.013	0.012	0.013	0.011	0.027	0.013	0.040	0.530	0.312	0.021	0.021	0.017	0.014	0.025	0.019
Sep	0.039	0.570	0.269	0.020	0.020	0.021	0.018	0.028	0.015	0.073	0.443	0.277	0.038	0.045	0.020	0.025	0.050	0.029
Oct	0.024	0.777	0.036	0.038	0.035	0.022	0.022	0.032	0.015	0.058	0.718	0.072	0.036	0.025	0.031	0.013	0.030	0.018
Nov	0.022	0.829	0.060	0.022	0.018	0.011	0.016	0.013	0.009	0.047	0.806	0.041	0.032	0.019	0.014	0.012	0.019	0.010
Dec	0.016	0.865	0.016	0.020	0.027	0.011	0.021	0.014	0.009	0.022	0.862	0.033	0.015	0.017	0.011	0.011	0.018	0.011

5.3.3 Approaches performance

The feature importance resulting from approaches 1 to 5 were introduced to FS procedure (w_i in Equation 2). For each approach, a new typical meteorological year weather file was generated with the new weighting factors for weather parameters. Energy simulations are conducted using new weather files to evaluate cooling and heating energy demand of the buildings. Figure 5.9 shows the simulation results with actual meteorological year weather files (fluctuating plot), the original CWEC file and the newly generated typical meteorological year weather files (horizontal lines) for Montreal and Vancouver.

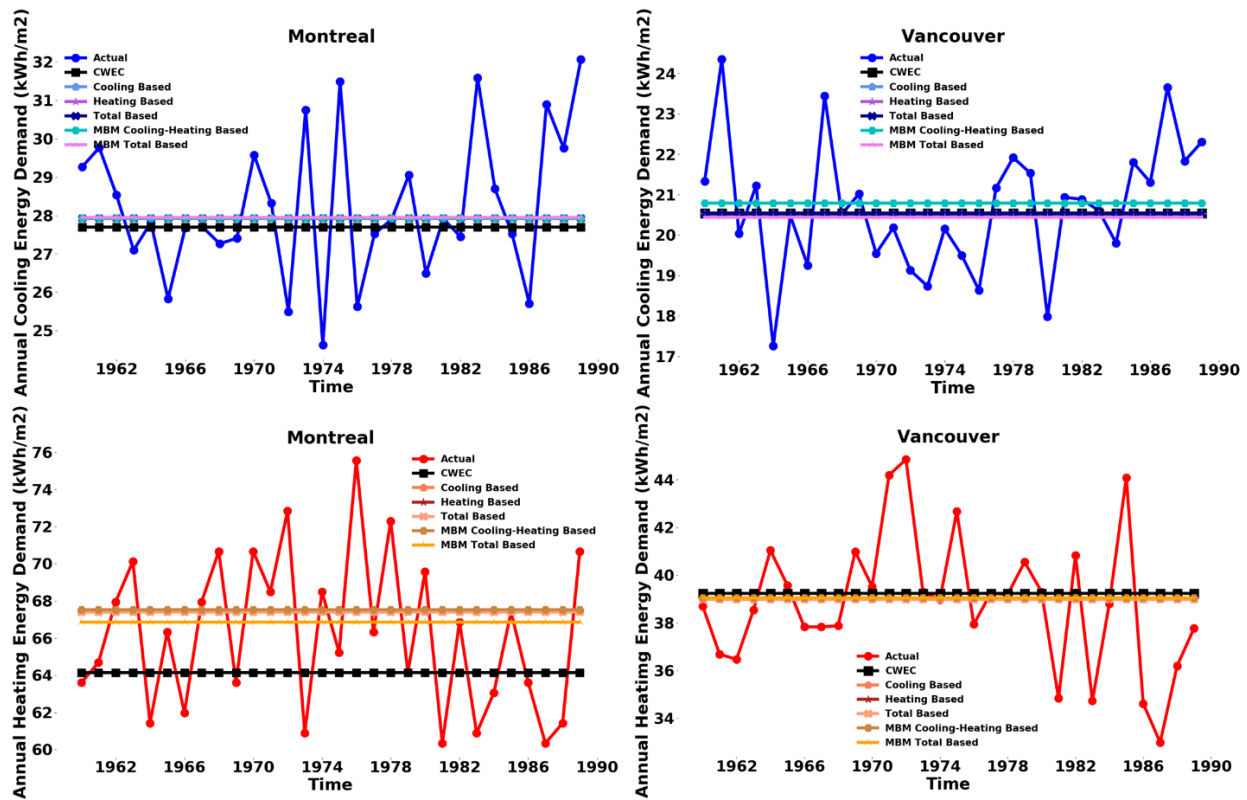


Figure 5.9: Simulation results based on actual meteorological years and the typical meteorological year weather files made up with weighting factors based on different approaches for Montreal and Vancouver.

As observed, different approaches lead to different simulation results although the deviation is minor in most of the cases. However, in the case for Montreal heating energy demand, the new approaches showed quite a significant difference. The heating energy demand resulted from CWEC showed about 64 kWh/m^2 , whereas the new approaches showed a demand in the range of 67 to 68 kWh/m^2 . The considerable reduction in deviation noted for Montreal can be explained by

the fact that, in this case, the dew point is highly correlated with dry-bulb temperature, as seen in Figure 5.1. Therefore, in order to reduce the variance of the model, the dew point parameter may be ignored or assigned a minor weighting factor. The latter is done by applying the new weighting factors generated from the Random Forest feature importance algorithm. However, in Vancouver, the dew point does not exhibit as high of a correlation with the dry-bulb temperature as Montreal. Therefore, removing the parameter or allocating a minor weighting factor to the parameter may not reduce the uncertainty of the model significantly.

When the year-to-year fluctuations in energy demand based on actual meteorological year weather are considered, the new approaches show more promising results. In order to estimate the deviation of each method from the simulation with actual years, the root mean square error for each approach is calculated using the difference between the actual years' results and the new approaches results. Table 5.5 shows the root mean square error calculated for the original CWEC file and the new approaches, where the root mean square error reduction of new approaches are calculated in comparison to the CWEC RMSE to show the improvement by each new approach.

Table 5.5: Root mean square error calculated for CWEC and the suggested approaches. RMSE reduction shows the reduction of root mean square error in comparison to CWEC method. Minus values show penalty; the numbers of percentages are rounded to zero digit. Blue and pink sections represent RMSE for cooling and heating respectively.

	CWEC	Annual-Cooling Based	Annual-Heating Based	Annual-Total Based	MBM-Cooling or Heating Energy Based	MBM-Total Energy Based	
Montreal	RMSE	2.54	2.47	2.47	2.47	2.47	
	RMSE reduction (%)		3	3	3	3	
	RMSE	6.00	5.50	5.50	5.50	5.40	
	RMSE reduction (%)		8	8	8	10	
	RMSE	4.52	3.76	3.76	3.76	3.89	3.70
	RMSE reduction (%)		16	16	16	14	18
Vancouver	RMSE	2.05	2.06	2.06	2.06	2.05	2.07
	RMSE reduction (%)		0	0	0	0	-1
	RMSE	3.63	3.60	3.60	3.60	3.61	3.61
	RMSE reduction (%)		1.00	1.00	1.00	1.00	1.00
	RMSE	2.64	2.63	2.63	2.63	2.65	2.63
	RMSE reduction (%)		0	0	0	0	0

As presented in Table 5.5, the new approaches of generating typical meteorological year weather file reduced the root mean square error as compared to conventional means such as that of CWEC. The RMSE of CWEC for cooling energy demand in Montreal is reduced from 2.54 to 2.47 showing

a 3% reduction with the new approaches. For the heating energy demand, CWEC RMSE of 6 is reduced to 5.5 for new approaches, a reduction of more than 8%. The largest reduction in RMSE for Montreal was noted for the total energy demand. The RMSE of CWEC for the total energy demand was 4.52, which was reduced to 3.70 using the MBM total energy based approach (approach 5), which represents an overall reduction of 18%.

For Vancouver, the new approach could achieve only a minor improvement of about 1% in heating and a negligible penalty of less than 1% for cooling.

The implication of applying the new approaches in generating typical meteorological year weather files can be summarized below:

- The proposed method has reduced the uncertainty associated with typical meteorological year and has improved the representativeness of the weather year generated from the Sandia method based on experts' judgments.
- The method showed better improvement when applied to the colder climate of Montreal as the RMSE observed both in cooling and heating were reduced considerably. This is partially due to the high correlation between the dew point and dry-bulb temperature, which implies one of the parameters can be removed or allocated with a minor weighting factor to reduce the uncertainty of the model. For the milder cold climate of Vancouver, the effect of applying the proposed method is negligible.
- All five approaches yielded similar results, in which the MBM total energy-based model for Montreal offered slightly better result. Approach 5 (MBM total energy based) centers on increased information being relayed from the dataset where more data having positive output as a result of summing up cooling and heating energy demands. The increase in information helps to improve the Random Forest model during the training.

5.4 Conclusion

This study introduced a novel method to systematically weigh the weather parameters for generating typical meteorological year. The “feature extraction” or “feature importance” method used in machine learning algorithms to reduce the uncertainty of modeling is applied in this study.

For the case of cooling and heating energy demand of a small office building, the weighting factors resulted from the novel method were different from those originally used in CWEC. For example, the weighting factor for solar radiation was overestimated whereas the average temperature was underestimated in CWEC. The new weighting factors resulting from the different approaches reduced the uncertainty and increased the typicality of the weather year which was reflected by RMSE reduction of almost 10% in one case. It is noted that a dynamic monthly set of weighting factors (12 sets or approaches 4 and 5) instead of a single annual set of weighting factors will most likely show a better result. This is due to the fact that multiple sets of weighting factors can make the Sandia method (which can be considered as a K nearest neighbor algorithm) more adaptive to the problem.

The proposed method is applicable to different climates and potentially offers better energy performance prediction and introduces significant energy saving for climatic locations with a dynamic temperature range. However, more studies are required to explore the effect of the proposed method for warm climates as well as for different building types and sizes.

5.5 References

- [1] M. Hosseini, B. Lee, and S. Vakilinia, “Energy performance of cool roofs under the impact of actual weather data,” *Energy Build.*, vol. 145, 284–292, 2017
- [2] S. Janjai and P. Deeyai, “Comparison of methods for generating typical meteorological year using meteorological data from a tropical environment,” *Appl. Energy*, vol. 86, no. 4, 528–537, 2009
- [3] National Climatic Center, “Typical Meteorological Year User’s Manual,” 1981
- [4] H. Lund, “The Design Reference Year Users Manual”. Thermal Insulation Laboratory, Technical University of Denmark, 1995
- [5] R. Festa and C. F. Ratto, “Proposal of a numerical procedure to select Reference Years,” *Sol. Energy*, vol. 50, no. 1, 9–17, 1993
- [6] International Organization for Standardization. “ISO Standard 15927-4, Hygrothermal performance of buildings — Calculation and presentation of climatic data — Part 4: Hourly data for assessing the annual energy use for heating and cooling”. 2005
- [7] W. Marion and K. Urban, “Users manual for TMY2s: Derived from the 1961--1990 National Solar

- Radiation Data Base,” 1995
- [8] S. Wilcox and W. Marion, “User’s manual for TMY3 data sets,” National Renewable Energy Laboratory Golden, 2008
 - [9] Environment and Climate Change Canada, “Canadian Weather Energy and Engineering Data Sets (CWEEDS files) and Canadian Weather Year for Energy Calculations (CWEC Files) User’s Manual,” 2008
 - [10] S. A. Kalogirou, “Generation of typical meteorological year (TMY-2) for Nicosia, Cyprus,” *Renew. Energy*, vol. 28, no. 15, 2317–2334, 2003
 - [11] A. L. S. Chan, “Generation of typical meteorological years using genetic algorithm for different energy systems,” *Renew. Energy*, vol. 90, 1–13, 2016
 - [12] G. Georgiou, M. Eftekhari, P. Eames, and M. Mourshed, “A Study of the effect of weighting indices for the development of TMY used for building simulation,” in *Building Simulation*, 922–929, 2013
 - [13] T. Kalamees *et al.*, “Development of weighting factors for climate variables for selecting the energy reference year according to the en ISO 15927-4 standard,” *Energy Build.*, vol. 47, 53–60, 2012
 - [14] L. Braiman, “Random Forests,” *Mach. Learn.*, vol. 45, no. 1, 5–32, 2001
 - [15] T. Hastie, R. Tibshirani, J. Friedman, “The Elements of Statistical Learning Data Mining, Inference, and Prediction”, Second Edition, New York, Springer, 2009, ISBN: 978-0-387-848570
 - [16] Saeys, Yvan, Thomas Abeel, and Yves Van de Peer. "Robust feature selection using ensemble feature selection techniques." *Joint European Conference on Machine Learning and Knowledge Discovery in Databases. Springer, Berlin, Heidelberg*, 2008
 - [17] Y. Lin and Y. Jeon, “Random Forests and Adaptive Nearest Neighbors,” *Am. Stat. Assoc.*, vol. 101, no. No. 474, 578–590, 2006
 - [18] United States Department of Energy, “Commercial Prototype Building Models,” 2019. [Online]. Available: https://www.energycodes.gov/development/commercial/prototype_models.
 - [19] NECB, “National Energy Code of Canada for Buildings 2015,” Natl. Res. Counc. Canada, 2015
 - [20] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification And Regression Trees*, 2nd ed. Boca Raton London New York Washington, D.C., 1984

Chapter 6. A PRILIMINARY STUDY OF CLIMATE CHANGE EFFECT ON BUILDING ENERGY PERFORMANCE

The nature of TMY in excluding weather extremes makes them less suitable to investigate the effect of potential climate change on building design as climate change likely increases the frequency and magnitude of those extreme conditions. The current practice of designing buildings has lacked a clear method to incorporate future climate change trends. An approach is used to compare present weather simulation results of a commercial building with varying roof reflectance and insulation thermal resistance parameters with future year-by-year results which are affected by potential climate change. Future weather data for year-by-year simulations is obtained by “morphing” historical weather data with a General Circulation Model (HadCM3). Mean energy consumption and optimal roof configurations are discussed with regards to climate change over the study period, and are compared to results obtained with TMY data.

This chapter was published in the journal of Building Engineering, Volume 17, Pages 107-114, M. Hosseini, F. Tardy, B. Lee, “Cooling and heating energy performance of a building with a variety of roof designs; the effects of future weather data in a cold climate“, © Elsevier Ltd. 2018

6.1 Introduction

Over the past decades, literature has indicated that a warming global climate is affecting various human activities ranging from crop production [1] to power plant output [2]. The practice of designing buildings to cope with potential climate change has lacked a clear method to incorporate this trend. Today’s buildings are designed to last several decades, and as weather patterns change over time, buildings designed for today’s climate may not withstand the potential changes during their useful lives.

Building designers should therefore take future climate predictions into account when assessing building energy performance in the subsequent building design process. Most building energy simulation programs use weather data which represents a single, typical meteorological year

(TMY). The implications of this practice are twofold. a) extreme weather conditions are excluded from the TMY weather data, and the use of TMY data might not be able to reflect future realities since the weather tends to become more extreme under the premise of climate change [3] and b) regardless of the different climate change scenarios, changes in the weather might not be adequately captured by a single TMY.

Therefore, even if building engineers today commonly use TMY weather data for design and analysis purposes, such data can not only lead to an under or overestimation of energy savings, but also does not support future weather modeling.

In this optic, the objective of this research is to:

1. Quantify and systematically demonstrate the effects of future climate changes on energy consumption.
2. Offer a path to building design which considers the effects of climate change.

This research described in this paper seeks to accomplish this objective by improving the thermal design of roofs in cold climates to reduce overall yearly energy consumption by anticipating the predicted effects of future climate change. Two factors inflecting roof design are studied: thermal insulation and solar reflectance. Simulations are conducted for several combinations of the two factors in order to comprehend underlying synergies and trends.

6.1.1 Climate change and its impact on building energy performance

Jentsch et al. [4] discussed the fact that many currently used TMY weather files for building energy performance are typically derived from historical weather data from the latter 20th century, and research by [5,6] has demonstrated that there exists discrepancies between this data and current weather trends. In Canada, the same issues arise from the use of Canadian Weather Year for Energy Calculation (CWEC) data [7], which is derived from historical weather data from 1961-1990. With the now widely accepted effects of climate change, the amount of energy used for building cooling should increase in the future. This is supported by the U.S. Global Change Research Program, which states that “warming will be accompanied by decreases in demand for heating energy and increases in demand for cooling energy. The latter will result in significant increases in electricity use and peak demand in most regions” [8].

While heating buildings can be achieved by using various energy sources at high efficiency, most commercial cooling devices operate solely on mechanical power, but several advances in passive cooling may be used to shift the cooling load to other measures [9] and the combination of various materials and colors and structure weights on roofs [10].

6.1.2 Climate change prediction models and their implication on building energy simulation

Building energy simulations involve hourly step calculations which reflect the complex interactions between HVAC systems, control systems, internal loads and external factors. Building energy simulation programs are commonly used to quantify the savings and/or penalties for a variety of techniques used to improve energy efficiency and to estimate the monthly and annual energy consumption of buildings. However, this quantification, might not be reliable due to a deterministic approach in simulation. Estimated energy performance based on typical meteorological years (such as TMY) and Canadian weather years for energy calculation (CWEC) may not reflect actual energy performance and their variations. In addition, TMY does not capture extreme weather conditions. Given the stochastic nature of building operation and weather patterns, exact predictions are difficult, if not impossible, to obtain. Furthermore, research by [11] documented that the city built environment and heat islands pose effects on the temperature and humidity of the surrounding environment, and suggests that these factors must be included in weather data for building simulations by using an urban downscaling methodology.

Researchers have increasingly been using General Circulation Models (GCMs) to predict future weather patterns affected by climate change. So far, several methodologies have been developed to integrate these predictions into weather files which is used to reliably prepare for the eventuality of climate change [12-14]. Jentsch et al. [15] discussed the importance of climate change adaptability in planning for future climate scenarios into the widely used TMY2 weather file formats. They chose to use the Hadley Center Coupled Model, version 3 (HadCM3) to predict future weather conditions, accounting for the effects of climate change. Instead of representing data predictions for a single weather station at a specific geographical location, the HadCM3 model covers a finite grid point model covering an area of 2.5° latitude by 3.75° longitude, with a resolution nearing $300 \times 300 \text{ km}^2$ worldwide. However, general circulation models such as HadCM3 provide monthly data whereas hourly data is required in building energy simulations. A

downscaling method should be applied to fine-tune to the required resolution. Jentsch et al. [15] [9] proposed using ‘morphing’ techniques to generate new TMY2 files from current Test Reference Years (TRY) and Summer Design Years (DSY).

With morphing methodology, originally developed by Belcher et al. [16], corrections for future hourly temperature, humidity and solar radiation GCM data are simply superimposed on existing EPW (EnergyPlus Weather) data. A modified weather file is generated which can be used by building designers with tools that are available in the public domain.

Zhu, et al. [17] applied the TMY morphing method to three cities in China over five future periods ending in 2089, and showed that average increases in temperature ranged from 3.0°C to 5.4°C in all three cities. In their analysis of a “Passivhaus”, or low energy design, McLeod et al. [18], determined peak load data from the worse of two distinct winter weather situations from a “morphing” method used with the Hadley Centre Regional Model 3 (HadRM3), which they have encapsulated in a software conveniently named the Passivhaus Planning Package (PHPP12).

Kikumoto et al. [14] proposed an improved method for obtaining future weather data by using a method referred to as “dynamical downscaling”. The often used “morphing” method, the authors argued, causes much data to be lost due to the GCM’s coarse resolution. It was also stressed that statistical manipulations led to the loss of information concerning interactions between various weather components, which is particularly important in determining extreme weather conditions. Boundary conditions originating from GCM data for prediction year climates are used to define a Regional Climate Model and the data is dynamically downscaled to produce weather predictions and standard data which can be used to simulate a building in the predicted climate for the year in question. The authors created a Weather Research and Forecasting (WRF) model for the 2030’s, which they used to conduct accurate predicted future building energy calculations in a simulated detached house, with room for improvements in accuracy and bias mitigation.

6.1.3 Adapting building design processes to adapt to climate change

With regards to the environmental or economic strategies surrounding the design and usage of buildings over their useful lives, climate change presents building designers with added constraints. Designers who do not take future conditions into account risk presenting future owners or occupants with buildings which might not effectively respond to local environmental conditions

at some point in time. Therefore, the focus of any climate adaptive design will be to analyze the energy patterns of a building with different simulations using weather predictions which span over this period [19].

The simplest energetic strategy to implement could be the one which produces the least deviations from normal conditions, but economic factors or business cases could make other strategies, such as retrofitting, more appealing. Loonen et al. [20] discussed the idea of a Climate Adaptive Building Shell (CABS), which was defined as a building which could “repeatedly and reversibly change some of its functions, features or behaviour over time in response to changing performance requirements and variable boundary conditions with the aim of improving overall building performance”. Robustness, adaptability, and multi-ability were defined and placed into context for energy efficiency. While their definition is not specifically geared towards global climate change, it does provide a blueprint towards defining building adaptability to varying indoor and outdoor conditions.

Robert and Kummert [21] generated future weather files to investigate if they affected the energy performance of an existing Net Zero Energy Building (NZEB) home in the northern climates of Montréal, QC and Massena, NY, and they found that the building does not attain net-zero energy status in future years. This has led to the argument that NZEB buildings should always be designed with weather data spanning over their entire useful life instead of with TMY data which might not even adequately reflect the first year of operation. In the same vein, McLeod et al. [18] argued that special attention needed to be placed on accurate local climate data to make Passivhaus designs relevant, data which is further complicated with climate change. For these specific cases, they considered that overheating and undercooling posed a significant risk to the mission and certification of the building.

6.1.4 Effects of highly reflective roofs (cool roofs) on buildings

The practice of installing cool roofs (surfaces which highly reflect solar radiation back to sky) in various climates have existed for millennia, as white roofs are very prevalent along the Mediterranean and in the Middle East. As effective as they are, they have only been sparsely used in western architecture. However, as techniques used to improve comfort have progressed, cooling has become as equally important to building design as heating is. Since buildings are static structures that cannot easily be adapted to changing seasonal weather patterns, a balance must be

found to smooth power demand and limit energy consumption throughout the year while taking advantage of the conditions present in the natural surroundings. Such a balance will often be unique to every type of building in every climate.

In hot environments, while limiting heat gain and increasing reflectivity on the roof can passively serve to increase comfort when outdoor temperatures are high, such measures can also serve to reduce cooling loads from air-conditioning devices. In the United States, the Department of Energy (DOE) began investigating the benefits of cool roofs in the 1980's, which were favorably adopted in California at the beginning of the 21st century as a method to reduce peak demand from air conditioning in the summer following an energy crisis. Several years later, such practices have become more common, with dedicated organizations such as the Cool Roof Rating Council making their appearance as a measure in green building certifications, such as LEED accreditations.

Currently, for a roof to be defined as a “cool roof”, it must possess a high solar reflectance factor of 0.55 after 3 years of use, effectively returning a majority of the solar radiation hitting the roof back to the sky in the form of infrared radiation [22, 23] although technology helped to produce dark-colour less reflective cool roofs [24]. Several studies have been made to determine deciding factors on their usability for various buildings particular climates. Piselli et al. [25] studied the use of cool roofs on buildings in five Italian climate zones with varying occupancies, building characteristics and HVAC systems, using optimized solar reflectance factors to minimize the energy consumption of the buildings. Their results showed that for warmer climates, the maximally considered solar reflectance of 0.8 was the optimal value. However, for heating dominant regions, the optimal solar reflectance depended on other parameters, such as the characteristics of the HVAC system. Furthermore, [26] devised a simple calculator to assist designers in determining the benefits of highly reflective roofs in varying conditions, while [27] proposed an advanced model which correlated daily accumulative inward heat in buildings with rooftop albedo, mentioning that increasing roof insulation can curtail air-conditioning requirements in the summer.

Several studies were conducted to determine the effects of varying albedo on roofs and their surroundings. A series of experiments were conducted by [28] to measure heating and cooling energy demand changes with varying roof albedo in various conditions and were validated with TRNSYS simulations. The authors found variations in reductions in air temperatures and

overheating hours, increased heating loads, and decreased cooling loads. Touchaei et al. [29] found that increasing the solar reflectance of surfaces such as roofs, walls and pavement in urban settings located in cold climates reduced solar heat gains in buildings and modified surrounding meteorological conditions. A DOE-2 building simulation was used to simulate four prototype commercial buildings with varying roof definitions: a dark roof control design, a white roof control design, and an albedo-enhanced roof. While calculating the difference in yearly heating and cooling energy consumption over white and albedo roof scenarios, it was found that the cooling energy savings from the white control roof cancelled the heating energy penalties for small offices and that heating energy losses from albedo-enhanced roofs outweighed cooling savings. However, as the size of the office building increased, cooling energy savings from white or albedo-enhanced roofs surpassed heating losses, and thus their presence became justified.

Beyond reflectance, increasing the thermal resistance of insulation on roofs to reduce energy consumption in the heating season, which is not specifically defined but which generally runs from October to May, can lead to contradicting effects during the cooling season, which generally runs from July to September. The minimal required total Thermal Resistance (RSI value) of insulation in buildings suggested by the national building codes for Montreal is 5.4 m²K/W. An experimental study conducted by Ramamurthy et al. [30] concluded that extra insulation on the roof might not always be beneficial during the cooling season, as heat accumulating inside a building can be prevented from exiting a highly insulated roof.

On its face value, the decision to use cool roofs in cold and northern climates may seem counterproductive due to the advantages obtained from solar heat gains in the winter. However, this is misleading. The presence of snow in most of the heating season, narrowed sun ray angles radiating during shorter days, increased cloud cover and nighttime heating schedules all contribute towards minimizing the impact a cool roof would have in preventing winter heat gains in the building. The effects of snow accumulation on cool roofs were studied by [31-33], who concluded that its presence significantly mitigated their heating penalty in the winter in both Montreal, QC and Anchorage, AK. Considering that weather patterns are dynamic and often deviate from standard definitions, roof design practices for winter remains largely unchanged due to the permanent factors discussed above. While predictions show that the length of the annual snowfall period will decrease with a warming climate, the frequency of heavy snowfall events have been

increasing of the past 30 years and it is estimated that it the trend should continue in the foreseeable future [34].

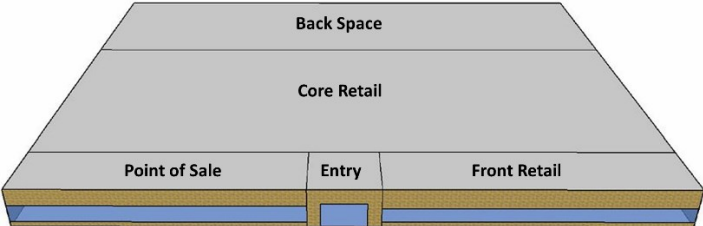
Considering the effects of climate change, the literature shows that, compared to previous years, warmer summers and winters with reductions in solar irradiation in the winter and increases in the summer, are expected in future years [21]. These variations can affect the energy performance of reflective roofs. Therefore, future weather data that accounts for the effects of climate change is used to demonstrate the energy performance of a variety of roof designs for the future.

6.2 Methodology

6.2.1 Adaptation of a climate change model to a base case building

To evaluate the effects of present and future weather conditions on a particular design, one of the building scenarios that was previously used in [7, 32], a one-storey commercial building with a 2299 m² floor area and a flat roof, was modeled with varying thermal resistance insulation and solar reflectance values. The DOE prototypical retail store reference building geometry and its characteristics are also available on the U.S Department of Energy website [35]. The original building enclosure characteristics are adapted to ASHRAE standard 90.1, upgraded to the Canadian National Energy Code for Buildings (NECB, 2011) since the case study is located in Montreal, Canada. The building is defined as having five zones (core, front, back space, point of sale, entry) and no plenum. This building model has been defined with 126 different roof configurations. The layout and overall characteristics of the building are summarized in Table 6.1.

Table 6.1: Building characteristics used for simulation

Item	Descriptions
Form	
Total Floor Area	2299 m ² (54.2 m x 42.3 m)
Building shape	
Window-to-Wall Ratio (WWR)	25.4% on the south facing facade
Envelope	
Exterior walls	
RSI-value (m ² K /W)	4.1
Solar reflectance	0.7
Roof	
RSI-value (m ² K /W)	2.4 to 15.4
Solar reflectance	0.1 to 0.9
Window	
RSI-value (m ² K /W)	0.5
SHGC	0.3
Foundation	
RSI-value(m ² K/W)	5.9
Air Barrier System	
Infiltration	0.001024 m ³ /s·m ² of above ground envelope surface area

EnergyPlus [36] is used in this paper to simulate building energy consumption. Total heating and cooling loads per unit area have been simulated for every simulation year over a lifespan of 20 years which begins in the year 2018. To reflect actual building energy consumption, a previously introduced COP of 2.93 has been applied to the cooling load to calculate cooling energy consumption, and the heating source for the building is assumed to be entirely electric, with an efficiency (η_{heating}) of 100%. Therefore, final results do not represent physical quantities of thermal energy, which in theory would negate each other, but instead represent the total yearly mechanical-electrical energy consumption required to heat and cool the building. While this assumption may not accurately predict the end-use electricity consumption, it is a common and acceptable method for estimation purposes. The reason behind this method is twofold. Firstly, determining HVAC system consumption in simulations requires much greater simulation resources. Secondly, depending on the building enclosure configuration (roof design), the size of the required HVAC

system could change, leading to variations in HVAC efficiency, affecting HVAC efficiency and therefore results, which is outside the scope of this study.

6.2.2 Future hourly weather generation

Six future possible scenarios relating to population of the world, economic condition, and the technology are introduced in the third and fourth reports from the IPCC (Intergovernmental Panel on Climate Change) [37]. Each of these scenarios predict that a specific quantity of greenhouse gas emissions are released up to the year 2100, based rely on data collected in the 80s and 90s. In the A2 scenario, the global population continuously increases, economic development is primarily regionally oriented, and capita economic growth and technological change are slower than in other scenarios. Additional information on these scenarios can be found in [15]. In this study, the A2 scenario is used as a representative of future years. World temperatures can increase or decrease in line with the greenhouse gas emissions presented by this scenario. Many other parameters can be affected by these changes, such as solar radiation, wind speed, cloud cover and relative humidity. Therefore, a general circulation model is required to mathematically predict such alterations. HadCM3, a general circulation model, covers grid points not only for Montreal, Canada, but over the entire planet. Since data from a general circulation model is limited to monthly averages, a downscaling method must be applied to convert the monthly averages to the hourly data required to use building energy simulation programs. In this paper, the typical meteorological weather file (CWEC) for Pierre Elliott Trudeau International Airport in Montreal is used as base weather data. The 'morphing' method was used to downscale the data from HadCM3 to generate hourly future typical horizon data for the 2020's, 2050's and 2080's decades from CWEC data. CCWorldWeatherGen a free Microsoft Excel-based tool developed by the University of Southampton, is used to transform base years into future climate change years, a method similar to that used in [21]. A set of 126 building simulations (126 roof configurations of varying thermal resistance and solar reflectance as discussed in 2.1) were conducted to evaluate the total energy consumption required for cooling and heating for each of the horizon future years and for the future typical years.

6.3 Results

A review of results from a simulation calculated with original CWEC data shown on the left side of Figure 6.1 indicate that heating energy consumption is dominant in the total energy consumption

balance. As is expected, heating energy consumption is highest when the building roof is simulated with low insulation values, and decreases when insulation is increased as per Fourier's law of heat conduction. Moreover, lower solar reflectance reduces the heating energy consumption.

By contrast, the building's yearly cooling energy consumption is lowest for a roof with high solar reflectance and low insulation, and highest with low reflectance and low insulation. Greater roof insulation reduces cooling energy consumption with low reflectance roofs, but increases it with high reflectance roofs, bridging the difference between albedo configurations. When added together, results show that the highest total energy consumption for this building occurs when solar reflectance and insulation values are low. Results also show that low solar reflectance on the roof is desirable during the heating season and undesirable in the cooling season.

This model is insightful but possesses the inherent flaw that it does not consider the enduring presence of snow, a highly reflective material, on flat roofs in the winter. Snow can therefore significantly increase the solar reflectance of a roof during the heating season. Snow also increases the overall thermal resistance of the roof contributing to less heating energy consumption, no matter with high or low reflective roof. Since snow is never present during the cooling season, it could be reasonable to assume that increasing the solar reflectance of a flat roof could have a positive effect in reducing the overall yearly energy consumption of a building. Furthermore, as is mentioned above, a building's roof is less exposed to sunlight during the winter months as it is during the summer months.

For simulations which were carried out with CWEC weather files that were morphed with GCM data for the 2020's and 2080's horizon years, Figure 6.1 – cooling, heating and total energy for a retail building using CWEC and horizon weather data shows that no matter which roof design is used, in future years, the cooling energy consumption will increase while the heating energy consumption will decrease.

		CWEC Cooling Consumption (kWh/m ²)														
		2	3	4	5	6	7	8	9	10	11	12	13	14	15	
Roof Solar Reflectance	0.1	15	14	14	14	13	13	13	13	13	13	13	13	13	13	
	0.2	14	14	13	13	13	13	13	13	13	13	13	13	13	13	
	0.3	13	13	13	13	13	13	13	13	13	13	13	13	13	13	
	0.4	13	13	13	13	13	13	13	13	13	13	13	13	13	13	
	0.5	12	12	12	12	12	12	12	12	12	12	12	12	12	12	
	0.6	11	12	12	12	12	12	12	12	12	12	12	12	12	12	
	0.7	11	11	11	12	12	12	12	12	12	12	12	12	12	12	
	0.8	10	10	11	11	11	12	12	12	12	12	12	12	12	12	
	0.9	9	10	10	11	11	11	11	12	12	12	12	12	12	12	

		2020's Cooling Consumption (kWh/m ²)														
		2	3	4	5	6	7	8	9	10	11	12	13	14	15	
Roof Solar Reflectance	0.1	20	19	18	18	18	17	17	17	17	17	17	17	17	17	
	0.2	19	18	18	17	17	17	17	17	17	17	17	17	17	17	
	0.3	18	18	17	17	17	17	17	17	17	17	17	17	17	17	
	0.4	18	17	17	17	17	17	16	16	16	16	16	16	16	16	
	0.5	17	16	16	16	16	16	16	16	16	16	16	16	16	16	
	0.6	16	16	16	16	16	16	16	16	16	16	16	16	16	16	
	0.7	15	15	15	16	16	16	16	16	16	16	16	16	16	16	
	0.8	14	14	15	15	15	15	15	15	15	16	16	16	16	16	
	0.9	13	14	14	15	15	15	15	15	15	15	15	15	15	15	

		2050's Cooling Consumption (kWh/m ²)														
		2	3	4	5	6	7	8	9	10	11	12	13	14	15	
Roof Solar Reflectance	0.1	26	24	23	23	22	22	22	22	22	21	21	21	21	21	
	0.2	24	23	22	22	22	22	22	21	21	21	21	21	21	21	
	0.3	23	22	22	22	22	21	21	21	21	21	21	21	21	21	
	0.4	22	22	21	21	21	21	21	21	21	21	21	21	21	21	
	0.5	22	21	21	21	21	21	21	21	21	21	21	21	21	21	
	0.6	21	20	20	20	20	20	20	20	20	20	20	20	20	20	
	0.7	20	20	20	20	20	20	20	20	20	20	20	20	20	20	
	0.8	19	19	19	20	20	20	20	20	20	20	20	20	20	20	
	0.9	18	18	19	19	19	19	19	20	20	20	20	20	20	20	

		CWEC Heating Consumption (kWh/m ²)														
		2	3	4	5	6	7	8	9	10	11	12	13	14	15	
Roof Solar Reflectance	0.1	102	94	89	86	83	82	80	79	79	78	77	77	76	76	
	0.2	103	95	89	86	84	82	81	80	79	78	77	77	76	76	
	0.3	104	95	90	87	84	83	81	80	79	78	78	77	77	76	
	0.4	106	96	91	87	85	83	82	80	79	79	78	77	77	77	
	0.5	107	97	92	88	85	83	82	81	80	79	78	78	77	77	
	0.6	109	99	93	89	86	84	82	81	80	79	79	78	77	77	
	0.7	110	100	94	89	87	85	83	82	81	80	79	78	78	77	
	0.8	112	101	95	90	87	85	83	82	81	80	79	79	78	78	
	0.9	114	102	96	91	88	86	84	83	81	80	80	79	78	78	

		2020's Heating Consumption (kWh/m ²)														
		2	3	4	5	6	7	8	9	10	11	12	13	14	15	
Roof Solar Reflectance	0.1	89	82	78	75	74	72	71	70	69	69	68	68	67	67	
	0.2	90	83	79	76	74	72	71	70	69	69	68	68	67	67	
	0.3	92	84	79	76	74	73	72	71	70	69	69	68	68	68	
	0.4	93	85	80	77	75	73	72	71	70	69	69	68	68	68	
	0.5	94	86	81	78	75	74	72	71	71	70	69	69	68	68	
	0.6	95	87	82	78	76	74	73	72	71	70	69	69	68	68	
	0.7	97	88	82	79	77	75	73	72	71	70	69	69	68	68	
	0.8	99	89	83	80	77	75	74	73	72	71	70	69	69	69	
	0.9	100	90	84	81	78	76	74	73	72	71	71	70	69	69	

		2050's Heating Consumption (kWh/m ²)														
		2	3	4	5	6	7	8	9	10	11	12	13	14	15	
Roof Solar Reflectance	0.1	81	75	71	69	67	65	64	64	63	62	62	62	61	61	
	0.2	82	76	72	69	67	66	65	64	63	63	62	62	61	61	
	0.3	83	76	72	70	68	66	65	64	63	63	62	62	62	61	
	0.4	85	77	73	70	68	67	65	64	64	63	63	62	62	61	
	0.5	86	78	74	71	68	67	66	65	64	63	63	62	62	62	
	0.6	87	79	74	71	69	67	66	65	64	64	63	63	62	62	
	0.7	88	80	75	72	69	68	66	65	64	63	63	62	62	62	
	0.8	90	81	76	72	70	68	67	66	65	64	64	63	63	62	
	0.9	92	82	77	73	71	69	67	66	65	65	64	63	63	63	

		CWEC Total Consumption (kWh/m ²)														
		2	3	4	5	6	7	8	9	10	11	12	13	14	15	
Roof Solar Reflectance	0.1	117	108	103	99	97	95	94	92	92	91	90	90	89	89	
	0.2	117	108	103	99	97	95	94	93	92	91	90	90	89	89	
	0.3	118	109	103	100	97	95	94	93	92	91	90	90	89	89	
	0.4	118	109	104	100	97	96	94	93	92	91	91	90	90	89	
	0.5	119	110	104	100	98	96	94	93	92	91	91	90	90	89	
	0.6	120	110	104	101	98	96	94	93	92	91	91	90	90	89	
	0.7	121	111	105	101	98	96	95	93	92	91	91	90	90	89	
	0.8	122	111	105	101	99	96	95	94	93	92	91	90	90	89	
	0.9	123	112	106	102	99	97	95	94	93	92	91	91	90	90	

		2020's Total Consumption (kWh/m ²)														
		2	3	4	5	6	7	8	9	10	11	12	13	14	15	
Roof Solar Reflectance	0.1	109	101	96	93	91	90	88	87	86	86	85	85	84	84	
	0.2	109	101	97	93	91	90	88	87	87	86	85	85	84	84	
	0.3	110	102	97	94	91	90	88	87	87	86	85	85	84	84	
	0.4	110	102	97	94	92	90	89	88	87	86	85	85	84	84	
	0.5	111	102	97	94	92	90	89	88	87	86	85	85	84	84	
	0.6	111	103	98	94	92	90	89	88	87	86	86	85	85	84	
	0.7	112	103	98	94	92	90	89	88	87	86	86	85	85	84	
	0.8	113	103	98	95	92	91	89	88	87	86	86	85	85	84	
	0.9	113	104	99	95	93	91	89	88	87	87	86	85	85	84	

		2050's Total Consumption (kWh/m ²)														
		2	3	4	5	6	7	8	9	10	11	12	13	14	15	
Roof Solar Reflectance	0.1	106	99	94	91	89	87	86	85	84	84	83	83	82	82	
	0.2	106	99	94	91	89	87	86	85	85	84	83	83	82	82	
	0.3	107	99	94	91	89	88	86	85	85	84	83	83	82	82	
	0.4	107	99	94	91	89	88	86	85	85	84	83	83	82	82	
	0.5	107	99	94	91	89	88	86	85	85	84	83	83	82	82	
	0.6	108	99	95	92	89	88	86	86	85	84	83	83	82	82	
	0.7	108	100	95	92	89	88	87	86	85	84	84	83	83	82	
	0.8	109	100	95	92	90	88	87	86	85	84	84	83	83	82	
	0.9	109	100	95	92	90	88	87	86	85	84	84	83	83	82	

Figure 6.1: Cooling, heating and total energy for a retail building using CWEC and horizon weather data

6.4 Conclusion

The research described in this paper shows that climate change will affect building energy consumption in future years and should be considered when designing HVAC systems today. Selecting reflectance and insulation values for a building roof should imply calculating heating and cooling consumption data over a period covering a building's lifespan to determine optimal configurations.

These results also show that heating energy consumption in a building is significantly reduced with higher levels of roof insulation and that increase in solar reflectance lead to reductions in cooling energy consumption. This indicates that cool roof designs are suitable for robust designs with respects to cooling energy. However, it should be noted that in cold climates like the one in Montreal, flat roofs are covered with snow during many of the heating days which leads to reduced solar effects. For these reasons, increased roof solar reflectance will have a minor effect on a building's heating energy performance in the winter. Therefore, in a robust building energy performance case, the total energy consumption variation will generally be affected by variations in cooling demand. In addition, using cool roofs would be even more attractive when larger scale benefits such as reduction in urban heat island effect is taken into account.

6.5 References

- [1] Y. Kang, S. Khan and X. Ma, "Climate change impacts on crop yield, crop water productivity and food security - A review," *Progress in Natural Science*, vol. 19, 1665-1674, 2009
- [2] U.S. Department of Energy, "U.S. Energy Sector Vulnerabilities to Climate Change and Extreme Weather," U.S. DOE, 2013
- [3] M. Snow and D. Prasad, "Climate Change Adaptation for Building Designers: An Introduction," *Environment Design Guide*, vol. 66, February 2011
- [4] M. F. Jentsch, P. A. James, L. Bourikas and A. S. Bahaj, "Transforming existing weather data for worldwide locations to enable energy and building performance simulation under future climates," *Renewable Energy*, vol. 55, 514-524, 2013
- [5] G. Chiesa and M. Grosso, "The Influence of Different Hourly Typical Meteorological Years on Dynamic Simulation of Buildings," *Energy Procedia*, vol. 78, 2560-2565, 2015
- [6] G. Sorrentino, G. Scaccianoce, M. Morale and V. Franzitta, "The importance of reliable climatic data in the energy evaluation," *Energy*, vol. 48, no. 1, 74-79, 2012
- [7] M. Hosseini, B. Lee and S. Vakilinia, "Energy performance of cool roofs under the impact of actual weather data," *Energy and Buildings*, vol. 145, 284-292, 2017
- [8] U.S. Global Change Research Program, *Global Climate Change Impacts in the United States*, Cambridge University Press, 2009
- [9] M. Santamouris, *Advances in passive cooling*, London: Earthscan, 2007
- [10] K. Al-Obaidi, M. Ismail and A. Rahman, "Passive cooling techniques through reflective and radiative roofs in tropical houses in Southeast Asia: A literature review," *Frontiers of Architectural Research*, vol. 3, no. 3, 283-297, 2014
- [11] M. Palme, L. Inostroza, G. Villacreses, A. Lobato-Corder and C. Carrasco, "From urban climate to energy consumption: enhancing building performance simulation by including the urban heat island effect.," *Energy and Buildings*, vol. 145, 107-120, 2017.
- [12] Y. Arima, R. Ooka, H. Kikumoto and T. Yamanaka, "Effect of climate change on building cooling loads in Tokyo in the summers of the 2030s using dynamically downscaled GCM data," *Energy and Buildings*, vol. 114, 123-129, 2016
- [13] M. Zhu, Y. Pan, Z. Huang and P. Xu, "An alternative method to predict future weather data for building energy demand simulation under global climate change," *Energy and Buildings*, vol. 113, 74-86, 2016
- [14] H. Kikumoto, R. Ooka, Y. Arima and T. Yamanaka, "Study on the future weather data considering the global and local climate change for building energy simulation," *Sustainable Cities and Society*, vol. 14, 404-413, 2014

- [15] M. F. Jentsch, C. K. Chang, P. A. B. James, A. S. Bahaj and Y. H. Yau, "Development of Climate Change Adapted Weather Files for Building Performance Simulation: Implications for Southeast Asia," in *The 3rd International Conference on "Sustainable Energy and Environment (SEE 2009)"*, Bangkok, 2009
- [16] S. E. Belcher, J. N. Hacker and D. S. Powell, "Constructing design weather data for future climates," *Building Services Engineering Research and Technology*, vol. 26, no. 1, 49-61, 2005
- [17] M. Zhu, Y. Pan, Z. Huang, P. Xu and H. Sha, "Future hourly weather files generation for studying the impact of climate change on building energy demand in China," in *13th Conference of International Building Performance Simulation Association*, Chambéry, 2013
- [18] R. S. McLeod, C. J. Hopfe and Y. Rezgui, "A proposed method for generating high resolution current and future climate data for Passivhaus design," *Energy and Buildings*, vol. 55, 481-493, 2012
- [19] D. Crawley, "Estimating the impacts of climate change and urbanization on building performance," *Journal of Building Performance Simulation*, 91-115, 2008
- [20] R. Loonen, M. Trčka and J. Hensen, "Climate adaptive building shells: State-of-the-art and future challenges," *Renewable and Sustainable Energy Reviews*, vol. 25, 483-493, 2013
- [21] A. Robert and M. Kummert, "Designing net-zero energy buildings for the future climate, not for the past," *Building and Environment*, vol. 55, 150-158, September 2012
- [22] American Society for Testing and Materials, "ASTM E1980, ASTM E1980: Standard Practice for Calculating Solar Reflectance Index of Horizontal and Low-Sloped Opaque Surfaces," Philadelphia, 1998
- [23] California Energy Commission, "Title 24. California Code of Regulations," in *Part 6: California Energy Commission*, Washington, 2010
- [24] H. Akbari and D. Kolokotsa, "Three decades of urban heat islands and mitigation technologies research," *Energy and Buildings*, vol. 133, 834-842, 2016
- [25] C. Piselli, M. Saffari, A. de Gracia, A. L. Pisello, F. Cotana and L. F. Cabeza, "Optimization of roof solar reflectance under different climate conditions, occupancy, building configuration and energy systems," *Energy and Buildings*, vol. 151, 81-97, 2017
- [26] V. Garg, S. Somal, R. Arumugam and A. Bhatia, "Development for cool roof calculator for India," *Energy and Buildings*, vol. 114, 136-142, 2016
- [27] Y. Qin, M. Zhang and J. E. Hiller, "Theoretical and experimental studies on the daily accumulative heat gain from cool roofs.," *Energy*, vol. 129, 138-147, 2017
- [28] M. Kolokotroni, B. L. Gowreesunker and R. Giridharan, "Cool roof technology in London: An experimental and modelling study.," *Energy and Buildings*, vol. 67, 658-667, 2013.

- [29] A. G. Touchaei, M. Hosseini and H. Akbari, "Energy savings potentials of commercial buildings by urban heat island reduction strategies in Montreal (Canada)," *Energy and Buildings*, vol. 110, 41-48, 2016
- [30] P. Ramamurthy, T. Sun, K. Rule and E. Bou-Zeid, "The joint influence of albedo and insulation on roof performance: An observational study," *Energy and Buildings*, vol. 93, 249-258, 2015
- [31] M. Hosseini and H. Akbari, "Heating energy penalties of cool roofs: the effect of snow accumulation on roofs," *Advanced Building Energy Research*, vol. 8, no. 1, 1-13, January 2014
- [32] M. Hosseini and H. Akbari, "Effect of cool roofs on commercial buildings energy use in cold climates," *Energy and Buildings*, vol. 114, 143-155, 29, 2016
- [33] J. Testa and M. Krarti, "A review of benefits and limitations of static and switchable cool roof systems," *Renewable and Sustainable Energy Reviews*, vol. 77, 451-460, 2017
- [34] City of Montreal, "Climate change Adaptation Plan for the Agglomeration of Montréal 2015-2020," Ville de Montréal, Montreal, 2015
- [35] "U.S. Department of Energy, Commercial Prototype Building Models," [Online].
- [36] "U.S. Department of Energy, EnergyPlus," [Online]. Available: <https://energy.gov/eere/buildings/downloads/energyplus-0>. [Accessed 09 2017]
- [37] "IPCC: Climate Change: Synthesis Report. Contribution of Working Groups I, II and III to the Fourth Assessment," IPCC, Geneva, Switzerland, 2007

Chapter 7. GENERATING FUTURE WEATHER FILES UNDER CLIMATE CHANGE SCENARIOS TO SUPPORT BUILDING ENERGY SIMULATION — A MACHINE LEARNING APPROACH

In the case of building energy performance, General circulation models (GCM) can be used to evaluate future building energy performance through simulations. However, there are key issues with the use of GCM data in building energy simulation. The first challenge is that the GCMs are usually biased, which means a considerable deviation can be found when the historical GCM data is compared to station observed weather data. The second challenge is that the GCM data has daily temporal resolution rather than the hourly resolution required in building energy simulation.

This chapter introduces a machine learning-based approach to process the climate change general circulation models (GCM) data for building energy simulation. The method statistically removes the bias in data and applies hybrid classification-regression machine learning algorithms to downscale the GCM data. The proposed workflow enables user to generate future weather files year by year under different climate change scenarios.

This chapter was published in the journal of Energy and Buildings, Volume 230, 110543, M. Hosseini, A. Bigtashi, B. Lee, “Generating Future Weather Files under Climate Change Scenarios to Support Building Energy Simulation — a Machine Learning Approach”, © Elsevier Ltd. 2021

7.1. Introduction

Historical data shows that Canada has experienced a warming rate of twice the global mean, while northern Canada has experienced triple the global mean [1]. According to the information presented by The Fifth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC) Working Group I [2], over the period of 1880 and 2012, the global temperature has increased by about 0.85 °C. A Canadian study reported the annual temperature rise ranging from 0.5 °C to 4 °C for 16 major cities in Canada over the period of 1900 to 2013. For Montreal, the average annual temperature rose about 2 °C, where a 1.4 °C rise was noted during the summer,

and a 2.7°C rise was recorded in winter [1]. Ongoing temperature rising with respect to climate change and its possible catastrophic consequences on building performance have made more and more engineers and architects consider sustainability in building designs. However, as a first step, it is important to clearly understand the possible problem and consequences in order to offer appropriate solutions. In the building and construction industry, the topic of climate change, and the possible impact on energy consumption of buildings has been studied by engineers and architects. Building energy simulation has been the main tool to evaluate the energy performance of buildings. In order to evaluate the building energy performance, meteorological data is required. Typical meteorological year weather files, which are used for simulations, reflect the historical weather condition and do not necessarily represent future weather data. In order to adequately represent future energy demand in building energy simulations, future weather data must be generated. The generated weather data must fit the purpose of the application. For example, in cases where the purpose of the building energy simulation is simply to estimate the future building energy consumption, a single future typical weather file representing a trend of multiple future years' weather might be enough. However, one important aspect of climate change is the increase in the number of heatwaves, which lead to significant consequences, including heatwaves related fatalities. As an example, during the summer of 2010 and 2018, heatwaves caused about 280 and 90 deaths, respectively, in the cold climate of Quebec, Canada [3]; of the 90 deaths in Quebec, 66 deaths occurred in Montreal, Quebec, Canada [4]. A survey conducted in 2011 showed that only half of the residential buildings were equipped with air-conditioning systems [5]. These facts might reflect the point that these buildings are not designed for extreme conditions, but rather, for typical cold climate weather conditions. This point directs the research toward creating future weather files that not only meet the intended initial use of finding typical building energy performance, but that also project the variability of performance for extreme conditions, including heatwaves. The generated weather files would enable architects, building engineers, and energy modelers to consider extreme future events at the building design stage in order to ensure that buildings and their associated energy systems could operate as expected under these conditions. There have been different types of weather files, which are used to evaluate typical and extreme weather conditions using energy simulation, that are reviewed in a previous study [6]. There are many challenges in handling climate change data for building energy simulation, which are explained in the following section.

Table 7.1: Nomenclature for phrases used in figures, tables, and texts

Term	Abbreviation	Explanation
Maximum dry-bulb temperature	max_temp	Maximum dry-bulb temperature in a day (°C)
Minimum dry-bulb temperature	min_temp	Minimum dry-bulb temperature in a day (°C)
Mean dry-bulb temperature	temp	Mean daily dry-bulb temperature (°C)
Maximum dew point temperature	max_dew	Maximum dew point temperature in a day (°C)
Minimum dew point temperature	min_dew	Minimum dew point temperature in a day (°C)
Mean dew point temperature	dew_temp	Mean daily dew point temperature (°C)
Maximum wind speed	max_wind	Maximum wind speed in a day (m/s)
Mean wind speed	wind	Mean daily wind speed (m/s)
Solar radiation	rad	Mean daily global horizontal solar irradiance (W/m ²)

7.1.1 Future climate change data and challenges relating to building energy simulation application

General circulation models (GCM) have been used by researchers to assess the effect of climate change in different fields of study. GCMs mathematically simulate atmospheric, oceanic, and biotic interactions and combine them with radiative forcing scenarios to evaluate the future climates. The models consist of grid cells resulted from latitude and longitudinal divisions, in which the meteorological data is calculated [7]. Although these models help considerate the impact of climate change, the output data of the models can't be directly used for building energy simulation. The challenges in applying the data are explained in the following section

GCM covers a vast geographical area, and the historical data based on the model is expected to deviate from the observed data of a specific location. Furthermore, the output of the circulation models has a coarse resolution, in terms of both spatial and temporal dimensions. The spatial resolution of the output from the circulation models is larger than 100 km x 100 km. As for the temporal resolution, typically, the output has a resolution of daily-average. This poses an issue for building energy simulation, which requires hourly meteorological data. In order to have a finer resolution, the data must be further processed. Finally, the GCM provides data for most weather parameters that used in typical meteorological weather files, with a few exceptions, notably, dew point.

Moreover, GCM suffers from systematic bias, which means that there is a considerable deviation when the historical data of GCM is compared to observed data at stations. The GCM systematic bias, coarse resolution, and limited available weather parameters make it difficult to apply the output data, without further processing, for energy simulation purposes.

Therefore, in summary, there are three main challenges to use GCM data for building simulation:

- Coarse spatial resolution
- Coarse temporal resolution
- Bias in the model data

From simple to sophisticated methods, previous studies used various approaches to address these challenges.

7.1.2 Previous approaches to address the challenges

Generally, there are two main approaches to process the GCM outputs; dynamical and/or statistical downscaling. Dynamical downscaling relies on further application of physics-based models for finer resolution outputs. Statistical downscaling, however, relies on the application of statistical rules and correlation for further processing. Each of these two approaches has pros and cons, which are discussed in the following section.

7.1.2.1 Dynamical downscaling

Depending on the availability of resources and expertise, different approaches are used to process the data in order to be used for building energy simulation. Some studies used dynamical downscaling, a method that requires an additional computationally intensive physics-based process based on a specific regional or local climate model. In this method, a regional or local climate model is nested in the GCM to use the GCM output as a boundary condition and take into account the hydrology, topography, and vegetation of the region to create finer resolution data. Kikumoto et al. [8] used GCM data as a boundary condition for regional climate models, namely, Model for Interdisciplinary Research on Climate (MIROC) and the Weather Research and Forecasting (WRF) to generate the future weather data of 2030s for a Japanese climate. The downscaled data was based on dynamical downscaling. For August, the sensible cooling load for a detached residential building increased by 15% from 2007 to 2030, mainly attributed to an average 1.52 °C rise in temperature comparing the two years [8].

Burger et al. [9] assessed the effect of climate change on the cooling and heating demand of an office building built during three different epochs: before World War I, after World War II, and from 2000 onward, in Vienna, Austria. They used REMO UBA regional climate model to

dynamically downscale A1B climate change scenarios with a resolution of about 10 km x 10 km. They evaluated the results for two timeframes ranging from 2011–2040 and 2036–2065. The downscaled data was used for building energy simulation. In one case, the results showed an average 41% increase in annual cooling and 56% decrease in annual heating compared to the period of 1961-1990 as a result of a temperature rise of about 3°C [9].

Although regional or local climate models provide finer temporal and spatial data, nesting of these models in the GCM models could be challenging as the time steps, and the grid resolution in the regional climate models are different from those in the GCMs. Moreover, physics-based calculations of finer resolution data would require expensive computational resources. In addition, when using dynamical downscaling, the bias of GCM data may not be removed completely, and further processing may still be required. Due to the mentioned restrictions, statistical downscaling techniques have been used to process the GCM data for various applications in different fields.

7.1.2.2 Statistical downscaling

Other studies used statistical downscaling methods. These methods rely on the availability of observed weather data and can be used for all scenarios of climate change. Statistical methods provide point-scale climatic parameters and can be applied to regional and/or global models. In the statistical approach, a critical assumption made is stationary, which means that, while the climate changes, the statistical relations among the meteorological parameters remain constant over time [10-11].

Generally, statistical downscaling is categorized into three main groups: linear methods, weather generators, and weather classification. The first two groups have been used for building simulation in previous studies and are briefly explained in the following sections, while the last group — weather classification, is one of the subjects of interest in this study.

Linear methods

Linear methods, including the delta method (also known as morphing in building simulation) [12], are easy to use and are widely applied in previous studies. In morphing, a changing factor is calculated by comparing the daily values of historical and future data and then applying it directly to the hourly observed data to achieve future hourly data. These factors can be additive or multiplicative or a combination of the two depending on the weather parameter. For example, the

downscaling of atmospheric pressure is additive (adding the change of means to each value for all hours), and the downscaling of the wind is multiplicative (multiplying the mean change to each value for all hours). These applications are done separately for each weather parameter and, therefore, the correlations among the weather parameters are ignored. Therefore, the morphing method can change the mean and variance of the historical data to the GCM data, but other statistics of the data such as the 25th or 75th percentile of the data do not necessarily change. In other words, the morphing method might transfer the intensity of the GCM data to the downscaled data, but it may not transfer the frequency.

Due to its simplicity, the method is widely used to predict the future climate condition, including in American climate [13-16], Canadian climate [17-19], Swiss climate [20], Swedish climate [21-22], Spanish climate [23], Italian climate [24], British climate [25], and Chinese climate [26].

The morphing method is designed to use a TMY file, and the same is applied in these studies as well. The TMY is generated from historical data with extreme events removed [11]. The application of morphing on a single TMY might lead to project a single annual building cooling or heating energy consumptions for a long period of time. The mentioned feature makes the morphing method suitable as long as the goal of the application is to estimate the building energy consumptions over a long period of time. However, in cases where multiple weather years are required, the morphing method may not be suitable. There are some tools that use the morphing method to downscale the future GCM weather data, such as “CCWorldWeatherGen” or “WeatherShift™ tool”, which are used in building simulation [27-29]. One important point in using these tools is respecting the time frame of the base TMY file as it should be between the years 1961-1990 for the former and 1976-2005 for the later tool; otherwise, the output would be overestimated.

However, when the application is to optimize the design e.g. robust design, reliable design, or other design aspects, multiple weather years, are required. This is because of the fact that these design methods rely on statistics requiring multiple samples. For example, for designing a weather-robust building, providing multiple weather years would be an essential requirement for design optimization. As another example, for reliability-based design, extreme events must be taken into account. However, these events are disregarded by the morphing method on TMY, as it ignores the frequencies in the processing of data (e.g. frequencies of multiple warm days/heat waves) [20].

Moreover, using the morphing method may not remove all the existing biases between the GCM historical data and the observation data. As such, weather generators are also used to generate multiple future weather years stochastically to capture the yearly variations over a long period of time.

7.1.2.2.1 Weather generators

Other studies used weather generators to evaluate the effect of climate change on buildings [26, 30-32]. Weather generators are used for temporal downscaling. These statistical models generate numerous possible time-series weather parameters using statistics of several-years historical data of weather parameters applied to the GCM model output. Due to the stochastic nature of data generation, the generated data might change when the process is repeated and may require further processing to make a single weather file. The data generated from the generators are different from the observed data and only keep the statistical characteristics of mean and variance of the GCM data. Therefore, all the generated data is artificial data that synoptically have the same statistics as the GCM data. Among the weather generators, only a few are able to consider the relationship between the weather elements when multiple parameters are predicted [11], and they are able to generate only a few weather parameters such as temperature and solar radiation [31]. Furthermore, thousands of different generated data is produced by weather generators. Therefore, in order to adapt the data for building energy simulation, further processing is required [26].

To prepare the generated data for extreme analysis, Nik [22] suggested creating a typical downscaled year (TDY) together with one extreme cold year (ECY) and one extreme warm year (EWY). In order to create these weather years, the same procedure of making a typical meteorological year (TMY) is used, except that only the temperature variable is considered in the procedure e.g., one year with all months having typical temperature values (TDY), one year with all months having low-temperature values (ECY), and one year with all months having high-temperature values (EWY). The reason behind considering only the temperature was stated to be “the difficulties and uncertainties in weighting the climatic variables”. The same method was used in a newer study to downscale regional climate model data in order to consider typical and extreme conditions for the projection of weather data for different future timeframes, namely, 2010-2039, 2040-2069, and 2070-2099 [20]. Although the method reduced the required number of simulations for building design optimization, the method seems to create artificial extreme weather years that

may never happen in the future, especially without the bias-correction of the RCM data that might overestimate the extreme conditions.

The issue of the weighting of the climatic variables mentioned by the previous two studies [20, 22] is recently addressed with a novel method in the previous study conducted by the authors for building energy simulation application [33] (chapter 5).

7.1.3 Research gap

From the literature review, it is perceived that the morphing method may not be used for design optimization of the building and HVAC systems as it ignores the frequency of data, and therefore, it ignores the extreme conditions such as heatwaves in summers. Moreover, the morphing techniques are applied to each of the weather parameters separately and consequently ignoring the correlations among the weather parameters.

Weather generators, on the other hand, may be used for design optimization; however, the method generates a large amount of artificial data that only maintains the mean and variance of the GCM data for each weather parameter separately, ignoring once again the correlation between the parameters. The literature review showed that previous studies mostly used morphing or weather generators. In other industries, weather classification or weather typing schemes is another method used for downscaling the GCM climate change data. However, due to certain limitations, this method has not been deployed in the building industry.

The other method of downscaling the GCM climate change data in other industries is weather classification or weather typing schemes. Weather classification relates a class of future weather patterns to locally observed weather data, and the future weather data are synoptically selected from the observed weather data. In this method, the effect of climate change is estimated by evaluating the frequency and intensity of the change of the weather pattern parameters from the GCM output. In other words, the future GCM data is compared to historically observed data, and a class of historical data is selected according to statistical resemblance. The selected historical weather pattern values will then represent the future climatic weather condition. The method can be used for normal and non-normal weather parameters such as temperature and wind speed. However, a large set of historical observations is required [11]. This method is more sophisticated

than the simple morphing method in the case of data analysis but can create real data selected from the past and therefore keeps the real relationship among the weather parameters in hourly values.

In contrast to the morphing method, this method can transfer all the information of GCM data (including percentiles) to the downscaled data. This method has been used in hydrological studies [34-35], but there is a lack of literature in applying the method to building energy performance. In the building energy simulation industry, weather typing or classification is used to create a representation of the long-term (usually 10 to 30 years) weather pattern from the historical data using a typical meteorological year. The Sandia method, a method for generating typical meteorological year files, can be categorized in this group. However, the method is not used to downscale the future GCM data. The main disadvantage of this method is its inability to predict new values that are beyond the range of the historically observed data (increase in intensity). Building energy consumption is highly correlated with outdoor air temperature, which is expected to increase in the future unprecedentedly. Therefore, the weather classification alone, may not be used to downscale the future climate data for building energy consumption. In this study, a hybrid classification-regression model is proposed to downscale climate change GCM data.

7.1.4 A proposed method to use weather classification downscaling for building simulation

Weather classification can retain all the statistics of the GCM data; however, it is incapable of keeping the intensity of data when a new out-of-range of historical data is in the GCM data. Therefore, the combination of the weather classification with a regression model as a hybrid model can be a promising method to downscale the GCM or RCM data to be used for building energy simulation. The method can partially apply an algorithm that is currently already used in the building simulation industry and uses real historical weather data rather than artificially generated values. The regression model will be trained using observed historical data and will be used only for conditions when the GCM data is higher or less than any previously experienced data. The hybrid application of this method can transfer all the statistics of the GCM data while keeping the intensity of the data as well.

There is a shortage of studies (if any) that apply weather classification as a method to downscale the GCM data to be used for building simulation. The reason for that, as mentioned before, is that the building energy performance is highly dependent on outdoor air temperature, which is expected to increase in the future; the classification model is not properly downscaling the new values that

are higher or less than any observed values. Therefore, a regression model will be combined with the classification model for new unseen values larger or smaller than the observed values. This method is introduced in this study as part of the methodology.

7.2 Objective and organization of the study

This study attends to the issues in applying GCMs to generate future weather files, namely removing bias within data, as well as, the application of a coarse resolution general model to a specific location and scaling of the daily data to hourly data.

The objective of this study is to provide a systematic workflow to create hourly weather files for individual future years under different climate change scenarios using GCM future data. The workflow can be applied to regional climate models (RCM) data as well.

This study introduces a new hybrid machine learning algorithm pairing the weather classification model with a regression model to process the climate change data to be used for building energy simulation. The classification model used in this study is introduced as the main model to downscale the GCM data for building simulation. An auxiliary regression model is also trained on the observed station data to predict (regress) the climate change values for cases when the values are larger than any observed values. The latter solves the challenge of using the classification method to downscale the GCM data for building simulation.

The proposed methodology would be suitable to create multiple year by year future hourly weather under four climate change scenarios that can be used by architects, building energy modelers, and engineers to consider typical and extreme weather conditions at the design stage of the buildings. The generated weather years can be used for building and energy system optimization, especially those designs that use statistics e.g. reliability-based design or robust-based design optimization.

The current study is outlined as follows:

Section 7.3.1 explains the process for selecting weather parameters deemed essential for building energy simulation, which is then weighted for further processing. This is done due to the fact that in cases where weather parameters are highly correlated, processing only one of these parameters is sufficient. The weather parameters weights and data will be used in the weather classification model. In section 7.3.2 the bias-correction method is explained to remove the existing bias in the

GCM data statistically. This step is essential since the GCM models ignore the topography, vegetation, and urban effects, and bias correction can address these issues.

Section 7.3.3 explains the spatio-temporal method used to downscale the bias-corrected GCM data to be used for simulation. The classification and the regression algorithms used for downscaling are explained in this section.

Section 7.4 presents the results followed by the discussion and a summary of the contributions of this study.

7.3 Methodology

Geophysical Fluid Dynamics Laboratory Coupled Model Earth System Model (GFDL-ESM2M) is a global climate model developed at the NOAA Geophysical Fluid Dynamics Laboratory and is consistent with the IPCC Fifth Assessment Report (AR5) [36]. The model provides a range of weather parameters including temperature, solar radiation, wind speed, and many other parameters with a resolution of about $2.0^\circ \times 2.5^\circ$ along the latitude and longitude. The model output includes the dataset for the four Representative Concentration Pathways (RCP) 2.6, 4.5, 6, and 8.5 W/m^2 , ranging up to the year 2100. Figure 7.1 shows an example of near-surface air temperature data in $^\circ\text{C}$ from GFDL-ESM2M for the 21st of July 2020 and 2049. A bilinear interpolation (interpolation among 4 points rather than 2 points) is conducted to extract the data at Montréal–Pierre Elliott Trudeau airport with a longitude of -73.75° and latitude of 45.47° .

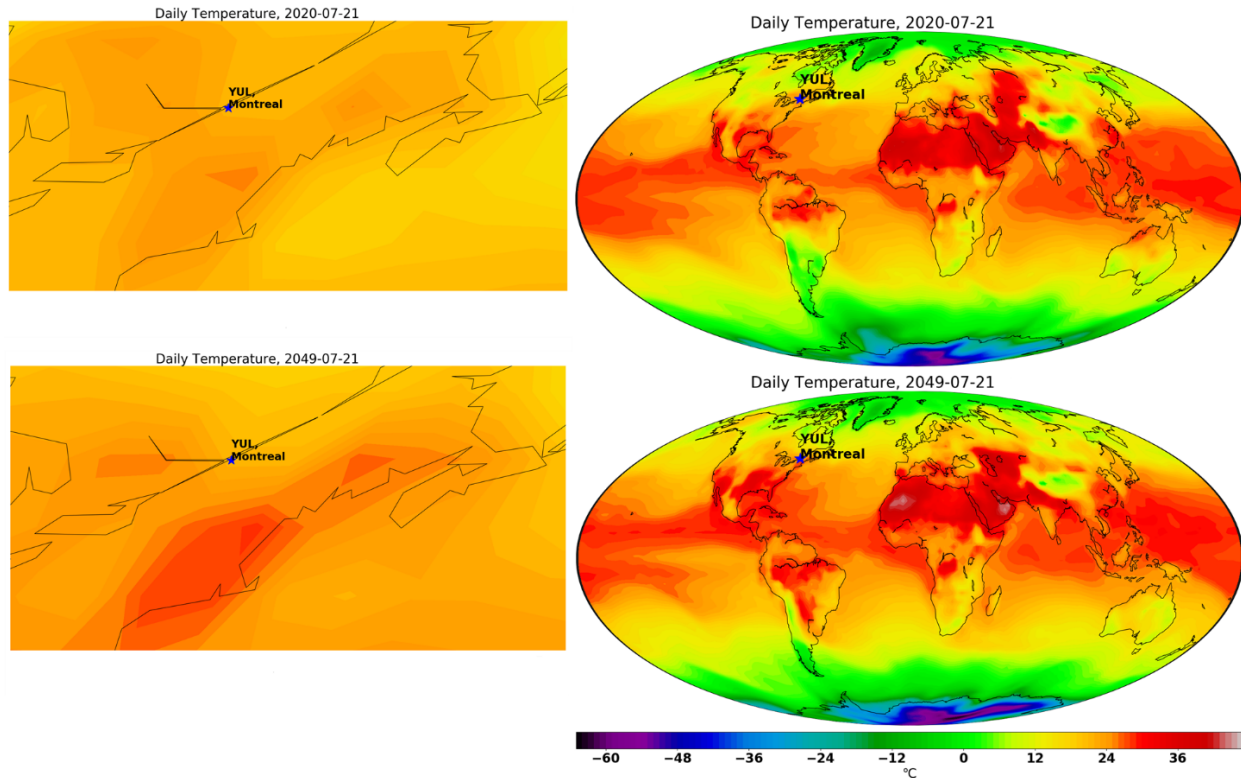


Figure 7.1: The GFDL-ESM2M, near-surface daily air temperature for 2020-07-21 and 2049-07-21 under RCP 8.5, left figures show the location of Montréal–Pierre Elliott Trudeau airport.

Typical meteorological year weather files being used for building energy simulations usually include a set of weather parameters such as horizontal solar irradiance, dry bulb temperature, dew point temperature, atmospheric pressure, and wind speed [37-38]. These weather files are used in building energy simulation as a means to evaluate building energy performance. In the previous study of the authors [33] the effect of each weather parameter on cooling, heating, and total energy consumption of the buildings are statistically evaluated. The GCM provides some of these weather parameters, with quite a few missing such as dew point temperature. To attend to the issue, an analysis is presented in the next section to select the required climatic data from the GCM data for further processing. Since some of the weather parameters are highly correlated, only one parameter among the correlated ones is selected for further processing, e.g. classification; this is explained in section 7.3.1.

Table 7.2 provides an overview of the workflow. Detailed implementation of each step is provided in subsequent sections.

Table 7.2 Overview of the workflow

Step	Background	Description
Weather parameter selection	To reduce the number of evaluated parameters and avoid redundancy as some parameters are highly correlated.	The correlation between the weather parameters are evaluated. For each case of highly correlated parameters, a “representative” or candidate parameter is selected for further processing.
Bias-correction	To address the existing discrepancy between the historical GCM data and the observed data due to low spatial resolution.	Application of quantile-quantile method, which corrects the bias by determining the difference between the CDF of the historical and observation data is then applied to the future data, quantile by quantile.
Spatio-temporal downscaling	To downscale the weather data to the required resolution for simulation purposes.	Application of two machine learning algorithms to process the climatic data. Each algorithm provides the required output data. A threshold is selected in order to evaluate the deviation of the historical month temperature from the future temperature. In the case that the threshold for temperature is not met, the regression model data is selected (Random Forest), otherwise, the K-nearest all data is selected.
K-nearest-neighbor	Algorithm 1	The weighted K-nearest neighbor algorithm is used to select the most similar historical month weather pattern to the future month. It is done by applying the FS statistics method to the CDF of each of the weather parameters. The feature importance of each parameter is applied and the most similar historical data is selected, with the corresponding hourly output data.
Random Forest Regression	Algorithm 2	A regression model is trained and tested using historical data. Historical hourly observed temperature is resampled to average daily and the corresponding hourly data is provided as training output. The regression model is then trained to provide corresponding hourly data for each month of daily average GCM data input.

7.3.1 Weather parameters selection

In order to reduce the dimensionality of data, the correlation coefficient of ten weather parameters are determined in order to eliminate the parameters with high correlations. Reducing the number of parameters contributes to better handling of the data and a more accurate prediction model in classification. The selected weather parameters will define a class of weather in the classification model. Following this, a dataset that includes 10-years historical hourly weather data [39] is generated to find the correlation among the weather parameters. The entire dataset is then resampled to daily average data as the classification model finds a very similar class of daily weather data to the GCM daily data. The considered parameters are total global horizontal irradiance, average dry bulb temperature, maximum and minimum dry bulb temperature in a day, average dew point temperature, maximum and minimum dew point temperature in a day, atmospheric pressure, average wind speed, and maximum wind speed in a day. The heatmap

presenting the correlation coefficient between the parameters is plotted in Figure 7.2. Values approaching 1 or -1 show a more linear correlation among the weather parameters. A high linear correlation (more than 0.93) between the average temperature, the minimum and maximum temperature in a day, average dew point, the maximum and minimum dew point is noted, as represented in Figure 7.2. There is also a high linear correlation between the average wind speed, and the maximum wind speed in a day (0.86). Therefore, the average temperature is selected as a representative for all the temperature parameters, and the maximum wind is represented by average wind speed. The atmospheric pressure doesn't show a considerable correlation with others; therefore, it is also considered as one of the selected parameters.

The four weather parameters, average dry bulb air temperature, global horizontal solar irradiance, atmospheric air pressure, and wind speed are selected for further processing. The next step is to compare the historical data of GCM data with observed data over the same period of time for these four selected parameters. As mentioned in section 2, due to reasons such as limited spatial resolution, simplified physical and thermodynamic processes, the historical data of the GCM model is deviated from the weather station data, which implies that the GCM future data cannot directly be used. Therefore, a statistical bias-correction is proposed to remove the bias from the future model data.

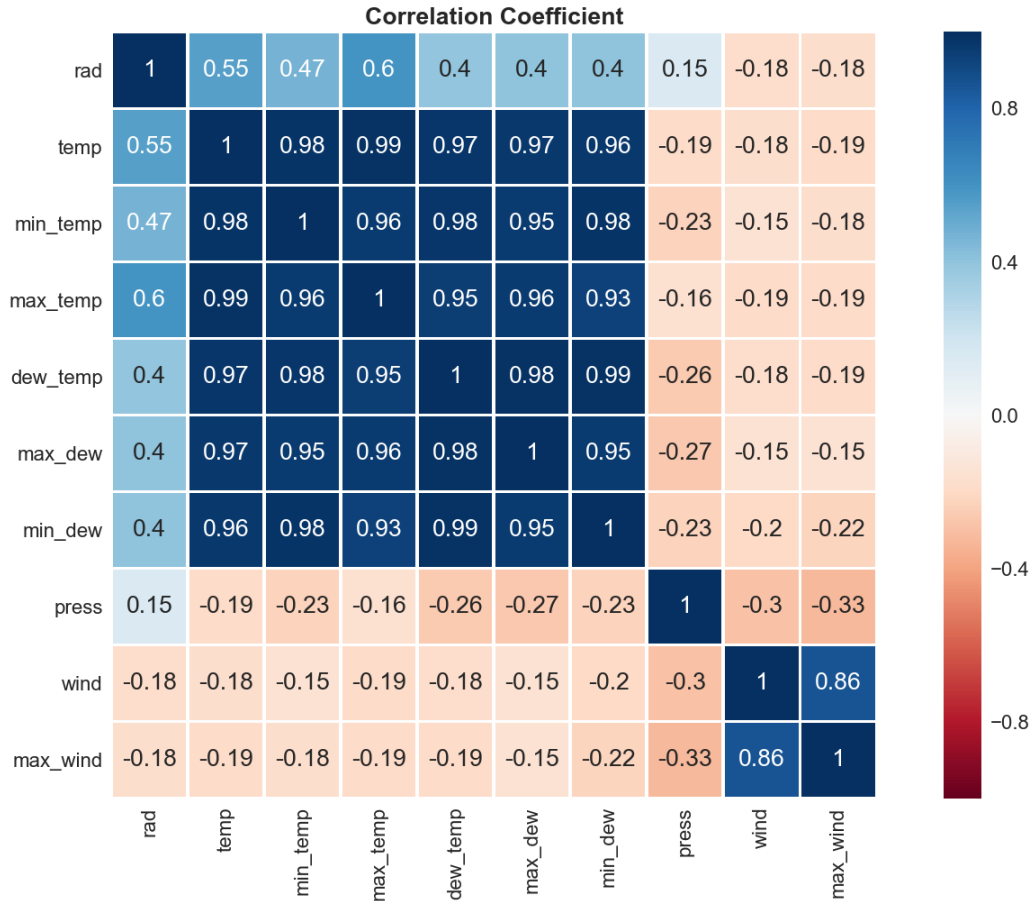


Figure 7.2: Correlation coefficient for parameters; darker colors represent stronger linear correlation.

7.3.2 Bias-Correction

GCM data doesn't include the effect of regional and local features such as vegetation and topography as the spatial resolution is low. This may lead to considerable differences when comparing GCM historical model data to observed data at a local station over the same period. In order to avoid this discrepancy, the GCM data requires calibration before proceeding to further processing. In older applications, two general approaches were applied; the first approach, known as the delta method, is to calculate the difference between the future and historical GCM data and add it to the observation data (station data). In the second approach, the bias can be removed by finding the difference between the observed station data and the historical GCM data and apply it to the future GCM data. Each of these approaches has to base on its own controversial assumption to apply. For the first approach, it is assumed that there is no change in the variation of the weather

data in the future. For the second approach, it is assumed that the future data provided by the model is accurate.

A newer approach, known as the quantile-quantile method, introduced by Amengual et al. [40] is used in this study. The method demonstrates more flexibility in bias-correction and is applicable to all climate parameters. In this method, instead of directly adding the difference to the future model or observed data, the difference is applied according to a statistical method — the future model data is corrected according to the cumulative distribution function (CDF). The difference between the corresponding CDFs for historical model data and the observations data are calculated and are then applied to the CDF for the future model data quantile by quantile. The advantage of this method is that all the statistics of data, including mean, variability, and shape of the GCM data will change according to the statistical changes in the calibration period. Amengual et al. [40] presented further details. Figure 7.3 shows an example of the function of quantile-quantile bias-correction for temperature. The blue color represents the CDF of the daily averaged observed station data, the red color represents the CDF of the GCM historical data, and the green color represents the CDF of the GCM future data. Under CDF values, the difference between observation and GCM historical data are calculated (red arrows which denote the area between the blue and red line), and are subsequently added to the corresponding values of the same CDFs of the GCM future data (green arrows which denote the area between the yellow and green line) to calculate bias-corrected GCM future data.

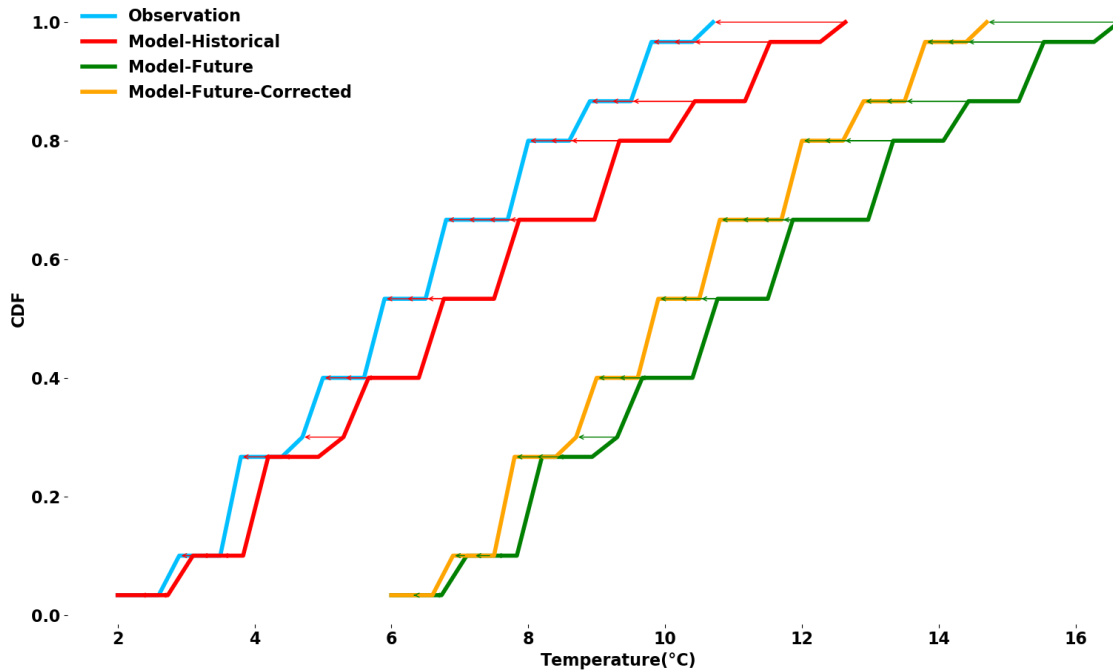


Figure 7.3: Quantile-Quantile bias-correction method; the future model data are statistically corrected according to historical deviation from observation.

In order to correct the existing bias in the GCM data from 2020 to 2049, a 30-year period of historical data, ranging between 1976 and 2005, is selected as the calibration period. Figure 7.4 shows the daily average temperature data for the calibration period and future data. From the figure, we observe that the GCM data is underestimated in maximum and minimum temperatures (red and blue) for most of the period. The underestimation is corrected for future GCM data. As it is shown, the corrected GCM future data (orange) surpasses the original GCM future data (green) in maximum and minimum temperatures.

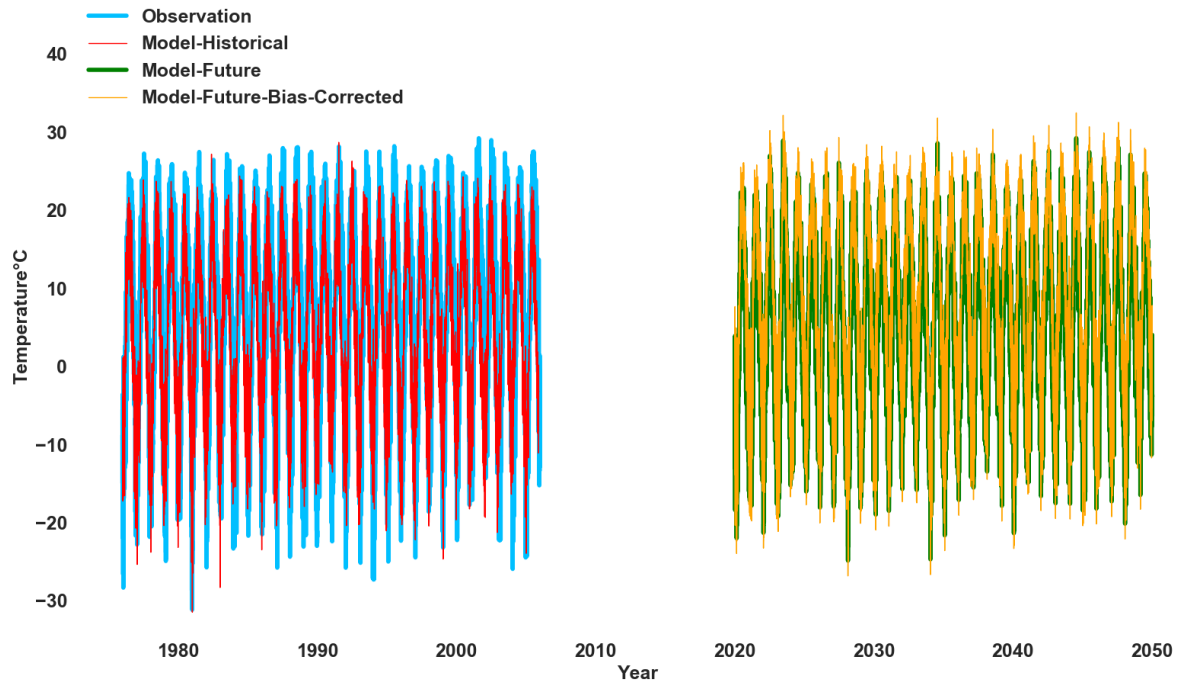


Figure 7.4: The validation period historical GCM data and historical observed data from 1976 to 2005 (left) [39] versus original GCM future data and bias-corrected GCM data from 2020 to 2049 (right).

Figure 7.5 zooms into the first year of the calibration period and the future period data for clearer illustration. As seen in the figure, during the first four months of the calibration period, the GCM model (red) over and underestimates the temperature (compared to observation). The fluctuation between over and underestimation leads to the bias-corrected data (orange) surpassing the original GCM data (green). For the rest of the year, the GCM mostly underestimates the data, and therefore, the GCM bias-corrected data is mostly higher than the original future GCM data (green).

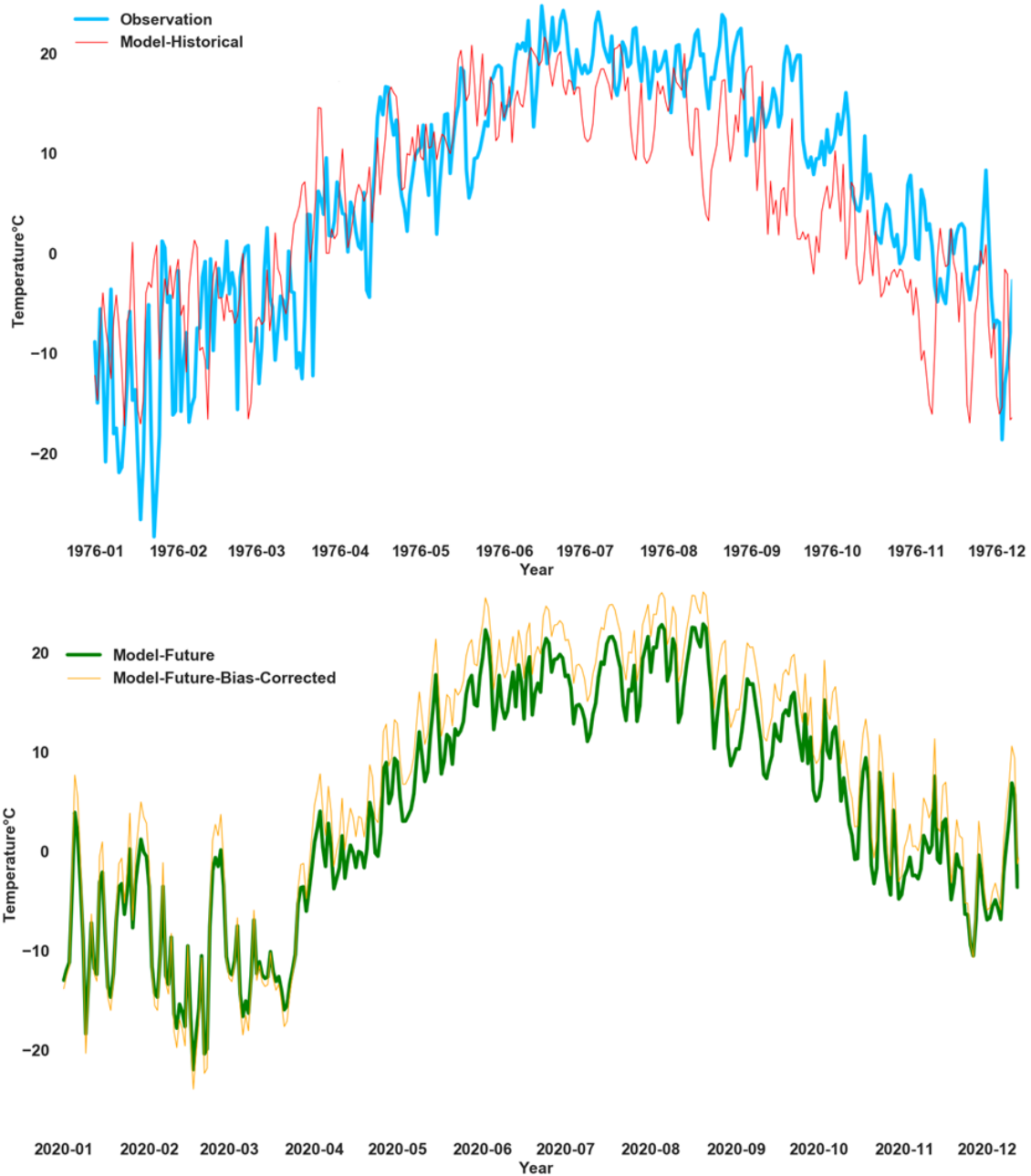


Figure 7.5: The first year (1976) of the validation period historical GCM data and historical observed data (top) versus the first year (2020) of original GCM future data and bias-corrected GCM data (bottom).

In Figure 7.6, the bias-corrected data is averaged yearly and compared to the corresponding original GCM future data for the future period.

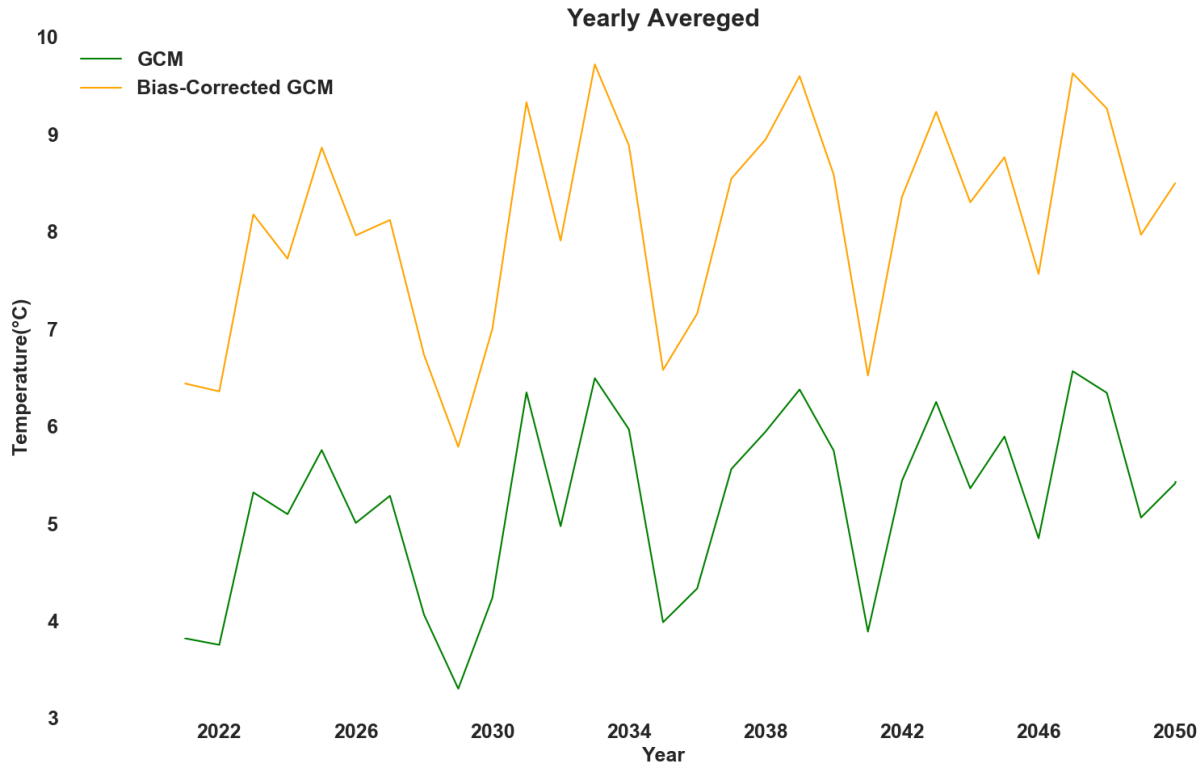


Figure 7.6: Yearly-averaged of original future GCM data versus the bias-corrected data for the period of 2020 to 2049.

The original GCM model shows annual average temperature within the range of about 3.5 to 6.5 °C, whereas the bias-correction increased the range to about 5.5 to 9.5 °C. As previously mentioned, the increase in temperature is due to the consideration of local station data, which account for the effects of topography and vegetation that are ignored in the original GCM data. For this study, the airport data is considered as the observed data; however, local station data located in an urban environment may also be used to consider the local climate effect.

For further capturing of the local effects from a larger scale, spatial downscaling is required. The output of GCM is set to the daily average, whereas hourly data is required for building energy simulation. Therefore, a spatio-temporal downscaling is applied to convert the daily-averaged data into hourly resolution data.

7.3.3 Spatio-temporal downscaling

A hybrid method of weather classification-regression is used to downscale the weather data, meaning that two machine learning algorithms are used to process climate change data. The reasons for combining the two suggested algorithms, as stated in section 7.1.3 are:

- The weather classification is capable of downscaling the data as long as the GCM data has values within the range of the observed data.
- In cases where the GCM data values are considerably higher or lower than any values in the observed data, a regression model is trained from the observed data to downscale the GCM daily data to hourly data.

Both algorithms use local station data as the observed data, and the data of GCM is extracted for the same location (Montreal international airport). The models disaggregate the data from daily average to hourly data; therefore, the current method can be referred to as spatio-temporal downscaling. This means that the data is spatially and temporally downscaled at the same time. The k-nearest-neighbor classification algorithm helps to select the most similar weather pattern from the observed station data and this can be valuable approach as the observed historical data is actual measured data which captured the characteristic of the local environment.

7.3.3.1 k-nearest-neighbor algorithm for weather classification

In pattern recognition, the k-nearest neighbor machine learning algorithm is used for classification, which doesn't require training and is, therefore, quite faster than the training-dependent algorithms. The algorithm is simple and easy to implement and is widely used in classification problems, especially in situations where the decision must be made for problems with non-normal distribution. This capability is helpful in this study with non-normal variables, including wind.

Depending on the type of dataset, different techniques, including root mean square error, can be applied to find the neighbors. Similarly, the Sandia method, which is used for generating typical meteorological weather files, uses the Finkelstein-Schafer (FS) model to statistically find similar neighbors [38]. The weighted k-nearest neighbor (KNN) algorithm is used to find the absolute value of the difference between predictor and the neighbor predictand. The predictor is the future GCM data (bias-corrected), and the predictand is the historical weather station observed data. In

other words, for each future month weather data, the most similar historical monthly weather pattern is selected as long as the selected month shows an acceptable level of closeness.

The following procedure is used to downscale future bias-corrected GCM daily-averaged data. For each future month of a year:

- i. The cumulative distribution function (CDF) of each of the weather parameters, namely, global horizontal irradiance, average temperature, average wind speed, and atmospheric pressure, are calculated.

- ii. The FS criteria are calculated for each parameter according to equation 1:

$$FS = (1/n) \sum_{i=1}^n \delta_i \quad (1)$$

Where: δ is the absolute difference between future parameter CDF and the historical candidate month parameter CDF, and n is the number of readings in a month (here is 30) (See Figure 7.7).

- iii. The most similar weather pattern is selected by weighting the parameters based on the importance of the parameters on the building energy demand. These importance factors are calculated by the feature importance technique used in machine learning and described in section 7.3.1.1. The minimum WS factor will present the most similar weather pattern of historical weather to the future model data, which will be selected according to equation 2:

$$WS = \sum w_i FS_i \quad (2)$$

Where: w_i is the weather parameter weighting factor (importance) shown in Figure 7.8, and FS_i is weather parameter FS statistics (from step ii).

- iv. The hourly weather data belonging to the minimum WS is used to fill up the corresponding month of the twelve-month weather file.
- v. The procedure is repeated for the next month of the future year until all 12 months of the future years are found from historical observations.
- vi. The whole previous steps are done for all the future years (30 years).

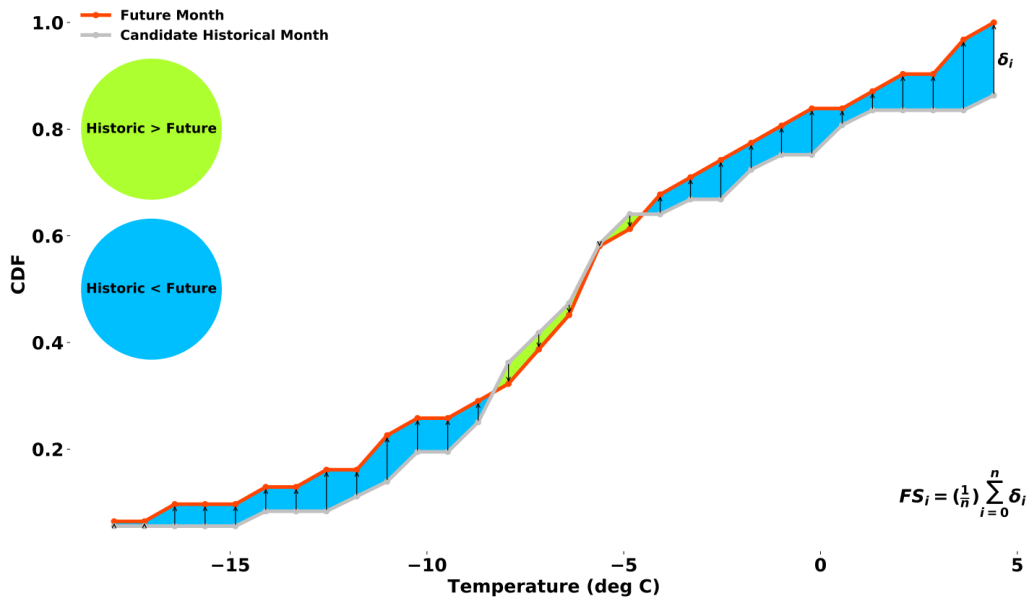


Figure 7.7: FS statistics for temperature, the absolute difference between the CDFs of the predictor and predictand over n bin points (30) is calculated (step ii).

Figure 7.7 shows the comparison of a future month GCM temperature data CDF (red) with a candidate month observed temperature data (grey). All the candidate observed months are compared to the future month and FS is calculated for all the candidate months. Since there are 4 selected weather parameters, for each month, FS for all the weather parameters are calculated, and the final selected month would be based on weighted-sum FS (WS in equation 2). The weighted sum approach in the equation helps the algorithm to select the historical months with the most statistical resemblance based on the most important weather parameters. The importance is relative to the effect on the output (i.e. cooling, heating, or total energy demand). In order to apply the WS, the weighting factors of each weather parameter are calculated, which is explained as follows.

7.3.3.1.1 Weighting factors of weather parameters for WS

The k-nearest neighbor algorithm, described in section 3.3.1, requires the importance of each weather parameter on the building energy demand to be evaluated (equation 2). The k-nearest neighbor and Random Forest algorithms have some similarities. Both cases make predictions from models constructed from training datasets by looking at neighborhood setup by applying weighting

factors [41]. The weighting factors calculated from Random Forest regression may be applied to KNN. These weighting factors help with the final decision-making process by finding the final overall most similar weather pattern to the future data from historically observed data. The procedure of finding the weather parameter importance is described in details in another study by the authors [33] (chapter 5). A brief description is provided in the following section, which includes the modifications made for this study.

To illustrate the procedure, a U.S. Department of Energy prototypical small office building [42] upgraded to National building Energy Code of Canada (2015) standards is simulated with 10-year historical weather data from 2009 to 2018, taken from the Montreal-Pierre Elliot Trudeau International airport. The simulation was conducted using EnergyPlus 8.9 building energy simulation program.

The Random Forest regression algorithm is used to statistically find relationships between the parameters, as it can capture the non-linear relationship among the parameters. It should be noted that, at this step, the Random Forest algorithm is only used to find the feature importance by fitting all the given data from 2009 to 2018. Since the goal at this step is not making a prediction, there is no test data at this step. Once the algorithm is fitted to the data, it is able to extract the feature importance from the data. To find the feature importance, the total energy demand is assumed to be the dependent variable, while the four weather parameters of average temperature, global horizontal solar irradiance, average wind speed, and the atmospheric pressure are considered as continuous independent variables.

The importance is defined as the importance of each of the weather parameters in predicting the total energy demand and is calculated based on variance reduction. Then the importance of each weather parameter is calculated. For more details, refer to the study by Hosseini et al. [33] (chapter 5).

Since the effect of each weather parameter on total energy demand in varying months might be different, a specific set of weighting factors is calculated for each month of the year. Figure 7.8 shows the feature importance of each weather parameter on predicting the total energy demand of the building.

As expected, the temperature gives the most information (variance) in predicting the total energy demand in all months of the year. The radiation is the second information contributor in predicting the total energy demand in six months of the year, although the radiation, atmospheric pressure, and the wind are observed to have similar importance and fairly small in all the months of the year. Following the overview of the k-nearest-neighbor algorithm and how the weather importance is calculated, the next section describes how the algorithm is applied for weather classification using weather data.

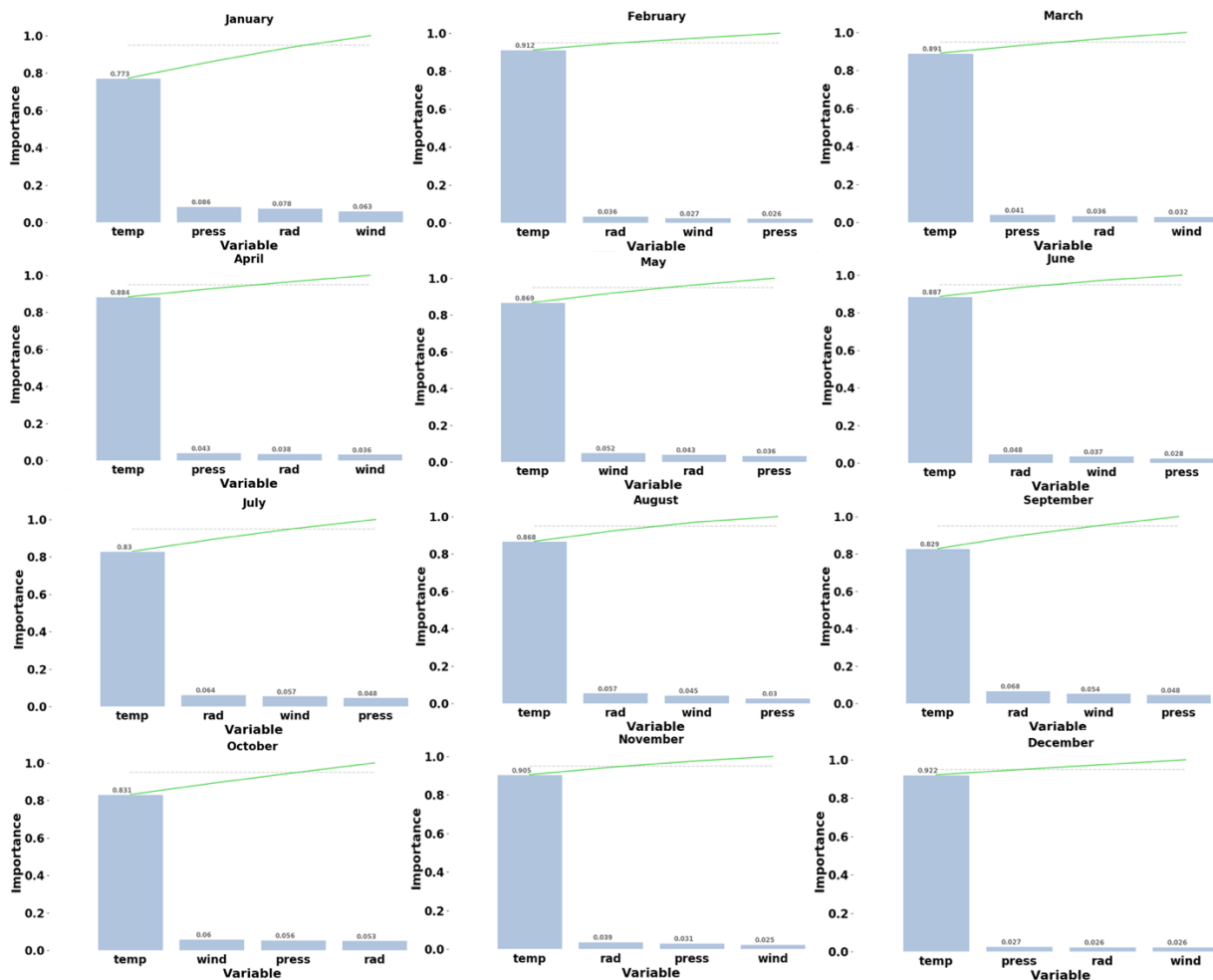


Figure 7.8: The feature importance (weather parameter weighting factors) extraction from Random Forest Regression to be used for KNN classification algorithm.

The previous sections explained the k-nearest-neighbor algorithm and how the weather importance calculated. The next section describes how the algorithm is applied for the weather classification purpose using weather parameters data.

7.3.3.1.2 Application of k-nearest-neighbor

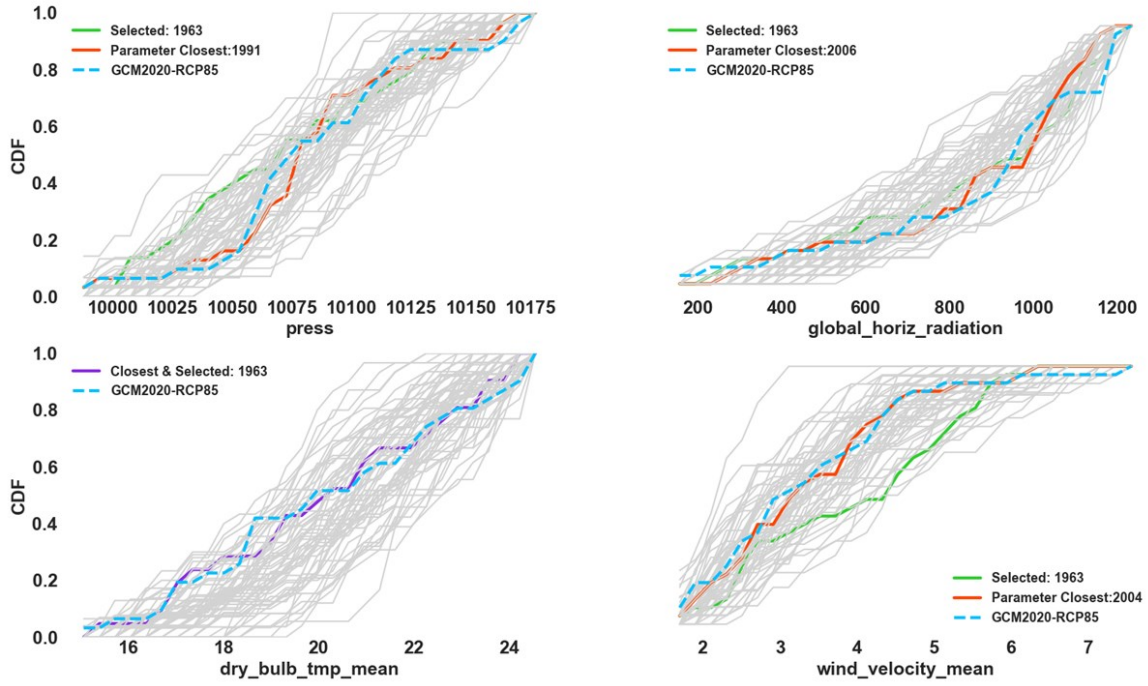
The weighting factors shown in Figure 7.8 are introduced to the classification model (KNN) to calculate WS from step iii.

Figure 7.9 shows the application of the FS statistics method for weather parameters. The blue color shows the CDF average daily values of a future month weather parameter. The red color shows the CDF of the most similar historical daily average weather data to future data, namely, the closest CDF to the blue color for only that particular weather parameter (before the application of weighting factors). The green color is the CDF of the overall closest weather pattern of historical observed weather data to the future month when all the weather parameters are considered with their weighting factors. For example, in Figure 7.9 (top), to disaggregate the future GCM data for July 2020 (blue), for the global horizontal solar irradiance, July of 2006 (red) is statistically the most similar to the future data, whereas for the average wind velocity, July of 2004 is the most similar data to the future data. However, when all the weather parameters are taken into account with their corresponding weighting factors, July 1963 has the most similar overall weather pattern to the future January (2020). The red and green colors are the results based on the FS method and the WS in steps i and ii. The results are shown in purple for cases where the results of FS and WS are the same, as is the case for average temperature. Following this, the hourly weather data of July 1963 is selected for July of the future year.

This method maintains the relationships among the weather parameters as the future hourly data is selected from actual observed historical data. In some cases, the temperature differs greatly from the historical candidates, which leads the CDF of the temperature of the future month to be far beyond the nearest candidate found by the algorithm; this is the limitation of weather classification mentioned in the literature review section 7.1.3. According to the data, this only occurs for the temperature parameter during, primarily, the winter, spring, and fall seasons. In other words, some of the future months during the aforementioned seasons can be considerably warmer or colder than any historical month, which may be relayed by the phenomenon of climate change. In order to account for the deviation, a regression model is incorporated into the workflow.

A regression model is developed to disaggregate (temporally downscale) the temperature for situations when the temperature of the most similar weather pattern in the past is considerably deviated from the future month temperature. The deviation is the absolute difference between the two CDFs. By observing the CDF of the temperature for multiple months, it is found that the two CDFs were relatively close as long as the deviation was under 0.05; therefore, three levels of thresholds, 0.03, 0.04, and 0.05 are considered as the criteria for replacing the temperature data of the selected historical month with the regression model data. If the selected historical month temperature is deviated from the future model more than the threshold, only the temperature of the selected month is replaced with the regression model data. For example, in Figure 7.9 (bottom) for the month of March 2021, the future model temperature CDF is higher than any historical observed temperature CDF, and the deviation is more than 0.05. Therefore, the temperature values resulted from the classification model are replaced with the data from the regression model. Section 7.3.3.2 describes the Random Forest regression model, which is applied to disaggregate the GCM bias-corrected temperature data in cases where the bias-corrected data is considerably higher or lower than any observed historical temperature.

July 2020, the most similar July weather pattern year: 1963



March 2021, the most similar March weather pattern year: 1984

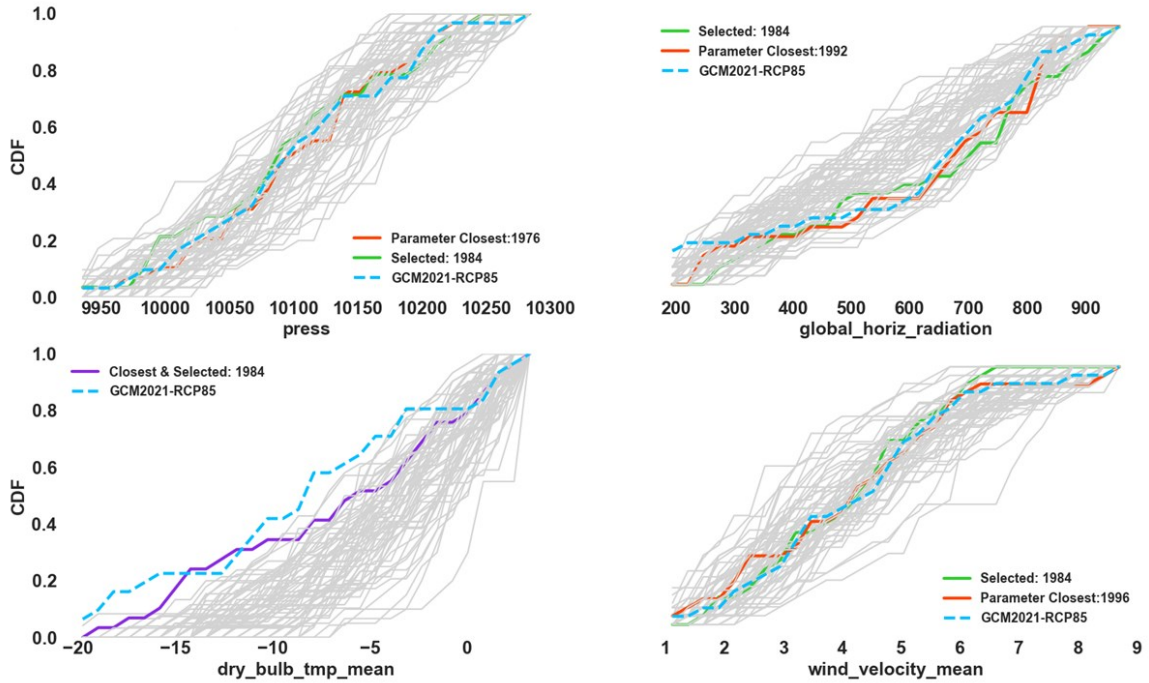


Figure 7.9: CDF of the future weather data in comparison to the historical observed weather data for July 2020 (top) and March 2021 (bottom).

7.3.3.2 Random Forest regression model for prediction (downscaling)

A Random Forest regression model is used due to its capability to model multiple targets (outputs). For disaggregation, the data input into the regression model is daily-average data, whereas the output is composed of hourly data for a 24-hour period. This way, the algorithm learns (train) the relationship between the daily-average input and the 24-hour (hourly) values. The original single-target algorithm of Random Forest is recursively developed based on finding the values of features where the most information is gained from a subset of the original dataset (variance for regression) by a minimization algorithm. The method is described in detail in another study by the authors of this study [33] (chapter 5) and Breiman et al. [43]. For a multi-target problem, Seagal [44], added a covariance weighting to include the multiple outputs to the objective of the optimization in the original Random Forest algorithm. Thus, the multiple outputs and their covariance are considered in the Random Forest algorithm during node splitting. For more details on the multi-target Random Forest algorithm, readers can refer to the study by Segal and Xiao [45].

In order to disaggregate the future GCM or RCM model data, first, all the historical hourly observed temperatures are resampled to daily-average data and introduced to the regression model as inputs of training data. Then, for each daily-averaged observed input, 24 hourly observed values are introduced to the model as training outputs. For each desired month, when the future daily-average GCM data is introduced to the model, 24 hourly values for each day of the desired month are predicted.

All the historical daily average and hourly temperature data for the period of 1953 to 2005 [39] is used to train the model, while data ranging between 2006 and 2018 is used for testing the model (around 20% of training). Figure 7.10 shows the data used to train and test the model, along with the output of the regression model for future data. The green color shows the training data, the light blue color shows the original data for the test period, the dark blue is the output of the Random Forest regression model output, and the red color shows the Random Forest output for the future data. A comparison of the light blue and dark blue hourly resolution data (Figure 7.10 top left) for the test period shows a very good agreement between the observed hourly data and the Random Forest output. For daily-average, monthly-average, and the yearly-average data, the observed and the Random Forest output overlap. The yearly-average data from the Random Forest model clearly shows the increase in average temperature in the future (Figure 7.10 bottom right).

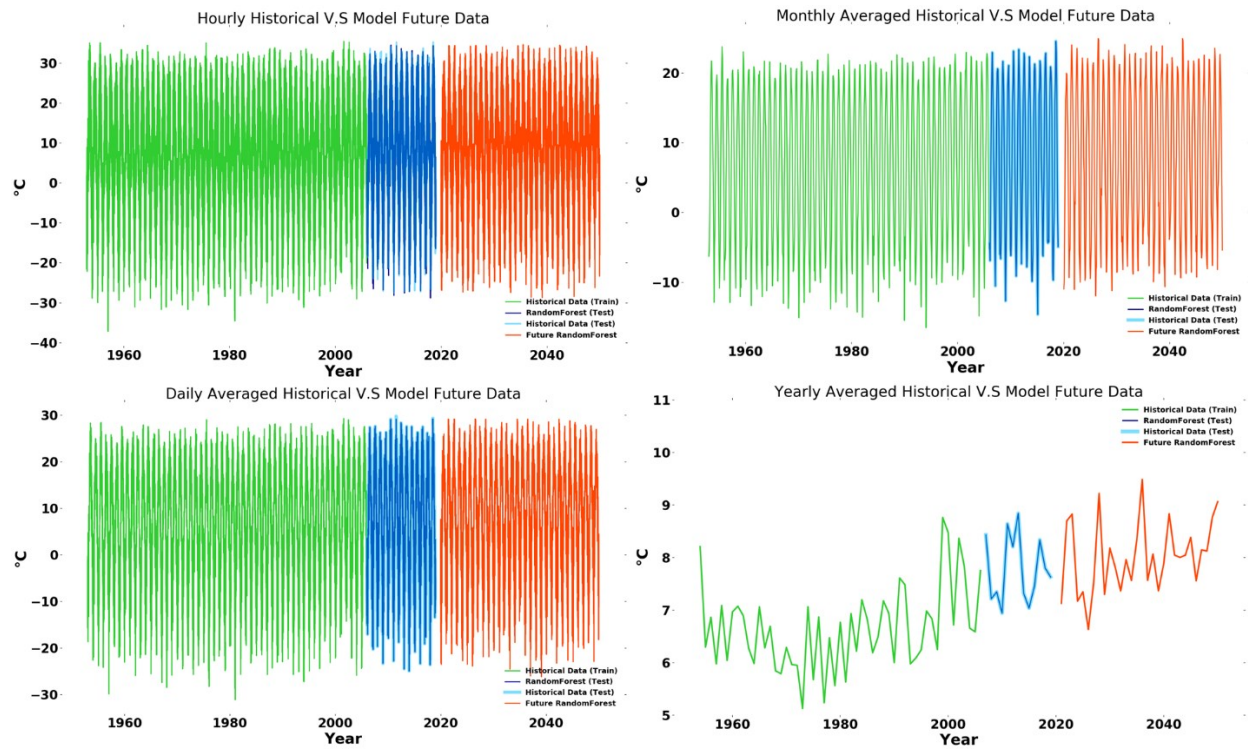


Figure 7.10: Train, test, and predicted result of the Random Forest regression model; hourly data (top left), daily-averaged data (bottom left), monthly-averaged data (top right), and yearly-averaged (bottom right).

For better presentation of the agreement between the observed data and the regression model output, the last 5 months of the training period, and the first 8 months of the test period is shown in Figure 7.11. As mentioned before, a very good agreement between the hourly observed data and the output of the regression model can be seen (Figure 7.11 top). Moreover, the daily-averaged observed data and the regression model completely overlap (Figure 7.11 bottom), which means that for daily and yearly resolution data, the regression model transfers all the statistics of the GCM bias-corrected data to the downscaled data. The accuracy of the regression model is quantified by comparing the output of the regression model with the actual hourly data (dark and light blue in Figure 7.10 top left). Table 7.3 reports the error between actual hourly values and the predicted hourly values using mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE), and R^2 score.

Table 7.3: accuracy of the regression model on test data

MSE	RMSE	MAE	R2
5.9	2.43	1.85	0.96

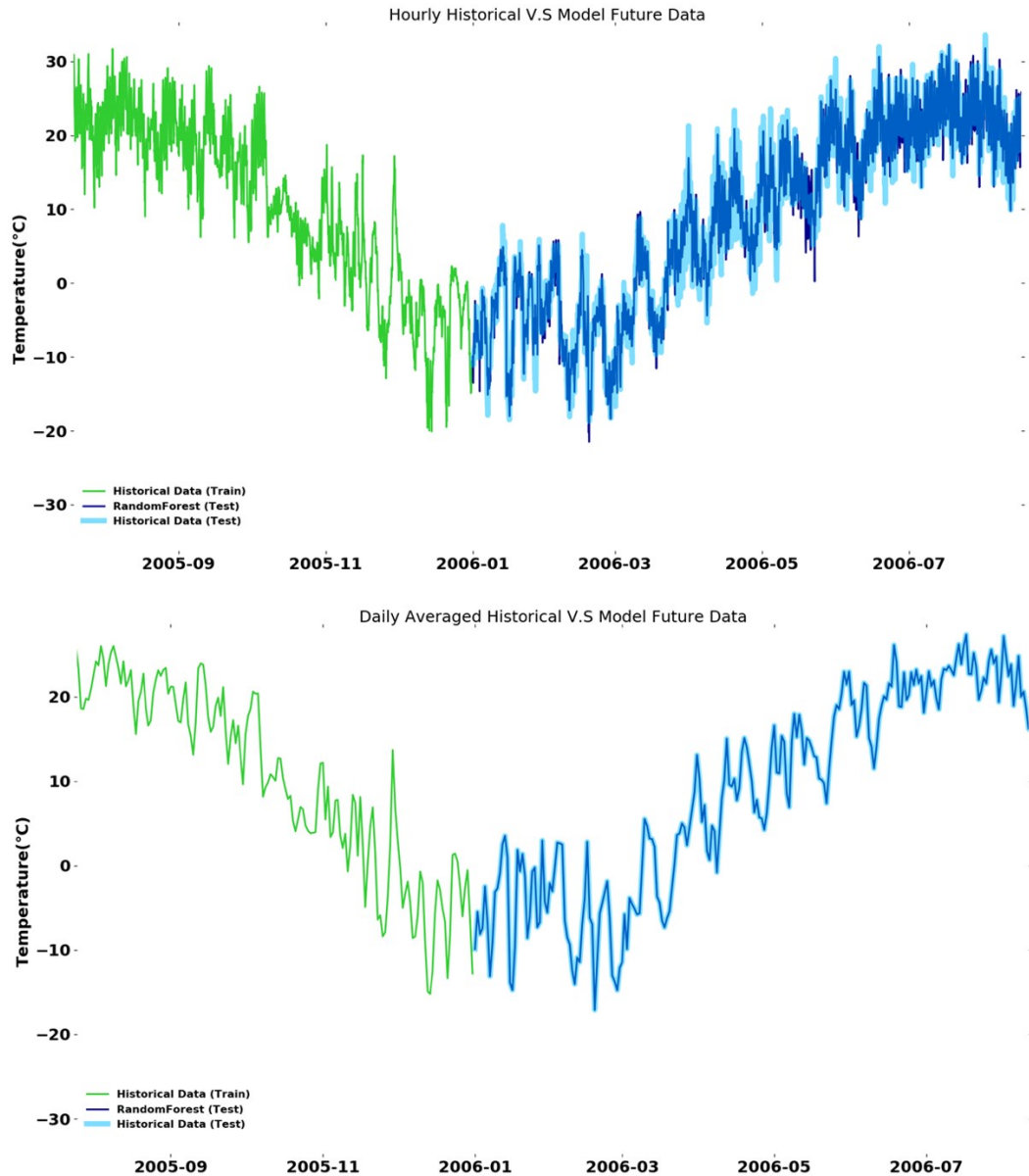


Figure 7.11: The last 5 months of the training period and the first 8 months of the test period with hourly resolution (top) and daily resolution (bottom).

As mentioned above, the regression model is used only when the temperature of the historical month selected by classification differs from the future month temperature with a deviation larger than the defined threshold (0.03 to 0.05). As an example, the data from March 1984 is selected for March 2021 (Figure 7.9 bottom) whereas, the CDF of the temperature of March 1984 differs considerably from the GCM data. Therefore, for the current future month, the temperature data of

March 1984 is replaced with data from the Random Forest regression model. Figure 7.12 shows the temperature values selected by the classification model for March 2021 (March 1984) and the corrected values using the regression model. The red color shows the daily average value of the most similar month of the observed temperature to the future month (March 1984). The light blue color shows the future GCM data, and the dark blue shows the regression model output when it is resampled to daily-average values. It can be seen that the daily data from the GCM and the daily-average data from the regression model overlap.

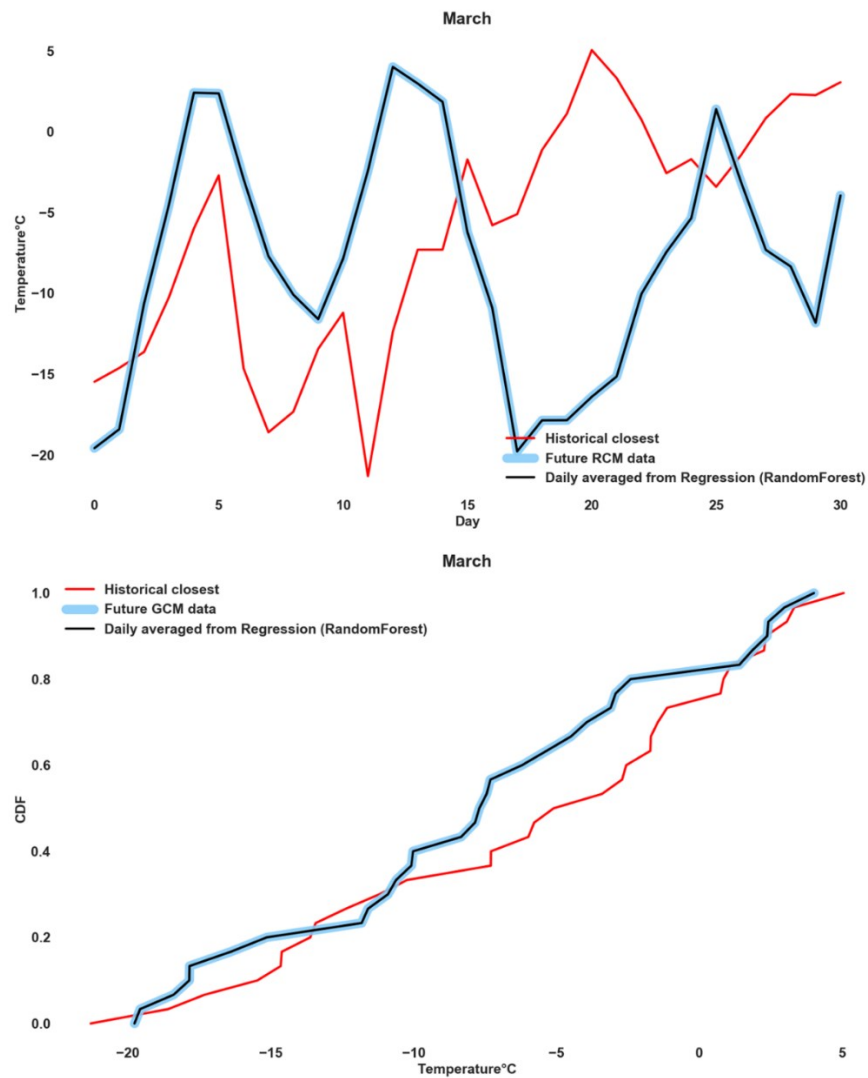


Figure 7.12: Temperature values selected by classification (March 1984) and daily-averaged of corrected values using regression model for March 2021. Top: daily-averaged of all days of the month; bottom: CDF of the daily values.

Figure 7.13 shows the flowchart of the workflow used to process the GCM data to be used for building energy simulation.

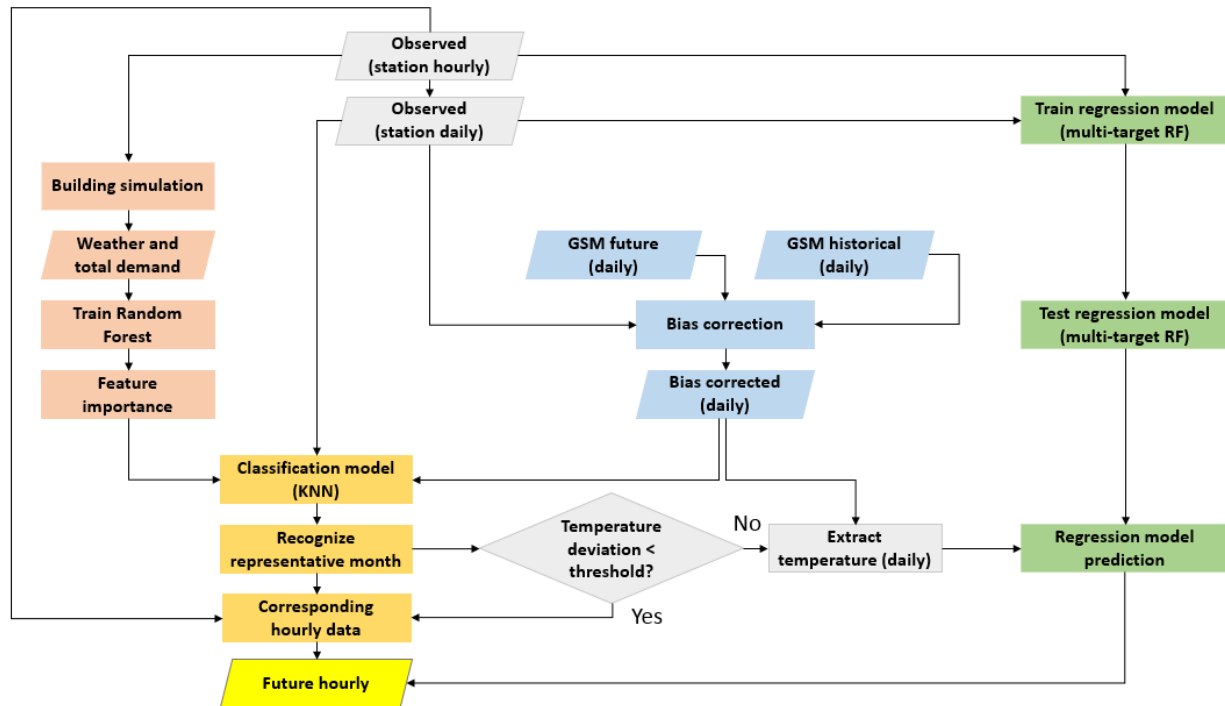


Figure 7.13: Flowchart of the workflow to process the GCM data to be used for building energy simulation; the Random Forest algorithm is used twice; once in the feature importance sub-process (shown in pink) with single target of total demand; and once as a regression model (shown in green) with multi-target of 24 hour-temperature and input of daily-averaged temperature.

As seen in Figure 7.13, the blue blocks in the middle of the flowchart show the bias-correction process. The sub-process is done to remove the bias in the GCM data. The observed daily-average and GCM historical data are used to correct the existing bias in the future GCM data. The output of this sub-process is the bias-corrected GCM daily data for four weather parameters; temperature, solar irradiance, wind speed, and the atmospheric pressure. The output data is then used by the classification model or the temperature by the regression model.

The whole pink-colored blocks on the left side of the flowchart show the feature importance sub-process. The sub-process is done in order to determine the weighting factors for all four weather parameters used in the classification model (equation 2). Multi-years hourly observed data is used

to simulate a small prototypical office building with EnergyPlus building energy simulation. The daily-average weather data and corresponding total energy demand is then used to train the Random Forest (RF) model, and the feature importance of each weather parameter is found. The goal of using RF at this step is to create a good fit to the data in order to find the importance of each weather parameter in predicting the output (daily-average total demand). In other words, in this case, the RF is not used for prediction, and therefore, a training phase is not required.

As stated, an auxiliary regression model is also used to downscale the GCM-daily data for situations when the GCM future bias-corrected temperature is considerably higher or lower than any observed data. For these situations, the consequent temperature CDF of the bias-corrected data is considerably deviated from the CDF of the observed daily-averaged temperature. The green blocks on the right of the flowchart show the train, test, and prediction steps for the RF multi-target regression model. The observed-daily averaged temperatures are used as the input, and the 24 hourly temperature values are used as the targets to train the regression model. In the testing phase, the model is tested with unseen observed values to ensure the model properly predicts the values.

The orange blocks show the classification sub-process. The daily-average observed data, along with the bias-corrected GCM data and feature importance of each weather parameter, are introduced in the KNN algorithm to find the most statistically similar weather pattern to the future data from the observed data. The statistical similarity is based on the comparison of CDFs of the four weather parameters, as explained in section 7.3.3.1. The date of the most similar weather pattern is noted. At this step, the deviation of the CDF of the selected month's temperature from the temperature CDF for the future GCM is examined. If the deviation is less than the specified threshold, all the corresponding hourly weather parameter data of the noted date, including temperature, are directly used as future hourly data. Otherwise, only the hourly temperature of the month found by KNN is replaced with the output of the multi-target Random Forest regression model (input is the bias-corrected temperature of the GCM).

The procedure is done for all the 12 months of the year from 2020 to 2049 for all four climate change scenarios. As a result, for each year of the future, 12 months of hourly historical weather data are generated. The three deviation thresholds defined in section 7.3.3.1 are considered to evaluate the effect of the threshold criteria in the workflow. Overall, 360 weather years representing the future 30 years are made for each of the four scenarios for all three thresholds.

The next section shows the results of the weather year and corresponding building energy simulations.

7.4 Results and Discussion

The workflow described in the previous section used to generate 360 weather years for the future 30 years under four climate change scenarios and three levels of thresholds. As an example, Table 7.3 shows the historical months from which weather data are used to generate the future hourly weather data of 2020 with RCP 2.6, 4.5, 6, and 8.5 W/m² and under the three thresholds of 0.03, 0.04, and 0.05. Depending on the RCP, various months from different years are selected as the corresponding months of the future year. As expected, the regression model (Shown with 'R') is less frequently used when a higher level of threshold is selected. As defined in previous sections, the threshold is the criteria for using the regression model temperature data instead of the temperature resulted from KNN. The deviation of the temperature CDF of the selected month by KNN from the temperature CDF of the GCM future data is compared. If the difference falls within the threshold, the KNN temperature data is kept, else the Random Forest predicted data is used. For example, under RCP of 2.6 the regression model is used only once with a threshold of 0.05 whereas, it is used 8 out of 12 times with a threshold of 0.03.

It should be noted that the regression model in the workflow can be used for the following two conditions; 1) When the future GCM month is warmer or colder than all the observed historical months and 2) when a very small threshold is selected, and the algorithm is interested to predict the data as close as possible to the GCM data. By continuously observing the selected data by KNN (Figure 7.10 as an example), it is realized that a deviation higher than 0.05 happens when the future GCM data is considerably warmer or colder than the warmest or coldest historical observed month. Considering the high weighting factor for the average temperature, it can be inferred that all months represented by 'R' and under the threshold of 0.05 in Table 7.4 are simply the months that are predicted to be warmer or colder than any of the observed historical months in the past by the GCM. However, it must be noted that deviations of 0.04 and 0.03 do not necessarily infer this, as it has been observed that in cases where the deviation is within the range of 0.03 to 0.04, there has been at least one colder or warmer month. In these conditions, the regression model is used only to make the results closer to the future data.

Table 7.4: The historical years from which hourly data are selected month-by-month for the future year of 2020; the months in which temperature is corrected by the values of regression model are shown with ‘R’.

		2020											
RCP	Threshold	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
		1984	1972	1968	1958	1999	1969	1968	2004	2000	1971	1958	1973
2.6	0.05		R										
	0.04	R	R		R		R						
	0.03	R	R	R	R	R	R				R		R
		2014	1975	1989	1956	1976	1976	1954	1964	1976	1957	1989	1975
4.5	0.05						R			R			
	0.04		R		R		R	R		R	R	R	
	0.03	R	R	R	R	R	R	R	R	R	R	R	R
		2008	1975	1968	2006	2004	2003	1963	2007	1974	1970	2002	1973
6.0	0.05												
	0.04	R	R					R	R				R
	0.03	R	R	R		R	R	R	R	R	R	R	R
		1964	1979	1992	1982	1959	2008	1963	2003	1999	1969	1986	1957
8.5	0.05	R		R					R				
	0.04	R	R	R		R			R				R
	0.03	R	R	R	R	R	R	R	R	R	R	R	R

To validate the method, the daily average of the downscaled hourly data is compared with the original bias-corrected GCM data. Figure 7.14 shows the daily average of the created downscaled hourly temperature under RCP 8.5 for all three thresholds for the year 2020, compared to the corresponding bias-corrected GCM temperature. The figure shows that the daily-average of the downscaled data are in very good agreement with the original bias-corrected data (red).

As it is seen from Table 7.3, for RCP 8.5, with a threshold of 0.03, all the temperature values selected by KNN are replaced with the regression model output for all the months of the year 2020. In Figure 7.14, the temperatures for the threshold of 0.03 and the GCM bias-corrected data (blue and red) overlap for all the months of the year. The lower threshold of 0.03 means the regression model is more frequently used to downscale the data. When the regression model is used for a month, the downscaled hourly daily-average data was found to be exactly equal to the mean of the daily GCM bias-corrected data, as observed in Figures 7.10 and 7.11. In fact, any threshold between 0.03 and 0.05 can be used by the user; however, the following tradeoffs must be considered:

- 1) A smaller temperature deviation (threshold closer to 0.03) from the GCM data will sacrifice the actual correlation between temperature and the other parameters.

2) A larger temperature deviation (threshold closer to 0.05) from the GCM data will maintain the actual correlation between the parameters.

In other words, in order for the downscaled temperature to be closer in value to the GCM bias-corrected temperature, the smaller threshold (0.03) sacrifices the correlation between the temperature and the other weather parameters.

For the threshold of 0.05, the regression model is only used for January, March, and August which means that except for these three months, the actual correlations between temperature and all the other weather variables are respected. During these months, the temperature data created with the threshold of 0.03 criteria overlap with the other thresholds and the GCM data. In other words, during these three months, the daily-average of downscaled hourly temperatures are identical to the GCM bias-corrected data. However, the correlations between the temperature and other weather parameters are not necessarily respected. For the threshold of 0.04, during the months of January, February, March, May, August, and December, the regression model is used, and the data overlaps with the 0.03 threshold and the GCM data. For other months, although the data with a threshold of 0.04 does not overlap with the GCM data, there is only a minor deviation between the threshold of 0.04 data and the GCM bias-corrected data.

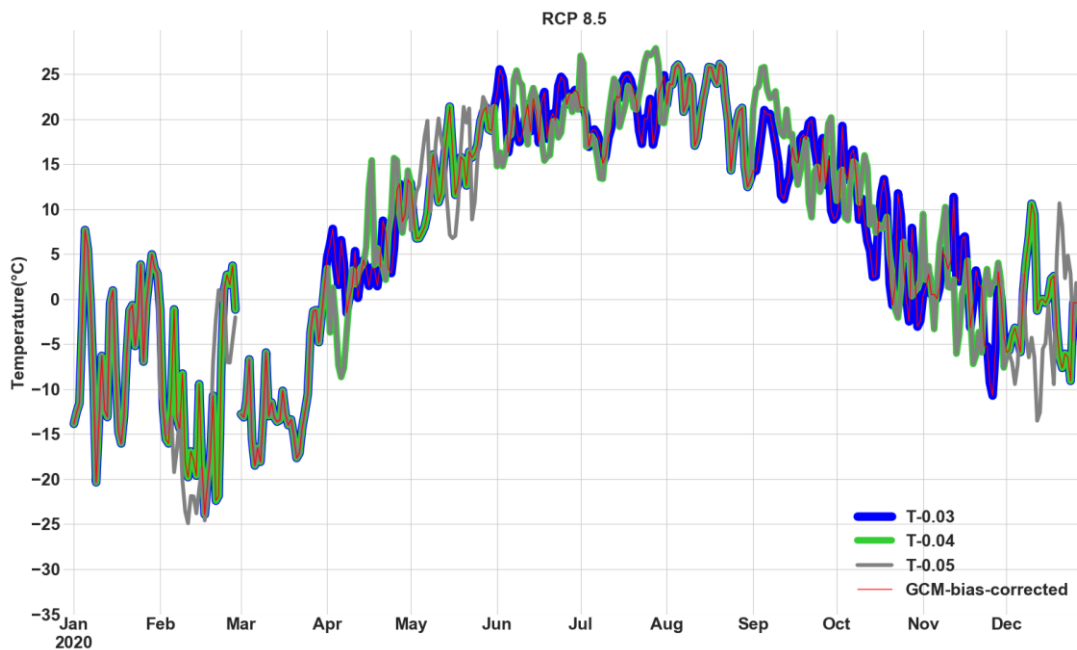


Figure 7.14: The daily-averaged temperature created under RCP 8.5 with three thresholds and the corresponding bias-corrected GCM temperature for the year 2020 (discontinuity at the end of February is due to the leap year).

In addition, the yearly-average temperatures disaggregated by the workflow with three different thresholds are compared with the bias-corrected GCM data under the four RCMs in Figure 7.15. For all the RCPs, the threshold of 0.03 shows a very close result to yearly-average GCM bias-corrected data. As expected, the threshold of 0.05 shows the largest deviation from the GCM bias-corrected data, although the deviation at most is within about 1°C. Moreover, RCP 4.5 and 8.5 show higher year by year change. Consequently, a drastic year by year change is expected in the cooling and heating energy demand of the building.

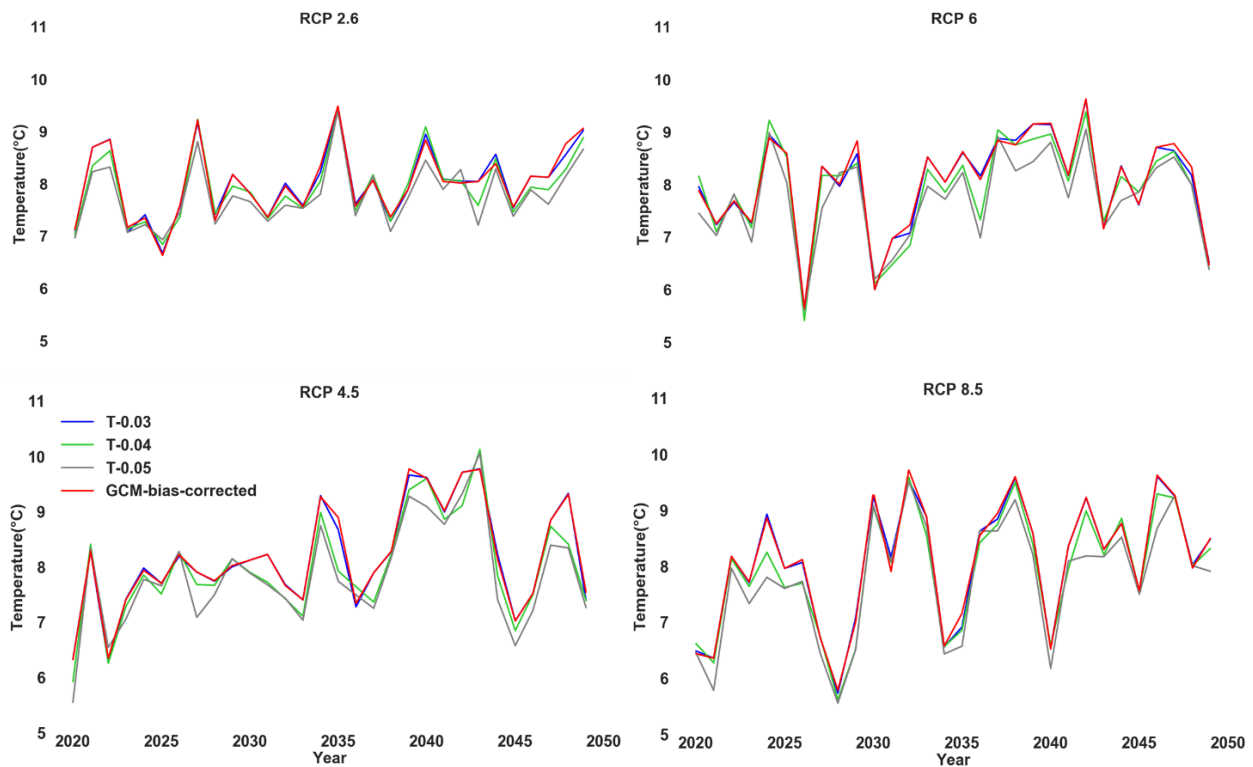


Figure 7.15: Yearly-averaged temperature of GCM bias-corrected data and temperature disaggregated with the three levels of threshold under the four RCPs.

Figure 7.16 presents the annual cooling and heating energy demand of the building simulated with weather files created with the three thresholds and under the four RCPs. For RCPs 6 and 8.5, a minor trend of increase in cooling energy demand can be seen. For RCP 2.6 almost a constant trend can be seen in cooling energy demand whereas, in RCP 4.5, the cooling energy demand trend slightly increases until around the year 2042, where it begins to decrease for the remaining years.

A cooling range of 50 to 70 kWh/m² is noted for most of the time for all the RCPs, although there are a few values passing this range, including about 45 kWh/m² in RCP 4.5 and about 78 kWh/m² in RCP 2.6 and 8.5.

For heating, a steep decrease is noted for all cases. Furthermore, a drastic year by year change in heating energy demand is found under RCP 8.5, which is in line with the drastic change of temperature under RCP 8.5 (Figure 7.15). A wide range of about 83 to 136 kWh/m² in heating energy demand can be seen under RCP 8.5. For RCP 2.6 and 6 the range is reduced to about 90 to 130 kWh/m².

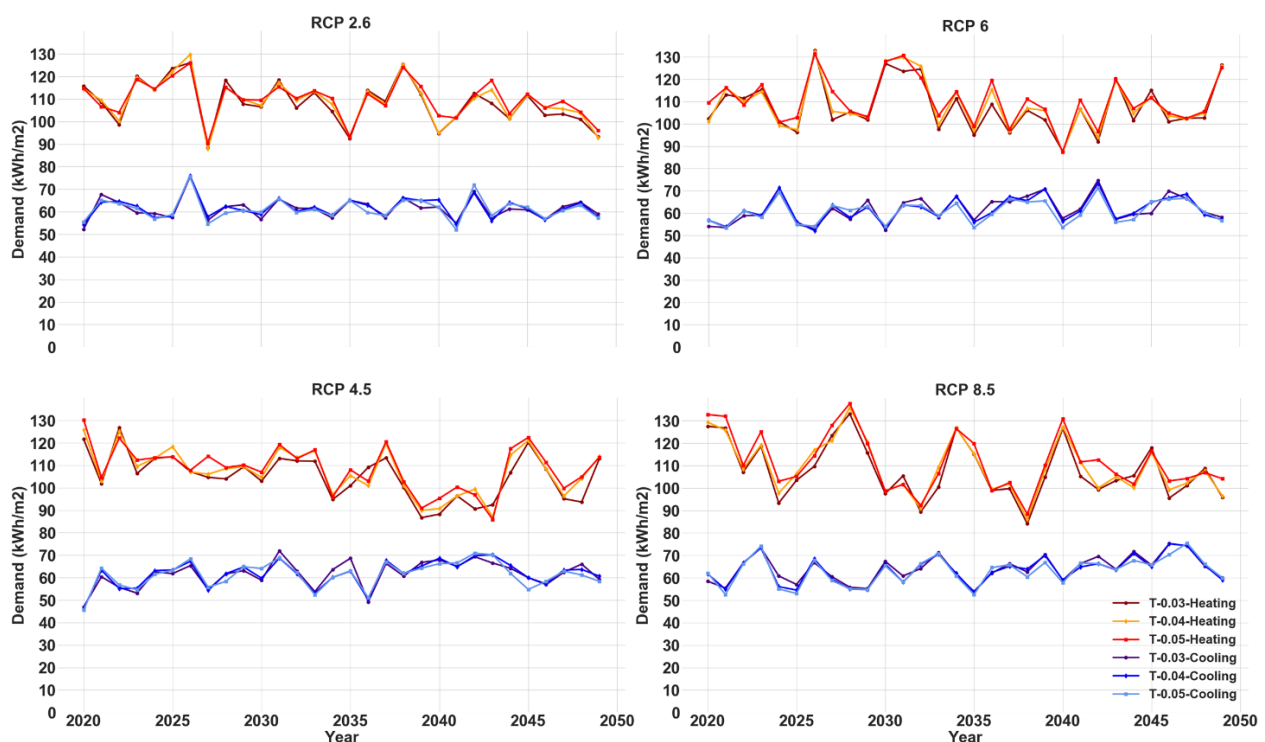


Figure 7.16: Annual cooling and heating energy demand of the building simulated with the weather files generated under the four RCPs and based on three thresholds.

Choosing different thresholds has a minor effect on annual cooling energy demand. The maximum deviation is found to be at most about 12 kWh/m² (or 10%) deviation in the year 2042 under RCP 8.5. This is mainly because of the fact that during the summer, solar irradiance can have a considerable effect on cooling energy demand, whereas the solar irradiance doesn't change under different thresholds. In other words, only temperature in the weather year might change with

different thresholds. On the other hand, for heating, choosing different thresholds leads to slightly higher deviations in annual energy demand. This is due to the fact that during winter, the days are shorter and cloudier, and the sun angle is lower in cold climates. Therefore, less solar radiation is available, and the total heating energy demand is sensitive to outside air temperature change. Hence, changing the threshold and consequent temperature data of the weather year leads to higher heating energy demand deviation. From Figures 7.14 to 7.16, with a threshold of 0.03, the regression model is used for most months of the weather file. Due to the selection of the regression model data, the relationship between the weather parameters may change. In contrast, with a threshold of 0.05, the regression model is barely used, and for most of the months, the temperature remains unchanged. However, the temperature selected from the classification might significantly deviate from the GCM temperature data. A threshold of 0.04 can be assumed as a conservative value, as it shows good results when comparing the corresponding output temperatures to the GCM bias-corrected temperature data. Furthermore, the specified threshold of 0.04 allows maintaining the relationship of the weather parameters, at the hourly resolution, for most cases. This is due to the fact that the real historical temperature from the classification model is kept rather than using the regression model.

In summary, the findings determined from this study are as follows:

- A statistical bias-correction method is used to remove the bias in the GCM data and calibrate it with station data; the station data can be a local weather station data as well.
- A weather classification model is introduced to downscale the GCM data for building energy simulation.
- The classification model keeps the actual correlations among the weather parameters as it uses whole weather parameters actual hourly data from observed data for future values.
- For those cases when the GCM future temperature is considerably higher or lower than any observed historical temperature, a regression model is trained on historical data to learn the relationship between the daily-averaged temperature data and hourly data; therefore the model is used to downscale the temperature and replace the temperature of the classification model with the regression output.

- A threshold is considered in order to evaluate the deviation of the observed temperature from the GCM temperature. In the case that the threshold for the temperature is not met, the regression model data is selected, otherwise, the classification model data is used.
- The thresholds can be selected by the user, however the following tradeoffs must be considered: 1) A smaller temperature deviation (closer to 0.03) from GCM data will sacrifice the actual correlation between temperatures and the other weather parameters and 2) A larger temperature deviation from the GCM data will maintain the actual correlation between the temperature and the other weather parameters.
- The proposed method keeps the statistics of the GCM data when data is downscaled which means that if the downscaled hourly data is averaged again, the statistics of the average data and the bias-corrected GCM data would be the same

The multiple year by year generated weather files can be used for simulation by architects, energy modelers, and engineers to evaluate and optimize their designs facing natural year by year weather variation and extreme weather conditions under climate change.

7.5 Conclusion

GCM data has been widely used as a means to assess the impact of climate change in various fields. In the case of building energy performance, the use of GCM data may allow to better estimate future building energy performance. However, GCM data is biased, where a considerable deviation can be found between historical GCM data and the observed weather station data, and do not have the hourly resolution required for building energy simulation. In order for GCM data to be used as a means to estimate future building energy performance, further processing is required.

The current study applied the quantile-quantile method, which statistically removes the bias in the GCM future data by comparing the historical GCM data with observed station weather data. The method can be applied to all weather parameters. Once the bias was corrected, the study applied a hybrid classification-regression model to downscale the bias-corrected GCM data. The method uses similarly observed weather data to represent future data. This method allows for the correlation among the weather parameters to be preserved with an hourly resolution. However, this K-nearest neighborhood machine-learned classification model cannot provide data for cases

that are outside of the observed data range. In other words, it may not be able to solve for cases with higher temperature intensity resulting from climate change. Therefore, the study applied a regression model for cases where the future GCM data is considerably warmer or colder than any observed weather data. In such situations, the temperature values from the classification model is replaced with hourly temperature data from the Random Forest regression model to follow the temperature rise due to climate change effect.

The suggested workflow allows generation of future weather files for each future year for different climate change scenarios. The workflow preserves the extreme weather characteristics and is suitable for reliability analysis.

7.6 Remark

The study introduces the weather classification method as a technique to downscale the GCM future data to hourly resolution data for building simulation. The current study uses a single GCM model (GFDL-ESM2M) containing all the four climate change scenarios RCPs. Two assumptions are made in this study:

1. the GCM data is stationary which means while the climate changes, the statistical relations among the meteorological parameters remain constant over time;
2. the GCM correctly model the future climate change scenarios after the bias-correction is conducted although it not guaranteed that this assumption will be actually true.

The goal of the study was to show the new method of bias-correction downscaling in building simulation. For, more reliable results, it is suggested to use the data based on the average of multiple GCM models. Furthermore, by following the proposed workflow, future weather files can be generated for a specific application, in this study, for building energy simulation. Other applications such as performance analysis of energy generation might require another set of future weather files generated for that purpose.

7.7 References

[1] Climate data and scenarios for Canada: Synthesis of recent observations and modelling results, Environment and Climate Change Canada, ISBN: 978-0-660-04262-6

- [2] IPCC, 2013. Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge, UK; and New York, USA: Cambridge University Press
- [3] G. Lebel, M. Dubé, R. Bustinza, “Surveillance des impacts des vagues de chaleur extreme sur la santé au Québec à l’été 2018, INSPQ”, (2019), Online: <https://www.inspq.qc.ca/bise/surveillance-des-impacts-des-vagues-de-chaleur-extreme-sur-la-sante-au-quebec-l-ete-2018#ref>
- [4] SANTÉ MONTRÉAL, retrieved online on May, 2019:
<https://santemontreal.qc.ca/en/public/news/news/extreme-heat-66-deaths-in-montreal-in-2018/>
- [5] D. Bélanger, P. Gosselin, B. Abdous, P. Valoir, “La climatisation à domicile dans les secteurs les plus défavorisés des grands centres urbains du Québec, INSPQ”, (2016), Online:
<https://www.inspq.qc.ca/bise/resume-la-climatisation-domicile-dans-les-secteurs-les-plus-defavorises-des-grands-centres-urbains-du-quebec>
- [6] M. Herrera, S. Natarajan, D.A. Coley, T. Kershaw, A.P. Ramallo-Gonzalez, M. Eames, D. Fosas, M. Wood, “A review of current and future weather data for building simulation”, Building Service Engineering Research & Technology, Vol. 38(5) 602–627, 2017
- [7] R. L. Wilby, J. Troni, Y. Biot, L. Tedd, B.C. Hewitson, D.M. Smith, R.T. Sutton, “A review of climate risk information for adaptation and development planning.” International Journal of Climatology 29, 1193-1215, 2009
- [8] H. Kikumoto, R. Ookaa, Y. Arima, T. Yamanaka, “Study on the future weather data considering the global and local climate change for building energy simulation”, Sustainable Cities and Society 14, 404–413, 2014
- [9] T. Berger, C. Amann, H. Formayer, A. Korjenic, B. Pospischal, C. Neururer, R. Smutny, “Impacts of climate change upon cooling and heating energy demand of office buildings in Vienna, Austria”, Energy and Buildings 80, 517–530, 2014
- [10] E. Zorita, H. von Storch, “The analog method as a simple statistical downscaling technique: comparison with more complicated methods”. Journal of Climate 12(8), 2474-2489, 1999
- [11] S. Trzaska, E. Schnarr, “A Review of Downscaling Methods for Climate Change Projections”, U.S. Agency for International Development (USAID), September 2014
- [12] S.E. Belcher, J.N. Hacker, D.S. Powell, “Constructing design weather data for future climates”, Building Services Engineering Research and Technology, 26, 1, 49-61, 2005

- [13] P. Shen, “Impacts of climate change on U.S. building energy use by using downscaled hourly future weather data”, *Energy and Buildings* 134, 61–70, 2017
- [14] J. Huang, K.R. Gurney, “The variation of climate change impact on building energy consumption to building type and spatiotemporal scale”, *Energy* 111, 137-153, 2016
- [15] A. Jiang, Y. Zhu, A. Elsafty, M. Tumeo, “Effects of Global Climate Change on Building Energy Consumption and Its Implications in Florida”, *International Journal of Construction Education and Research*, 14(1) 22-45, 2018
- [16] L. Wang, X. Liu, H. Brown, “Prediction of the impacts of climate change on energy consumption for a medium-size office building with two climate models”, *Energy and Buildings* 157, 218–226, 2017
- [17] A. Robert, M. Kummert, “Designing net-zero energy buildings for the future climate, not for the past”, *Building and Environment* 55, 150-158, 2012
- [18] M. Hosseini, F. Tardy, B. Lee, “Cooling and heating energy performance of a building with a variety of roof designs; the effects of future weather data in a cold climate”, *Journal of Building Engineering*, 17, 107-114, 2018
- [19] U. Berardi, P. Jafarpur, “Assessing the impact of climate change on building heating and cooling energy demand in Canada”, *Renewable and Sustainable Energy Reviews*, 121, 109681, 2020
- [20] A. Moazami, V. M. Nik, S. Carlucci, S. Geving, “Impacts of future weather data typology on building energy performance –Investigating long-term patterns of climate change and extreme weather conditions”, *Applied Energy* 238, 696–720, 2019
- [21] U. Yao, A. Tettey, A. Doodoo, L. Gustavsson, “Energy use implications of different design strategies for multi-storey residential buildings under future climates”, *Energy* 138, 846-860, 2017
- [22] V. M. Nik, “Making energy simulation easier for future climate – Synthesizing typical and extreme weather data sets out of regional climate models (RCMs)”, *Applied Energy* 177, 204–226, 2016
- [23] J.M. Rey-Hernández , C. Yousif , D. Gatt , E. Velasco-Gómez , J. San José-Alonso, F.J. Rey-Martínez, “Modelling the long-term effect of climate change on a zero energy and carbon dioxide building through energy efficiency and renewable”, *Energy & Buildings* 174, 85–96, 2018
- [24] M. Cellura, F. Guarino, S. Longo, G. Tumminia, “Climate change and the building sector: Modelling and energy implications to an office building in southern Europe”, *Energy for Sustainable Development* 45, 46–65, 2018

- [25] M.F. Jentsch, A.S. Bahaj, P.A.B. James, “Climate change future proofing of buildings—Generation and assessment of building simulation weather files”, *Energy and Buildings* 40, 2148–2168, 2008
- [26] W. Tian, P. de Wilde, “Thermal building simulation using the UKCP09 probabilistic climate projections”, *Journal of Building Performance Simulation* 4(2), 105-124, 2011
- [27] BB. R. Dickinson, “Generating future weather files for resilience”, In: Pablo La Roche MS, editor. *International Conference on Passive and Low Energy Architecture, PLEA 2016 - Cities, Buildings, People: Towards Regenerative Environments*. Los Angeles 2016.
- [28] DF. L. Troup, “Morphing Climate Data to Simulate Building Energy Consumption”, *The ASHRAE and IBPSA-USA SimBuild 2016: Building Performance Modeling Conference*. Salt Lake City, Utah 2016.
- [29] MF. Jentsch, PAB. James, L. Bourikas, AS. Bahaj, “Transforming existing weather data for worldwide locations to enable energy and building performance simulation under future climates”, *Renewable Energy*, 55:514-24, 2013
- [30] B.D. Lee, Y. Sun, H. Hu, G. Augenbroe, C.J.J. Paredis, “A Framework for generating stochastic meteorological years for risk-conscious design of buildings”, *National Conference of IBPSA-USA*, August 2012, Madison
- [31] A.H.C. Van Passan, Q.X. Luo, “Weather data generator to study climate change on buildings”, *Building Services Engineering Research and Technology*, 23, 4, 251–258, 2002
- [32] M. Semenov, LARS-WG stochastic weather generator. United Kingdom, Department of Computational and Systems Biology (2012), available at:
<http://resources.rothamsted.ac.uk/sites/default/files/groups/mas-models/download/LARS-WG-Manual.pdf>
- [33] M. Hosseini, A. Bigtashi, B. Lee, “A systematic approach in constructing typical meteorological year weather file using machine learning”, *Energy and Building*, Volume 226, 110375, 2020
- [34] W.A.R. Brinkmann, “Modification of a correlation-based circulation pattern classification to reduce within-type variability of temperature and precipitation”. *International Journal of Climatology*, 20, 839–852, 2000
- [35] H. J. Fowler, S. Blenkinsop, and C. Tebaldi, “Linking climate change modelling to impacts studies: Recent advances in downscaling techniques for hydrological modelling”, *International Journal of Climatology*, 27, 1547–1578, 2007

- [36] J.P. Dunne, J.G. John, S. Shevliakova, R.J. Stouffer, J.P. Krasting, S.L. Malyshev, P.C.D. Milly, L.T. Sentman, A.J. Adcroft, W. Cooke, K.A. Dunne, S.M. Griffies, R.W. Hallberg, M.J. Harrison, H. Levy, A.T. Wittenberg, P.J. Phillips, N. Zadeh., “GFDL’s ESM2 Global Coupled Climate–Carbon Earth System Models. Part II: Carbon”, *Journal of Climate*, 26, 7, 2247-2267, 2013
- [37] I.J. Hall, R. Prairie, H. Anderson, E. Boes, “Generation of a typical meteorological year. Analysis for solar heating and cooling”. San Diego, CA, USA: Sandia Labs., Albuquerque, NM (USA); 1978.
- [38] S. Wilcox and W. Marion, “User’s manual for TMY3 data sets,” National Renewable Energy Laboratory Golden, 2008
- [39] Environment Canada, Canadian Weather Energy and Engineering Datasets (CWEEDS), available at: https://climate.weather.gc.ca/prods_servs/engineering_e.html
- [40] A. Amengual, V. Homar, R. Romero, S. Alonso, C. Ramis, “A Statistical Adjustment of Regional Climate Model Outputs to Local Scales: Application to Platja de Palma, Spain” *Journal of Climate*, 25, 939-957, 2012
- [41] Y. Lin, Y. Jeon, “Random Forests and Adaptive Nearest Neighbors”, *Journal of the American Statistical Association*, Vol. 101, No. 474, 578-590, 2006
- [42] U.S. Department of Energy, Building Energy Codes Program, Commercial Prototype Building Models, available at: https://www.energycodes.gov/development/commercial/prototype_models
- [43] L. Breiman, “Random Forests”, *Journal of Machine Learning*, Volume 45 Issue 1, 5-32, 2001
- [44] M.R. Segal, “Tree-structured methods for longitudinal data”, *Journal of the American Statistical Association*, 87, 418, 407-418, 1992
- [45] M.R. Segal, Y. Xiao, “Multivariate random forests”, *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 1, 1, (2011), 80-87, DOI: 10.1002/widm.12

Chapter 8. DESIGN SOLUTIONS UNDER WEATHER VARIATION AND CLIMATE CHANGE

Building design solutions under climate change suggested by previous studies were highly related to the method of climate change assessment. In most studies, a single representative year was used to show a couple of decades' climatic conditions. Therefore, finding the design solutions methodology was limited to optimizations without considering the fluctuations and variations resulting from the weather's natural variation. This chapter aims to introduce a novel method to select the design variable that meets the requirements with a high probability of success under different climate change scenarios. In other words, at the design step of building, those designs meet the energy efficiency levels most of the time, excel the designs with only a few times meeting the targets, and are better candidates for the final design selection. The workflow helps architects and designers find the best design solutions under uncertain conditions of climate change at the buildings' design stage. The workflow applies the constructed future climate change weather files under different scenarios (from chapter 7) and predicts multiple design options' energy performance under the future climate. The workflow applies a novel method to find the design solutions that reduce the building's cooling and heating demand and the building keeps the performance under different climate change scenarios.

8.1 Introduction:

In the recent years, considerable researches focused on climate change in building sector. Many researches focused on assessment of climate change impact on existing buildings energy consumption including low energy buildings [1,2]. This included numerous studies for different building types and climates which are helpful to demonstrate the vulnerability of building sector to climate change. Fonseca et al. [3] quantified the effect of climate change on building energy consumption across 96 cities in the United States and found a rise in the energy use intensity of buildings for most cities and climate zones. They suggested prioritization for building energy efficiency of warm and humid climates. For Toronto, a cold Canadian climate, Berardi and Jafarpour [4] studied the effect of climate change on 16 prototype buildings and found a decrease of heating EUI by 18–33% and an increase of cooling EUI by 15%–126% by 2070s. Further studies considered other aspects resulted from climate change. Dino and Akgül [5] studied the effect of

climate change on existing residential building stock in Turkey in case of energy use, greenhouse gas emissions and occupant comfort; they concluded that cooling was dominate thermal loads due to increased temperatures and it was even worse in warmer climates. Their analysis showed severe overheating for poorly ventilated and highly insulated buildings. Other studies considered the effect of climate change on performance of buildings utilized renewable energies; in many studies the modelled present zero-energy building in different climates would not achieve net zero target in future under climate change [6-8].

Other studies focused on strategies toward the mitigation of the climate change impact on buildings using various techniques at design stage or retrofiting. Shen et al. [9] showed that considering the effect of climate change can significantly change the strategies of selecting energy efficiency measures during the buildings retrofit as a result of increasing cooling however, only one climate change scenario was considered in the study.

At the early design stage of a net-zero energy building, Chai et al. [10] optimized the sizing of a net-zero energy building air-conditioning system, the PV system and the electrical storage system under climate change without optimization of building enclosure. Gercek and Arsan [11] conducted sensitivity analysis for a residential building design parameters in Turkey under climate change; their analysis for a warm and humid climate indicated that the most important design parameters related to energy and environment were those with the transparent surfaces of the building enclosure, i.e. solar heat gain coefficients (SHGC), and heat transfer coefficients (U-value) of transparent surfaces of the building.

The effect of architectural design parameters in mitigating the effect of climate change is also investigated in the literature. Roshan et al. [12] downscaled two climate variables namely, temperature and relative humidity of a general circulation model (GCM) under three scenarios for 10 cities representing 10 climatic condition in Iran. Having the two climatic condition data, the authors investigated various design strategies suggested by Givani [13] bioclimatic chart in the past and the future time horizon for the 10 climatic conditions. The design strategies were 16 design strategies suggested by Givani [13] including active heating/cooling, passive cooling, shading, etc that could be used by architects to design the buildings according to climatic condition. Their study showed that compared to the past, the design strategies changed when the future

climate were considered in some climates. However, the authors did not consider the solar radiation effect in their investigation.

Rubio-Bellido et al. [14] optimized annual energy demand of an office building in 9 Chilean climate zones under an extreme climate change scenario; the authors optimized the energy demand using the parameters of form ratio (FR) and window-to-wall ratio (WWR) and concluded that for different climates, different climate change scenarios, and different future time period, the optimum value of FR and WWR can vary.

The design solutions suggested by the literature were highly related to the method of climate change assessment. In all of the literature, a single representative year was used to show the climatic condition of couple of decades. Most of the literature used morphing method that apply the super imposed effect of the climate change on a single typical meteorological year [15-22]; consequently, the methodology of design solutions were limited to optimizations without considering the fluctuations and variations resulted from the natural variation of weather.

8.1.1 System Engineering designs to mitigate climate change

8.1.1.1 Robust design

Although energy consumption reduction might be a major objective in the design of a sustainable energy building, it may not be the only concern in design. Generally, the uncertain parameters such as occupancy patterns and weather condition variation can lead to a significant variation in energy performance. Even multi-year energy simulations with historical weather data show considerable fluctuation in a period of time. In these situations, the simulations with some weather years might lead to meeting the energy efficiency target whereas they may not meet the target when the simulations are conducted with other weather years. This is where an engineering design can be significantly useful to cope with such relatively complicated problems.

Robust design optimization has been one way of dealing with selecting designs under uncertain parameters due to occupancy-related or weather condition [23-27]. The concept of robustness in design was associated with selecting the designs that show the least sensitivity to uncertain condition.

In Taguchi method, the sensitivity of uncertain condition is defined by signal-to-noise ratio (S/N) with unit of decibel (db) expressed in the following [28]:

$$S/N = 10 \log_{10} \left(\frac{\mu^2}{\sigma^2} \right) \quad (1)$$

Where: μ is the signal factor and σ is noise factor

In order to address robust design solution against climate change, a previous study [29] suggested robust design optimization for finding design solutions. In the study, the climate change was considered as a noise to the system therefore, used two objective function for a genetic optimization process; the first objective was to optimize signal-to-noise ratio (S/N), and the second objective function was to minimize the energy use. The study used 3 weather files, a constructed typical (TDY), an extreme cold year (ECY), and an extreme warm (EWY) future years as input.

The S/N was adopted to the problem by defining first the variability in response expressed by mean squared deviation (MSD) which is average squared difference of energy simulated with ECY and EWY from TDY as reference values. The second parameter of S/N was defined by energy performance of the design with TDY. The whole objective function was formulated with following equation:

$$f(x, u_i) = -10 \log_{10} \left[\frac{(\sum_{i=1}^p E_{TDY}(u_i, i))^2}{MSD} \right] \quad (2)$$

Where: p is temporal resolution (12 for monthly, 8760 for hourly), $E_{TDY}(u_i)$ is the energy performance of each design with TDY, and MSD is the average squared difference of energy simulated with ECY and EWY from TDY.

By minimizing the two objective functions, the authors found the robust design solutions against the climate change.

Although optimization makes the process faster, it doesn't provide observing all the design options over long-term. In engineering problems, designers might interested in more output characteristics; for example, designers may want to achieve more chance of success or meeting a design goal that can be achieved through axiomatic design concept.

8.1.1.2 Axiomatic Design

The axiomatic design focuses on mapping between design parameters and functional requirements. In axiomatic design, the functional requirement can be met through a systematic formulation of design parameters [30]:

$$\{FR\} = [A] \cdot \{DP\} \quad (3)$$

Where: FR is matrix of requirement, A is called design matrix and DP is matrix of design parameters.

The name of axiomatic design comes from design principle or axiom in which two axioms should be respected:

1: The Independence Axiom:

Maintaining the independence of the functional requirements (FRs).

2: The Information Axiom:

Minimize the information content of the design or maximize the probability of success [30].

According to the axiomatic design, good designs are those respecting the two axioms. The axiomatic designs were previously used in many industries [31-33] including HVAC design [34]. This study introduces the axiomatic design concept in building energy performance under uncertain condition of climate change. In green building industry, the functional requirement can be total energy consumption intensity, cooling and heating demand, Greenhouse gas emissions, etc. The design parameters are building envelope, HVAC, and architectural design parameters. The probability of success also can be considered as probability of a design to meet a specific target or functional requirement under uncertain condition.

According to the independent axiom, the good design is the design in which changing the value of a design parameter, does not affect more than one functional requirement; this axiom in practice, requires to have design parameters equal or higher than the numbers of functional requirements, otherwise the design would be called a coupled design which is not acceptable by the theory. If the number of design parameters is more than the number of functional requirement, the system is either coupled or redundant; in a redundant system, by aggregating the design parameters and tuning the values, a decoupled design can be achieved which is considered a good design by axiomatic design [30]. For the case of building cooling and heating demand efficiency, many design efficiency measures contribute to either cooling or heating efficiency whereas at the same time it contributes to a penalty in the other if used alone. This means that changing value of a design parameter can change both functional requirement. For example, using a single efficiency

measure of reflective color for walls reduces the cooling energy demand in summer, whereas, it increases the heating energy demand in winter; these savings and penalties varies in different climates and building characteristics. However, if multiple design parameters are applied, by aggregating the effects of multiple design parameters a good design fulfilling the functional requirement can be achieved.

Moreover, in many guidelines and standards like those that are based on performance, the cooling and heating demand must be less than an absolute target [35] which can be considered as functional requirement in axiomatic design concept. However, what is less considered (if any), is the probability of the success or probability to meet the targets. At the design stage of a sustainable and energy-efficient building, the decision about the design should be made under uncertain condition related to the project. There can be a considerable uncertainty in building energy consumption corresponding to variable weather condition.

The axiomatic design is identified through quantification of probability of success; this means the probability that a design meets the functional requirements. Considering the policies and design goal, the axiomatic design tackles the uncertainty introduced by weather variation and climate change by finding the designs that meet the functional requirement with the most probability over the expected lifespan of the building. In order to find the best values for design parameters, the simulation tools can be used. With the novel machine learning algorithms, the consideration of all the design options under all the future climate change weather years doesn't necessarily require full simulation runs; therefore surrogate models trained on significantly lower number of simulations can be used to cover all design space under all the future weather years.

8.1.2 Objective and organization of the study

The goal of this study is to select the design variable that meet the requirements with high probability of success under different climate change scenarios. This means that at the design step of building, those designs meeting the energy efficiency levels for most of the time excel the designs with only a few times meeting the targets and are better candidate for final design. The workflow help architects and designer to find the best design solutions under uncertain condition of climate change at design stage of the buildings. In addition, policy makers can use the workflow to observe the energy performance of multiple building characteristics under climate change in

order to make appropriate national, provincial, and municipality codes both for building envelopes and energy performance of different buildings in different climate zones.

The rest of the study is organized into different sections:

In section 8.2.1 the simulation results of a baseline model under historical and future weather years are shown to explore the trend of cooling and heating over a long period of time. In section 8.2.2 the design of experiment as a method to create different design alternatives is explained. A deep neural network will be applied to reduce the number of required simulation runs. Therefore, preprocessing of input and simulation output is conducted in section 8.2.3 to prepare data for training and testing of the model. The model architecture, validation and test, together with post processing are explained in section 8.2.4. A performance indicator for different design alternatives is introduced in section 8.2.5; this is followed by a summary of the whole methodology used in this study in section 8.2.6. The results are shown in section 8.3 followed by discussion in section 8.4. Finally, the contribution of the study is summarized in section 8.5.

8.2 Methodology

8.2.1 Observing energy performance of a baseline over time

The historical and the future weather data under the four scenarios constructed in the previous chapter were used to simulate the case study building with NECB 2015 standard as a baseline model. The simulations are conducted for the historical period of 1953-2014 and the future years of 2020-2049. In addition, the simulation is conducted with Canadian weather year for energy calculation (CWEC) to compare the results with historical and future weathers. Figure 8.1 shows the cooling and heating demand of the building under historical, future, and CWEC weather years. From the figure, since the CWEC is a representative of the years 1960-1989, the cooling and heating demand result of CWEC (black line) show an acceptable agreement with the cooling and heating results of the actual years for the same period. However, after the 1989, the cooling shows an increasing trend whereas the heating shows an increasing trend.

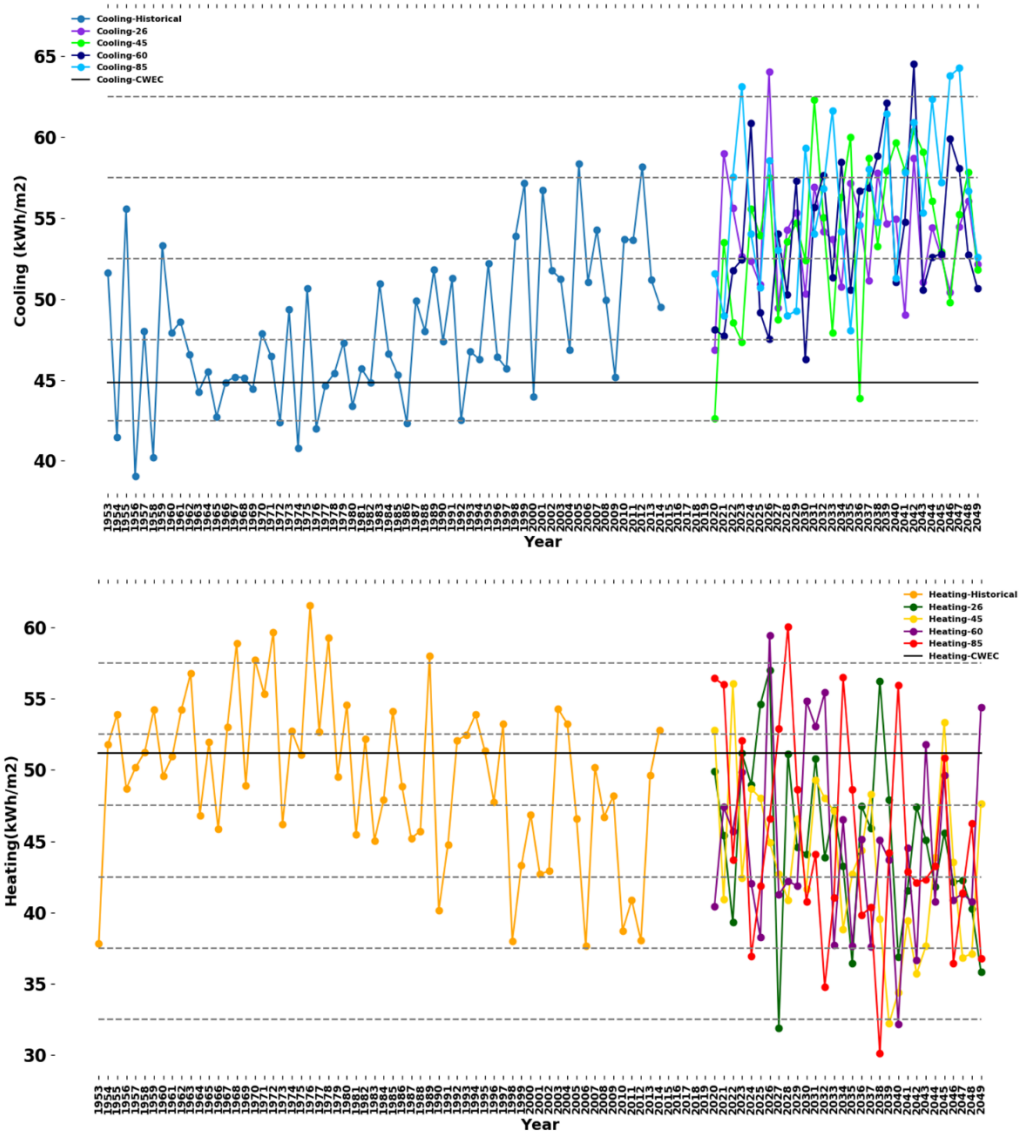


Figure 8.1: Annual cooling and heating of the small office building under historical, CWEC, and future weather years.

The result of simulation with CWEC shows a cooling demand of about 45 kWh/m²; the results of the actual years between 1960 and 1989 show cooling demand within the range of 41-51 kWh/m². The range of cooling demand increases to 43-58 kWh/m² after 1989. The simulation results with the future weather years show that except two years under RCP 45, all the cooling demands are higher than 45 kWh/m² namely, between 46 and 65 kWh/m².

For the heating, the simulation with CWEC shows heating demand of 51 kWh/m². For the period of 1960-1989 the simulations with actual weather years show heating demand within the range of 46-58 kWh/m². The heating demand gradually decreases to the range of 37-57 kWh/m² between 1990 and 2014. The range of heating demand reduces further to 30-60 kWh/m² under future weather years between 2020 and 2049.

The above were results of the baseline model under multiple historical and future years. However, some design combinations might present smaller energy consumptions with drastic fluctuations whereas, others might show larger energy consumption with a slight variation over time. Therefore, in order to investigate the performance of different design combinations, a large-scale building energy simulations including multiple design options are conducted that are explained in the following section.

8.2.2 Design of Experiments

In complex engineering problems where scientific theory is difficult to achieve if available, experimentation and observation is the only approach to find a solution. When several factors play a significant role in the performance of a system, the best strategy of experimentation is to design a type of factorial experiment. In a factorial experiment, factors are varied together to study the joint effects of several factors effect [36]. In fact, in each complete trial, all possible combinations of levels of the factors are investigated [37]. The experimental design has numerous applications including, product design configuration, process development, and optimization, performance testing, evaluation of materials and alternatives, component tolerance determination, reliability and life testing [36].

Building energy performance is significantly dependent on a range of architectural efficiency measures and their interactions. Therefore, evaluation of simultaneous effects would be quite difficult to predict unless a large-scale full factorial simulation is carried out [38]. In this study, a factorial energy simulation are conducted to fully explore the energy performance resulted from design space. In order to have an efficient way of investigation, the design factors are limited to architectural and enclosure characteristics which contribute a considerable performance change. Based on the author experience in numerous simulations and the literatures, wall and roof thermal insulation, wall and roof solar reflectance, window-to-wall ratio, and window type are among the most influential and practical in building envelope design. These parameters could be also found

using sensitivity analysis techniques [38]. Table 8.1 shows the design factors and the levels for each factor.

Table 8.1: Design parameters and levels selected for the case study

	Roof Solar Reflectance	Wall Solar Reflectance	WWR	Window Type (U Factor in W/m ² -K)	Roof Thermal Resistance (m ² -K/W)	Wall Thermal Resistance (m ² -K/W)
Levels	0.4	0.4	0.2	Type 1: U=0.39, SHGC=0.40	5.4	4
	0.6	0.6	0.4	Type2: U=0.33, SHGC=0.40	7.4	6
	0.8	0.8	0.6	Type3: U=0.25, SHGC=0.40	9.4	8
			0.8	Type4: U=0.19, SHGC=0.20	11.4	10
				13.4	12	

As an experience, an energy simulation input file containing a design combination is constructed and introduced to the simulation program to run. The simulations are repeated year-by-year for the future years to evaluate the performance of each design over a period of time.

To evaluate the performance of all the design options, each design option must be simulated with 120 future weather years namely, 30 future years for each of the four climate change scenarios. Considering the assumed 3600 design options and 120 weather years, 432000 simulations are required to evaluate all the design options which is quite time consuming. This was also an issue in a previous study [29]; that is why the authors in previous study made three weather files, namely, extremely cold, extremely warm, and a typical future year, followed by optimization to find the best solution. In the current study however, surrogate models are used to evaluate the design options; in this case, deep neural network models are trained on the results of significantly shorter number of simulations; then the models are used to find the performance of all the design alternatives. Surrogate models have been used in previous studies as well and many of them showed promising accuracy compared with simulation results [39-41].

8.2.3 Data preprocessing

In order to prepare training data for the surrogate models, six representative years with different cooling demand levels and six representative years with heating demand levels are selected to run the full factorial design options. The representative years are selected in a way to cover at least the minimum, average, and maximum cooling and heating demands over the 62 historical actual years in addition to the 120 (4 scenarios* 30 years) future years with the NCEB 2015 standard (baseline) building. Figure 8.2 shows the representative years and corresponding cooling and heating demand

selected to be fully simulated to be used as input for the surrogate models. For cooling, the actual weather years of 1956, 1989, 1996, 2001 and the future weather years 2031 under RCP 4.5 and year 2042 under RCP 6 are selected for training of the models and the year 2014 is selected to test the models. For heating, the actual weather years of 1976, 1994, 1996, 2001, 2006 in addition to the future year 2038 under RCP 6 is selected for training of the models and the year 2014 is selected to test the models. Therefore, the full-factorial design options with the representative years are simulated using EnergyPlus energy simulation program coupled with Python programming.

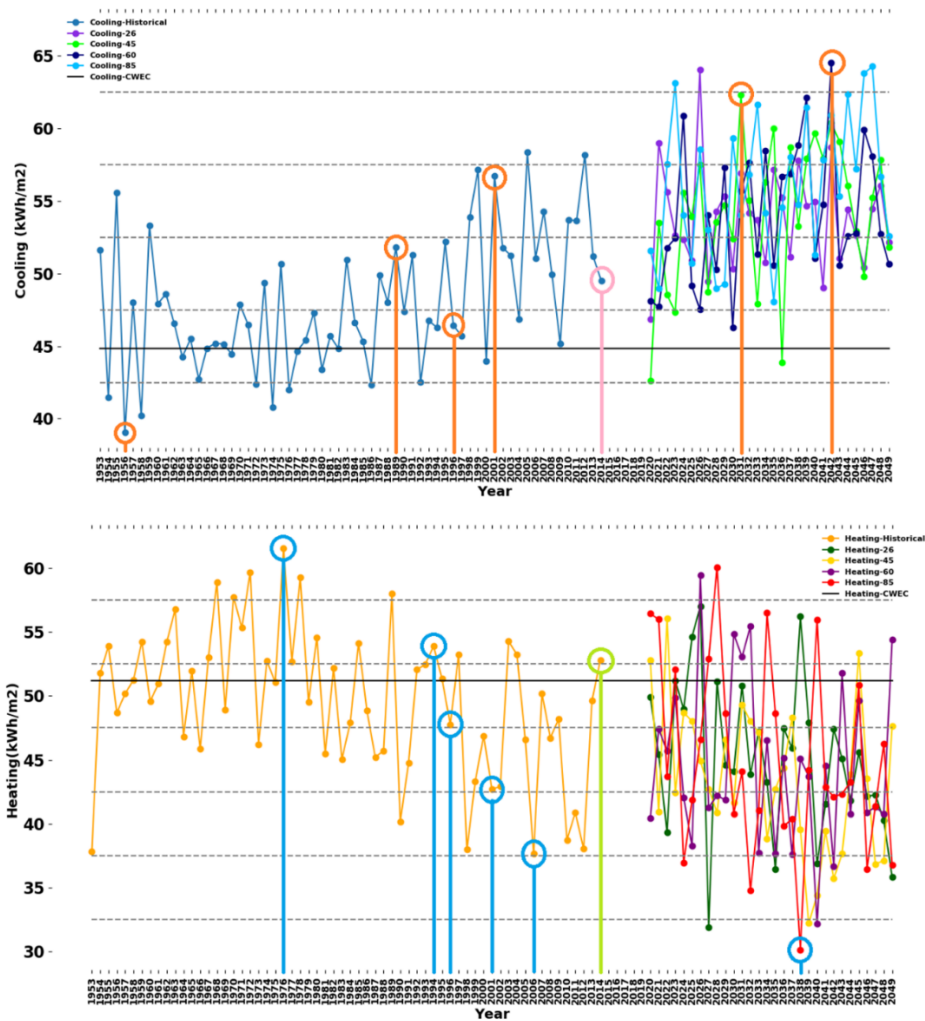


Figure 8.2: Annual cooling and heating of the small office building under historical, CWEC, and future weather years; orange and blue circles indicate the weather years selected for training of heating and cooling models respectively; the pink and green circles indicate the weather year for test of the cooling and heating models respectively.

The simulation results of all the design alternative with the representative weather years are used for training the surrogate models. To train the models, the simulation results need to be processed to convert to a language that can be learned by the machine.

In order to train the surrogate models, the input data is needed to be processed. The input data to the models include the building design data, temporal data, and weather data which are described further below.

8.2.3.1. Design data encoding

The design data related to the building enclosure, previously shown in Table 8.1, included numerical data such as solar reflectance, or categorical data such as window type. One Hot Encoding (OHE) technique is used to encode this type of data. The technique converts n unique levels of a design parameter (a column in the dataset of design) into n different new features (n columns) with new binary values of 0 or 1. This technique is specifically useful for design variables whose values do not necessarily follow an order or hierarchy in the dataset; otherwise, models might confuse during the training process. Therefore, the six design parameters (six columns in the dataset) are converted to 24 new features (columns) with 0 and 1.

8.2.3.2. Temporal data encoding

To consider the building's temporal features, namely, building schedules, the temporal data is also encoded. The temporal data plays a significant role in the modeling as it can capture the hourly schedule of the building during the days, weekday/weekend, and holidays in modeling.

The temporal data include hour of the day, day of the week, day of the month, and the month of the year. Therefore, there were four columns of temporal data in the original training data set. Hour of the day column has values 1 to 24; day of the week with values between 1 and 7, day of the month with values 1 to 31, and the month of year with values 1 to 12.

Similar to the design data, the temporal data is encoded using OHE because timestep $t+1$ does not necessarily have more/less cooling or heating demand than timestep n . Due to the computational memory limitation and better capturing seasonal effect the whole hourly data is divided into four seasons data (3 months data each). The first season is considered January-March, the second season April-June, the third season July-September, and the last season October-December.

For each season, the hour of the day converted to new 24 features, the day of the week converted to 7 new features, the day of the month converted to new 31 features, and the month of the year converted to new three features (each season data has 3 month data) all having values 0 or 1. Overall, instead of the 4 temporal features, 65 new features are used in the dataset.

8.2.3.3. Weather data encoding

The four weather parameters of dry bulb temperature, horizontal solar irradiance, atmospheric pressure, and wind speed were used as input to the models. The weather data were encoded using min-max scaling using the following equations to convert all the data into values between 0 and 1.

$$X_{\text{scaled}} = (X - X_{\text{min}}) / (X_{\text{max}} - X_{\text{min}}) \quad (4)$$

Where: X is each of the weather parameter data before scaling, X_{scaled} is each of the weather parameter data after scaling, X_{min} is the minimum value of each weather parameter, and X_{max} is the maximum value of each weather parameter in the dataset.

After processing the input parameters, the overall 93 input features, together with corresponding cooling or heating demand, are introduced to the models for training purposes.

8.2.4 Model

Feed Forward Neural Networks (FFN) have been previously used to build energy prediction as surrogate models or forecast actual energy consumptions of buildings with remarkable accuracy in previous studies [39-41]. There are various types of neural networks including Multi Perceptron neural networks (MLP), Recurrent neural network (RNN) including Long-Short-Term-Memory (LSTM), and Convolutional neural networks (CNN) that can be used for time-series prediction and forecasting.

The MLP FFN is more straightforward compared to the other two; there is no feedback from output to input, and the information moves only in one direction. Unlike LSTM, the MLP is a good autoregressive model that is quite necessary for the current problem because it includes multiple design options and multiple weather years from different climate change scenarios. This study's final goal is to find the best design solution against climate change; some designs might result in very similar outputs, whereas others might result in significantly different outputs. The building with multiple design options can be considered a complex non-linear system with different behaviors (designs),

requiring accurate models to predict cooling and heating demands. In this study, eight Multi Perceptron (MLP) feed-forward deep neural network models are trained to predict the building's cooling and heating demand. The eight models include four models for the cooling demand of four season data and four heating demand models for four seasons. The reasons for selecting eight different models instead of one model are twofold:

The training data is quite large, and importing all data at one step is very challenging, if possible, in case of computation memory. Dividing the whole weather data into different seasonal data can make the models more accurate, as the data's variance reduces.

There are 93 features in each model's input, including all the design features, temporal features, and the weather features; in the output, there is hourly cooling or heating demand values.

There is no particular rule for selecting the number of hidden layers and neurons; however, from the literature, for n input features, two hidden layers with $2n+1$ neurons in the first hidden layer and n neurons in the second layer showed accurate outputs. Therefore, two hidden layers are considered for the model. The first hidden layer has 187 neurons. The second hidden layer has 93 neurons; the output layer is a dense layer with one neuron for cooling or heating. Figure 8.3 shows the architecture of the models.

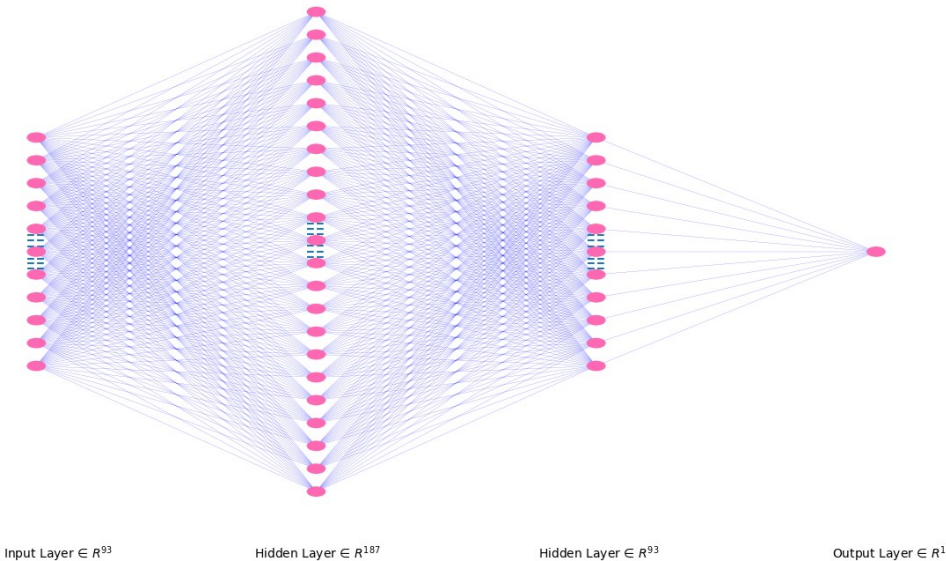


Figure 8.3: The architecture of the deep neural network model with two hidden layers; there are 187 neurons in the first hidden layer and 93 neurons in the second hidden layer. Note: the number of neurons in the layers are reduced in the figure for better presentation.

Two hyper-parameters in training the networks are number of epochs and batch size. Epoch is the number of forward and backward passes of introducing samples to the model for training; batch size is the number of samples in each epoch. 80 epochs with batch size of 200 were considered for models.

8.2.4.1 Model validation

The models are validated using 20% of random data in the train dataset in each epoch of each model's training. Figures 8.4 and 8.5 show the history of the training process. The models are very well trained in the data as mean absolute error and mean squared error show reduction trends over epochs and the errors on train and validation dataset converge.

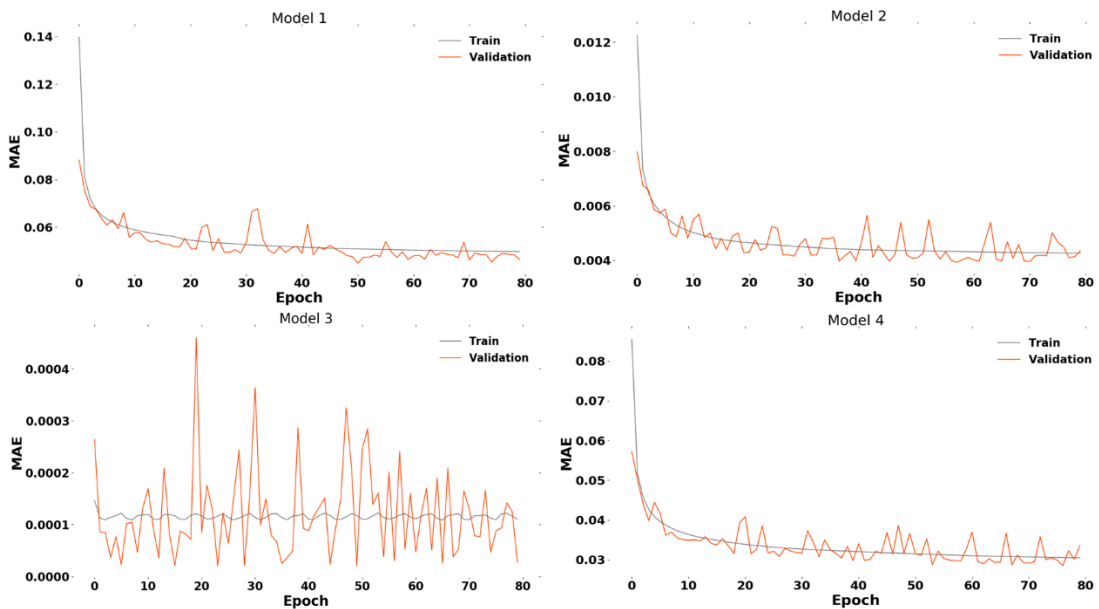


Figure 8.4: History of models trained for heating demands over epochs.

The four models are trained on four different seasons heating demand data; therefore, the mean absolute errors have different ranges. Furthermore, the Model 3 is trained on data from the beginning of July and the end of September, where heating demand is zero or relatively small. Although the figure shows fluctuations, the MAE reduces with further epochs, and the range of the MAE values are quite negligible.

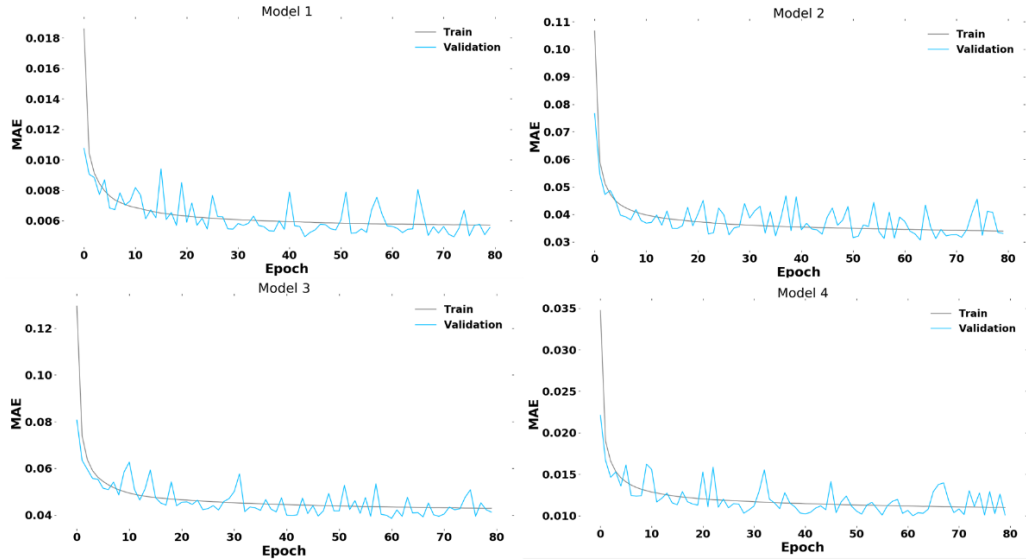


Figure 8.5: History of models trained for cooling demands over epochs.

8.2.4.2 Model test

In addition to model validation, all the models' performance is tested with test data that included weather data that was not introduced to the models during training process. In the test process, the weather data of the year 2014 and temporal and design data are given to the models for predictions. Afterward, the eight models' output is compared with the simulation results with the weather year 2014; the mean absolute error, mean squared error, root mean squared error, and R^2 of the models are calculated and shown in table 8.2.

Table 8.2: The performance of the eight trained models in predicting the cooling and heating for 3600 design with the weather year 2014 (test data).

Cooling				Heating			
MAE	MSE	RMSE	R^2	MAE	MSE	RMSE	R^2
57.12	4235.6	65.08	0.9998	44.18	2870.1	53.57	0.9997

Table 8.2 shows the eight models' performance in predicting annual cooling and heating of all the 3600 designs for the year 2014 compared to the simulation results. From the table, the models are entirely accurate in predicting the cooling and heating demands of the buildings for the 3600 design options. The models' annual cooling and heating demand predictions on test data are shown in

figures 8.6 and 8.7 with blue and orange colors. The simulation results for the same design options and weather year 2014 are shown with the grey color. As can be seen, there are fair agreements between the model's output and the simulation results.

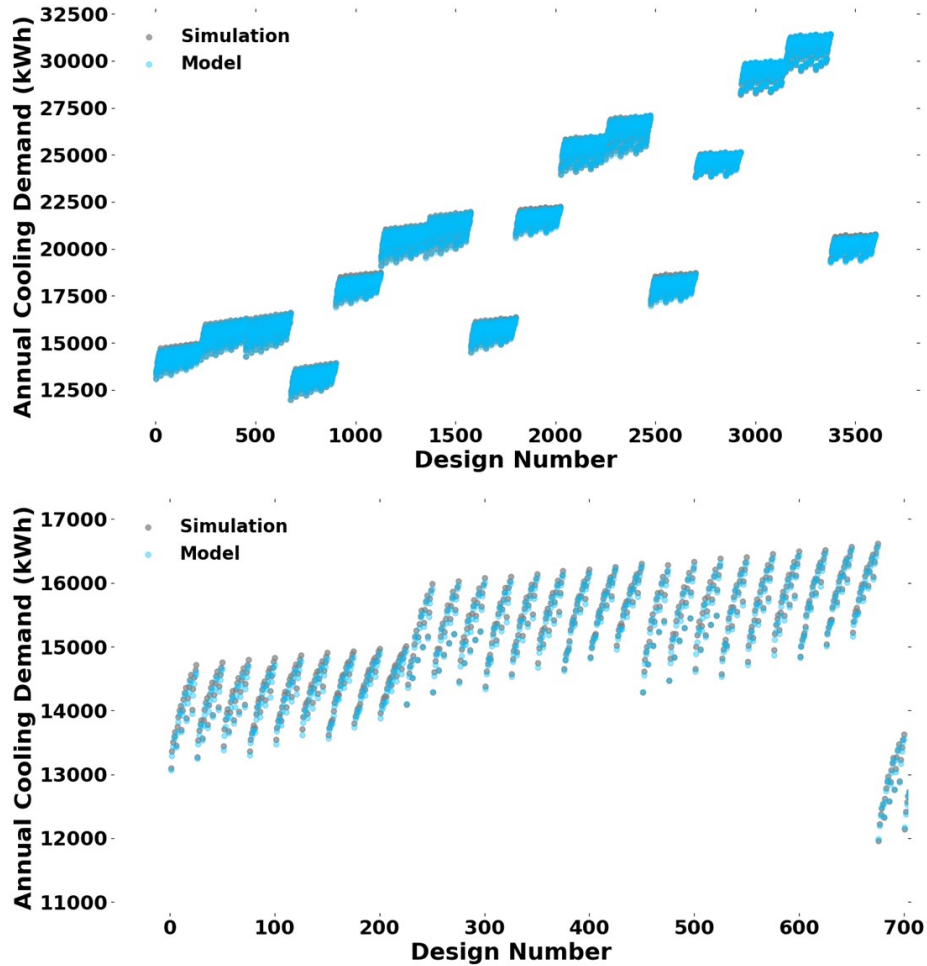


Figure 8.6: Predicted annual cooling demands of the trained models with weather year 2014 (test data) in comparison to the simulation results; top: the whole 3600 designs, bottom: the first about 700 designs.

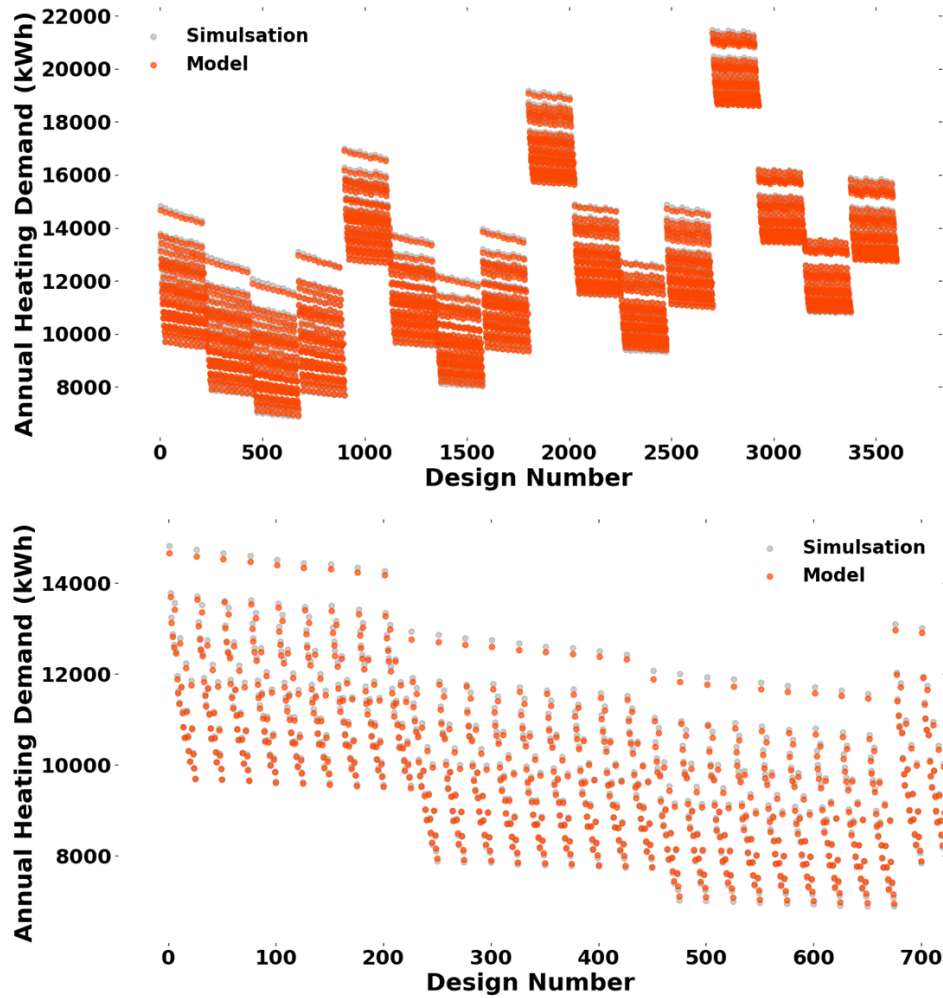


Figure 8.7: Predicted annual heating demands of the trained models with weather year 2014 (test data) in comparison to the simulation results; top: the whole 3600 designs, bottom: the first about 700 designs.

8.2.4.3 Data post-processing

After training and testing the models, all the future weather years and all the design options and temporal features are given to the models to predict all the future weather years under the four RCPs. For each design option, the hourly results are resampled and summed annually for each year.

8.2.5 Performance indicator

The designs' information content is expressed as a performance indicator based on the probability of success to consider the information axiom.

$$I = \sum_i \log_2\left(\frac{1}{p_i}\right) = - \sum_i \log_2(p_i) \quad i \in \{\text{RCP2.5, RCP4.5, RCP6, RCP8.5}\} \quad (5)$$

Where: I is the information content of a design under the four RCP scenarios with unit of bits, and p is the probability that a design meet target under a specific RCP or number of years the design meet the targets over all the future years (e.g. 30 years for each scenario).

As mentioned before, according to the axiomatic design, the best designs are those having the minimum information contents; in other words, the design that contributes to the smallest uncertain information.

8.2.6 Workflow summary

Figure 8.8 shows the workflow of the methodology. The case study with NECB 2015 standard (baseline) is simulated with historical and future weather years. The results of the baseline simulation are used to select representative weather years. Six representative weather years for training purposes and one weather year for testing purposes are selected. The representative weather years then are used for full factorial design simulations (orange blocks). The simulation results are preprocessed and divided into train and test data to be used for training and testing of the models (blue blocks). Once the cooling and heating models are trained, they are tested with test data, and the results are compared with the simulation results for weather year selected for testing. After making sure that the models showed acceptable performance, the models are used to predict all the design alternatives with all the future weather years. The output of the models are then post-processed; for each design option, the hourly results are resampled and summed annually; this is followed by applying the information axiom to find the best design solutions using equation 2 from the design space.

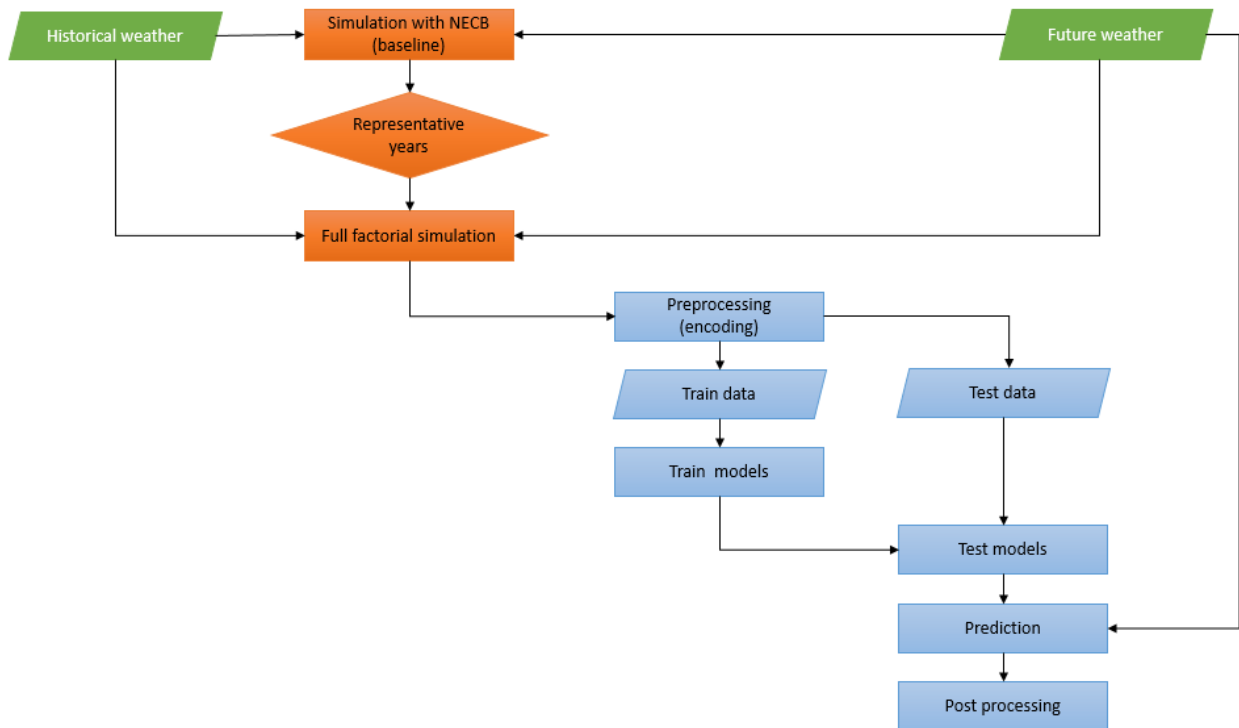


Figure 8.8: The workflow of the methodology

8.3 Results

The models' outputs of cooling and heating demands of all designs for each future year under the four RCPs are shown in figure 8.8 for better comparison. The cooling and heating demands of all designs with the CWEC weather are shown with orange color.

In all the scenarios, the heating demands resulting from CWEC are generally lower than the future years. The cooling demands are higher than the future years for the same design options. With CWEC, the largest heating and cooling demands are about 60 kWh/m², and the smallest cooling and heating demand of about 22 kWh/m² and 17 kWh/m², respectively. Whereas, certain designs show a cooling demand of up to about 75 kWh/m² under RCP 6 and heating demand of up to about 50 kWh/m² under RCP 6 and 8.5, and minimum cooling of about 20 kWh/m² under RCP 4.5 and minimum heating demand of about 7 kWh/m² under RCP 2.6.

However, designs show different performance under different scenarios. Therefore, the designs with the lowest cooling and heating demand under different scenarios or the designs with the most probability to meet the performance targets can be found through design space exploration.

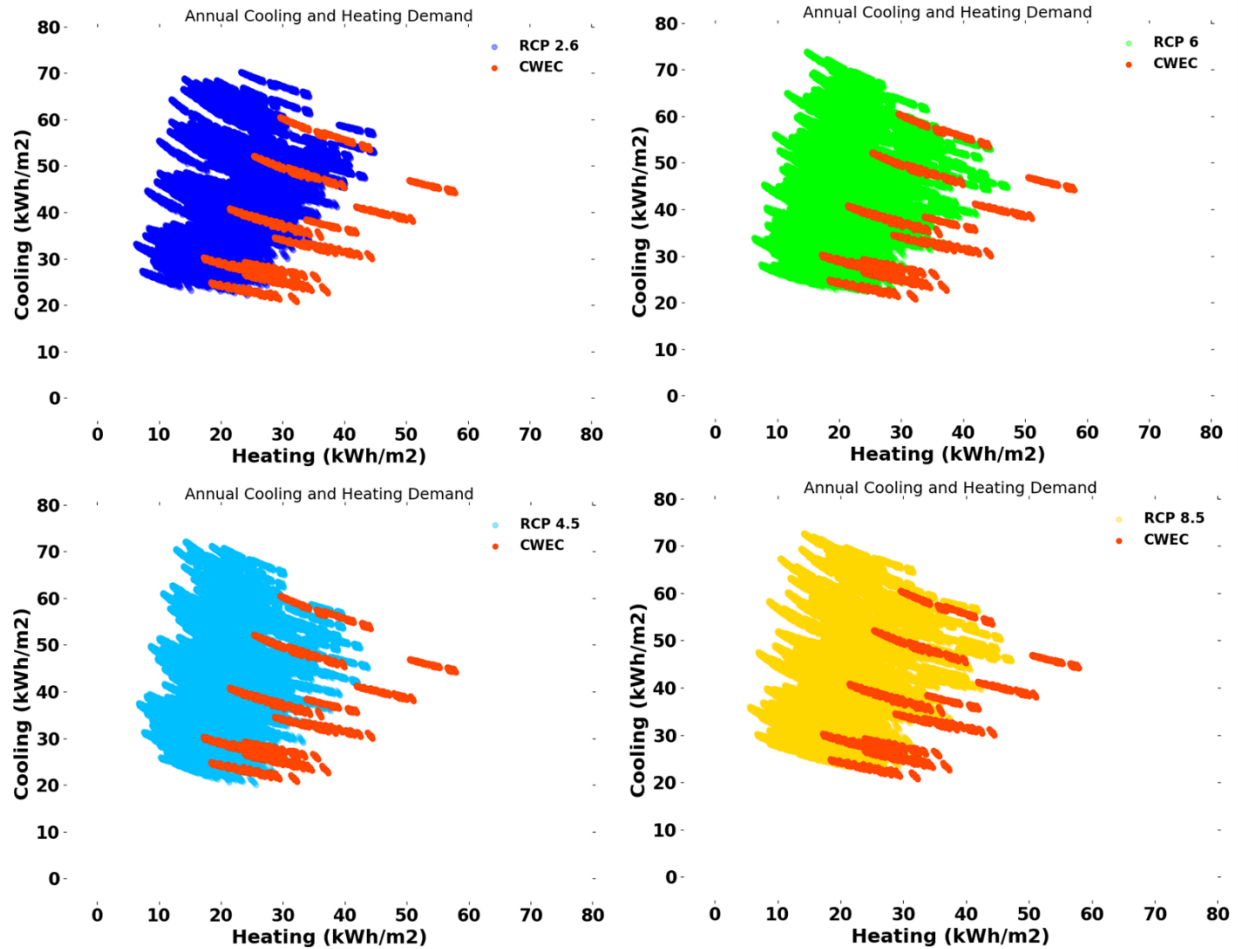


Figure 8.9: Annual cooling and heating demand under four climate change scenarios; under each scenario; 108000 points represent the demand of 3600 design combination for 30 future years; the 3600 orange points represent the same design combinations under CWEC weather condition.

8.3.1 Design space exploration

The results obtained from the models contained 3600 different design combinations with different cooling and heating demands. The best design options are selected with inspiration from a system design methodology called axiomatic design, which transfers customers' needs into functional requirements that can be achieved by manipulating design parameters [30].

The probability of meeting the targets can be defined as the number of years the heating or cooling corresponding of design is less than a target over the next 30 years under each RCP using the following equation. For example, if the target of cooling demand is defined as 30 kWh/m², the probability of meeting the cooling target of 60% under a specific RCP means the design contributes to a cooling demand less than 30 kWh/m² in 18 out of 30 future years under the specific RCP.

The design with NECB 2015 code showed in figure 8.1, indicated the next thirty years average cooling demand of 53.8, 54.1, 54, and 56 kWh/m² and heating demand of 45.1, 43.5, 44.6, and 45.1 kWh/m² under RCPs of 2.6, 4.5, 6, and 8.5 respectively. Therefore, the cooling and heating demand should not surpass 53.8 and 45.1 kWh/m² for respectively cooling and heating demand in most of the future years. With these two functional requirements and 60% probability of success, the results show 2528 design options meet these criteria, as shown in Figure 8.10; each point can represent the probability of success for multiple designs. The purple points with blue points in the center show that some designs meet the targets under a specific scenario but not necessarily under all the scenarios.

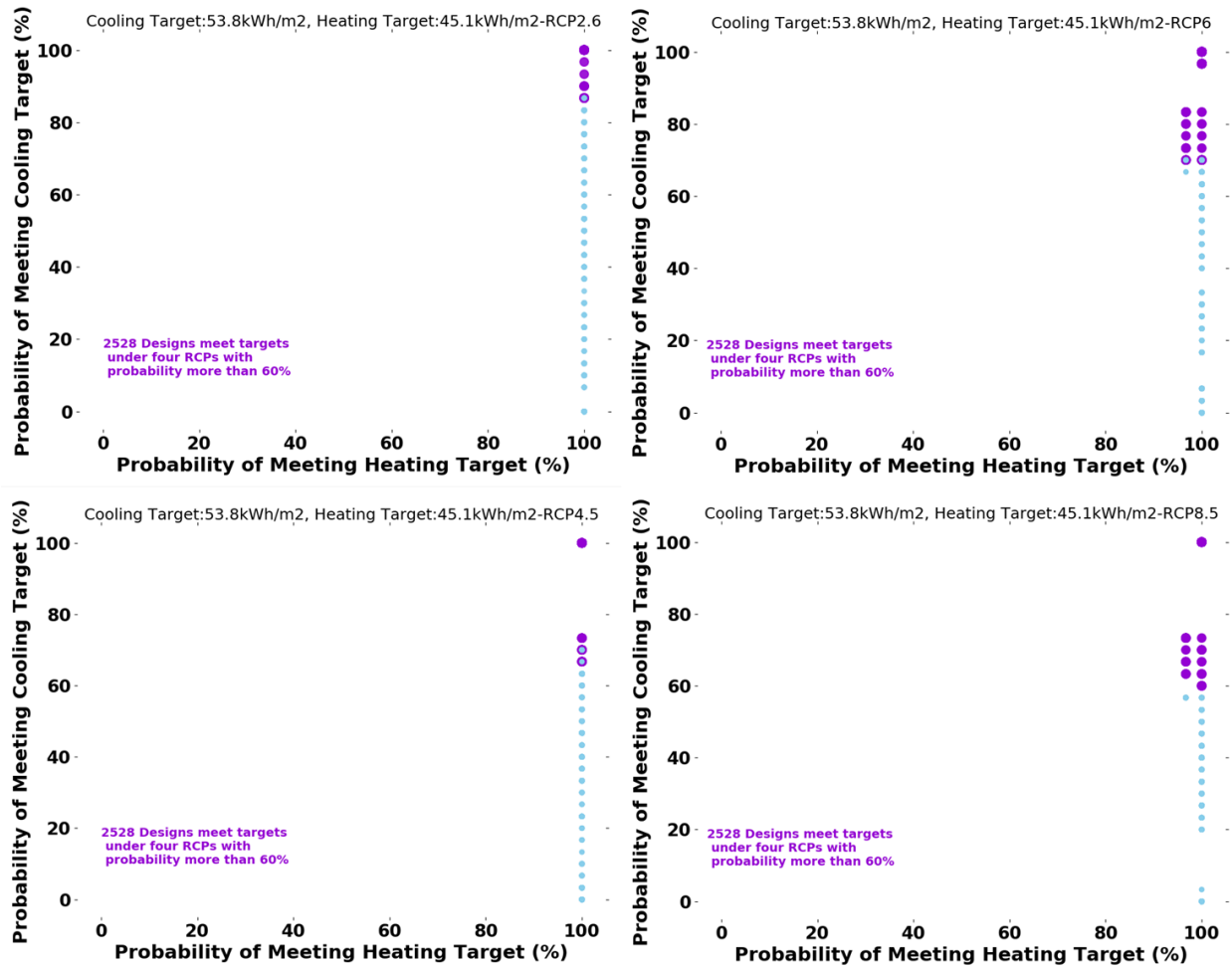


Figure 8.10: Probability of meeting targets (functional requirement) for cooling and heating for various designs. The blue points represent the probabilities under each RCP and the purple points show the performance of designs meeting the targets with more than 60% under all the four scenarios.

With tougher functional requirements (targets) of 30 kWh/m² for cooling and 25 kWh/m² for heating, the probability of meeting the targets under the RCPs are plotted in figure 8.11. A designer can be 80% confident, that 12 designs out of 3600 design options will meet the targets under all the four scenarios which are shown with solid purple points.

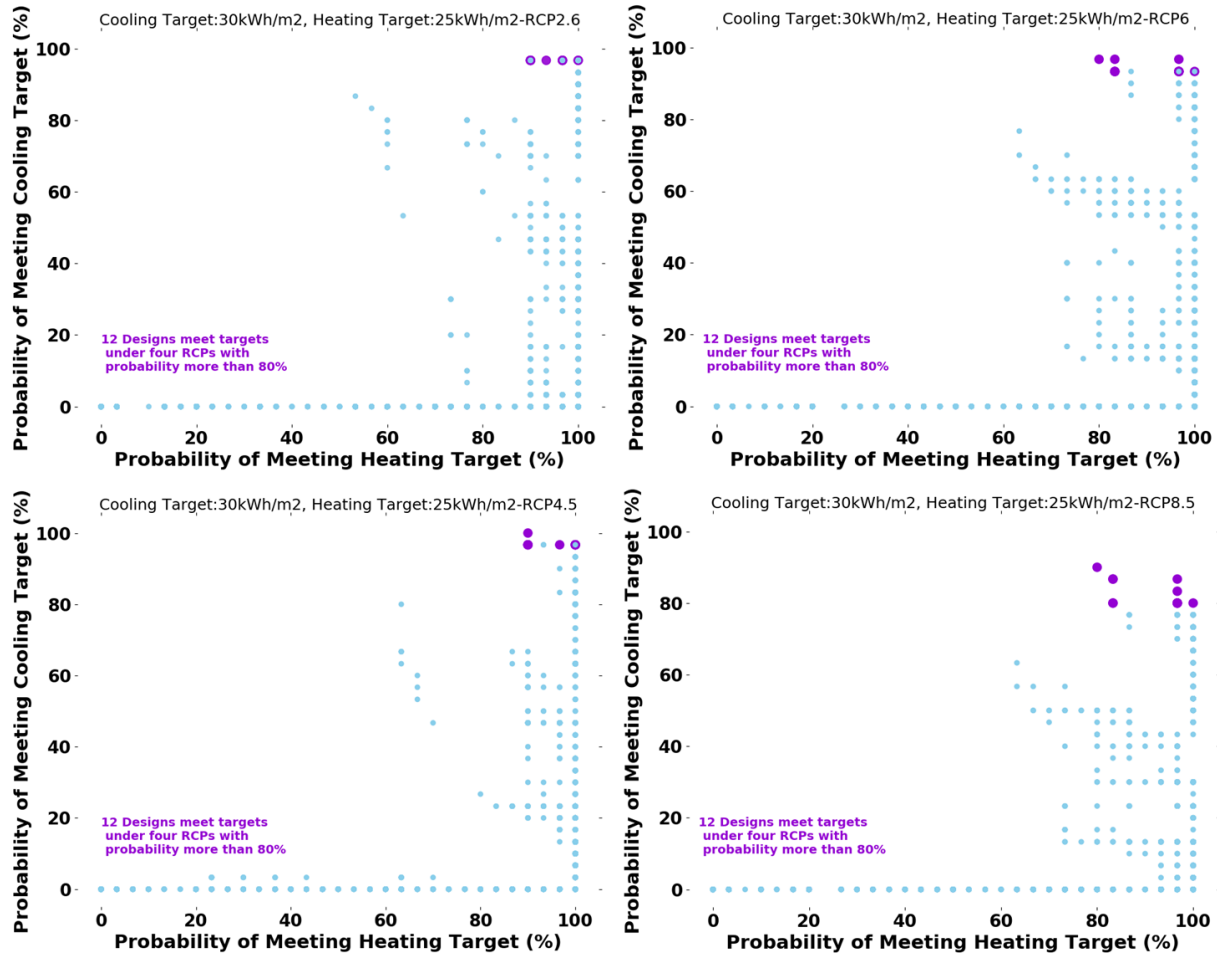


Figure 8.11: Probability of meeting targets (functional requirement) for cooling and heating for various designs. The blue points represent the probabilities under each RCP and the purple points show the performance of designs meeting the targets with more than 80% under all the four scenarios.

Numerous designs meet the cooling and heating demand target, at least in one future year under each scenario. Some designs meet the targets only once in the future years, whereas other designs meet the targets in most future years. Therefore, those designs that meet the targets more frequently or the design with more probability of success are considered better design according to the axiomatic design concept. For the same cooling and heating demand targets, the information content of all the designs that meet the targets at least in one future year is shown in figure 8.12. Those designs with lower information content are the designs with a higher probability of success. From the figure, design numbers between 700 and 900 show relatively smaller information contents (less than 4 bits) in comparison to other design alternatives.

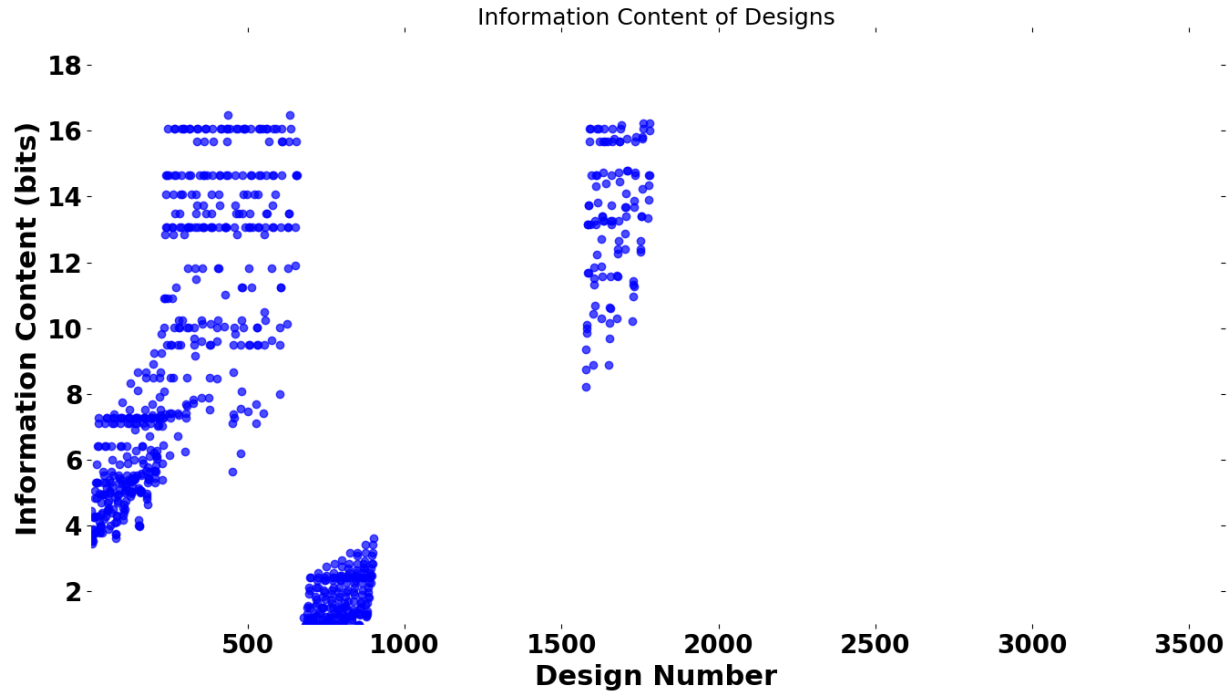


Figure 8.12: The information content of designs that meet cooling demand target of 30 kWh/m² and heating demand of 25 kWh/m² at least in one future year in each scenario.

The 12 designs meeting the targets under all the scenarios with probability more than 80% are summarized in Table 8.3.

Table 8.3: The 12 designs meeting the targets with more than 80% probability under all the four RCP scenarios

	WWR	Window Type (U Factor in W/m ² -K)	Roof Thermal Resistance (m ² -K/W)	Wall Thermal Resistance (m ² -K/W)	Roof Solar Reflectance	Wall Solar Reflectance	Information (bits)
1	0.2	Window_U_0.19_SHGC_0.20	5.4	4	0.6	0.6	1.1966899
2	0.2	Window_U_0.19_SHGC_0.20	5.4	6	0.6	0.6	0.5964203
3	0.2	Window_U_0.19_SHGC_0.20	5.4	8	0.6	0.6	0.6156279
4	0.2	Window_U_0.19_SHGC_0.20	5.4	10	0.6	0.6	0.5188035
5	0.2	Window_U_0.19_SHGC_0.20	7.4	4	0.6	0.6	0.6057236
6	0.2	Window_U_0.19_SHGC_0.20	5.4	4	0.6	0.4	1.1823621
7	0.2	Window_U_0.19_SHGC_0.20	5.4	6	0.6	0.4	0.6640401
8	0.2	Window_U_0.19_SHGC_0.20	7.4	4	0.6	0.4	0.6640401
9	0.2	Window_U_0.19_SHGC_0.20	5.4	4	0.6	0.2	1.3500329
10	0.2	Window_U_0.19_SHGC_0.20	5.4	4	0.4	0.6	1.2340009
11	0.2	Window_U_0.19_SHGC_0.20	5.4	6	0.4	0.6	0.6640401
12	0.2	Window_U_0.19_SHGC_0.20	5.4	4	0.4	0.4	1.3500329

From Table 8.3, while all the design options meet the targets with more than 80% probability under the four scenarios, some designs' information content is less than others and, therefore, are better designs according to the second axiom. All the design options suggest a window-to-wall ratio of 0.2 with window type 4, namely, with U factor 0.19 W/m²-K and SHGC 0.2. The reason must be because a smaller window with smaller solar heat gain prevents heat gain through the coming sunny warm days and prevents the heat loss during the cold days in the future.

In most designs, roof thermal resistance of 5.4 m²-K/W with a solar reflectance of 0.6 is suggested to reduce the building's cooling load. The low thermal resistance of 5.4 m²-K/W is because a high thermal resistance leads to a heat trap in the attic, which leads to a high cooling load for the building. However, walls thermal resistance is suggested to have a low thermal resistance of 4 W/m²-K to the high value of 10 W/m²-K yet, the high thermal resistance of 12 W/m²-K is not found in any of the 12 designs. The wall solar reflectance is also suggested to be a value of 0.2, 0.4, or 0.6; this shows that the building may be more sensitive to the roof than the wall, although the smallest information content belongs to the design number 4 with thermal resistance 10 W/m²-K and solar reflectance 0.6. Since the sun angle is lower and less solar radiation is available in winter, design number 4 helps thermal zones keep cooler in summer and warmer in winter under different climate change scenarios.

Overall, the roof solar reflectance is mostly suggested to be 0.6 while the roof thermal resistance is suggested to be the minimum value of 5.4 W/m², which is equal to the value suggested by the current NECB standard. The suggested values generally express a smaller window to wall ratio with lower U-value windows; besides, the smaller roof and wall thermal resistance indicate higher thermal resistance is not necessarily associated with sustainability in design. The higher roof solar reflectance in roof and wall indicates the cooling strategy of building through the roof and wall in summer.

8.4 Discussion

The probability calculated for the designs can have other advantages. The probability of success calculated in previous section can be used for other purposes such as risk assessment as well which is described in the following.

8.4.1 Risk definition

For a project investor and policymakers, an insight into the success of the project is significantly essential. Generally, the risk is defined as the expected consequence associated with an event. Therefore, the product of the probability and consequence of an event calculates the risk. Considering that many utilities have different policies and rates regarding the electricity price, some level of caution is suggested to calculate the risk. However, a fundamental way of calculating the partial risk related to the energy performance of design of a sustainable building can be the product of the probability of the failure of the design ($P_f = 1 - P$) and the electricity cost.

8.4.2 Risk treatment

The risk analysis can be conducted based on uncertain parameters that can be because of:

- a. natural variation (in this case, natural variation of weather under different climate change scenarios).
- b. modeling, which can be either related to the parameters that are ignored in the modeling or related to the model's mathematics.
- c. the statistical uncertainty that is related to the techniques of the probability calculation.

This study's objective is the item a. which is the natural variation and the effect of climate change contributing to the risk in sustainable building design.

8.4.3 Risk mitigation

Risk mitigation of a sustainable building under climate change and weather variation can be carried out by modifying building energy performance control at the operational stage. For example, when a very cold or warm year is anticipated, or a forecasted sky day is predicted, the occupants are asked to limit their miscellaneous electricity consumption to essential requirements. Predictive control strategies can implement this.

8.4.4 Risk reduction

The risk reduction is performed at the design stage by selecting a high probability of success as a target at the design stage. The risk can also be reduced by reducing the failure's consequence at the operational stage through control optimization.

8.4.5 Risk transfer

The risk calculation associated with the design can be used by the third party companies such as insurance, or financial centers.

8.4.6 Risk acceptance

The methodology used in this study enables the decision-makers to select the designs that comply with risk acceptance criteria.

8.5 Contribution

In this study, a workflow is suggested to consider climate change and its effects on the buildings' energy performance and suggest building designs that meet the predefined performance targets under uncertain climate change scenarios with probability defined by designers.

The proposed method provides the opportunity to observe the variability of each design solution's energy performance under the effect of climate change. The probability of success is introduced to design options at the design stage such that those design solutions with a minimum information content can be selected as the final design solutions. The proposed method helps to select the designs contributing to resilient performance against the effect of climate change; meanwhile, it will help avoid under/overdesign building thermal characteristics while keeping the targets. Depending on the project delivery method, this might be very important for the design-builder team or the owner who will be contractually responsible for the actual performance of the design to meet or exceed the design target. Furthermore, the probability of success calculated in the workflow can be used by the third party companies.

8.6 References

[1] V. Pérez-Andreu, C. Aparicio-Fernández, A. Martínez-Ibernón, J.L. Vivancos, Impact of climate change on heating and cooling energy demand in a residential building in a Mediterranean climate, *Energy*, Volume 165, Part A, 15, 63-74, 2018

[2] J. B. Dias, P.M.M. Soares, G.C. da Graça, The shape of days to come: Effects of climate change on low energy buildings, *Building and Environment*, 181, 15, 107125, 2020

[3] J. Fonseca, I. Nevat, G.W. Peters, Quantifying the uncertain effects of climate change on building energy consumption across the United States, *Applied Energy*, 277, 1, 115556, 2020

- [4] U. Berardi, P. Jafarpur, Assessing the impact of climate change on building heating and cooling energy demand in Canada, *Renewable and Sustainable Energy Reviews*, 121, 109681, 2020
- [5] I.G. Dino, C.M.al Akgül, Impact of climate change on the existing residential building stock in Turkey: An analysis on energy use, greenhouse gas emissions and occupant comfort, *Renewable Energy*, 141, 828-846, 2019
- [6] A. Robert, M. Kummert, “Designing net-zero energy buildings for the future climate, not for the past”, *Building and Environment* 55, 150-158, 2012
- [7] J.M. Rey-Hernández , C. Yousif , D. Gatt , E. Velasco-Gómez , J. San José-Alonso, F.J. Rey-Martínez, “Modelling the long-term effect of climate change on a zero energy and carbon dioxide building through energy efficiency and renewable”, *Energy & Buildings* 174, 85–96, 2018
- [8] P. Shen, N. Lior, Vulnerability to climate change impacts of present renewable energy systems designed for achieving net-zero energy buildings, *Energy*, 114, 1, 1288-1305, 2016
- [9] E.L.A. da Guarda, R.M.A. Domingos, S.H.M. Jorge, L.C. Durante, J.C.M. Sanches, M. Leão, I.J.A. Callejas, The influence of climate change on renewable energy systems designed to achieve zero energy buildings in the present: A case study in the Brazilian Savannah, *Sustainable Cities and Society*, 52, 101843, 2020
- [10] P. Shen, W. Braham, Y. Yic, The feasibility and importance of considering climate change impacts in building retrofit analysis, *Applied Energy*, 233–234, 1, 254-270, 2019
- [11] J. Chai, P. Huang, Y. Sun, Differential evolution - based system design optimization for net zero energy buildings under climate change, *Sustainable Cities and Society*, 55, 102037, 2020
- [12] M. Gercek, Z.D. Arsan, Energy and environmental performance based decision support process for early design stages of residential buildings under climate change, *Sustainable Cities and Society*, 48, 101580, 2019
- [13] Gh.R. Roshan, R. Oji, Sh. Attia, Projecting the impact of climate change on design recommendations for residential buildings in Iran, *Building and Environment*, 155, 15, 283-297, 2019
- [14] B. Givoni, Comfort, climate analysis and building design guidelines, *Energy and Buildings*. 18, 11–23, 1992
- [15] C. Rubio-Bellido, A. Pérez-Fargallo, J.A. Pulido-Arcas, Optimization of annual energy demand in office buildings under the influence of climate change in Chile, *Energy*, 114, 1, 569-585, 2016

- [16] P. Shen, “Impacts of climate change on U.S. building energy use by using downscaled hourly future weather data”, *Energy and Buildings*, 134, 61–70, 2017
- [17] A. Jiang, Y. Zhu, A. Elsafty, M. Tumeo, “Effects of Global Climate Change on Building Energy Consumption and Its Implications in Florida”, *International Journal of Construction Education and Research*, 14(1), 22-45, 2018
- [18] L. Wang, X. Liu, H. Brown, “Prediction of the impacts of climate change on energy consumption for a medium-size office building with two climate models”, *Energy and Buildings* 157, 218–226, 2017
- [19] M. Hosseini, F. Tardy, B. Lee, “Cooling and heating energy performance of a building with a variety of roof designs; the effects of future weather data in a cold climate”, *Journal of Building Engineering*, 17 107-114, 2018
- [20] U. Yao, A. Tettey, A. Dodoo, L. Gustavsson, “Energy use implications of different design strategies for multi-storey residential buildings under future climates”, *Energy* 138, 846-860, 2017
- [21] M. Cellura, F. Guarino, S. Longo, G. Tumminia, “Climate change and the building sector: Modelling and energy implications to an office building in southern Europe”, *Energy for Sustainable Development* 45 46–65, 2018
- [22] M.F. Jentsch, A.S. Bahaj, P.A.B. James, “Climate change future proofing of buildings—Generation and assessment of building simulation weather files”, *Energy and Buildings* 40, 2148–2168, 2008
- [23] M. Palme , A. Isalgué, H. Coch , Avoiding the possible impact of climate change on the built environment: the importance of the building’s energy robustness, *Buildings*, 3 (1), 191–204, 2013
- [24] G. Chinazzo , P. Rastogi , M. Andersen , Robustness assessment methodology for the evaluation of building performance with a view to climate uncertainties, in: *Proceedings of BS 2015*.
- [25] P. Hoes, J.L.M. Hensen, M.G.L.C. Loomans, B.de Vries, D. Bourgeois.”User behavior in whole building simulation, *Energy and Buildings*, 41 (3), 295–302, 2009
- [26] W. O’Brien, Occupant-proof buildings: can we design buildings that are robust against occupant behaviour, 13th Conference of International Building Performance Simulation Association, 2013.
- [27] R. Kotireddy , P. Hoes , J.L.M. Hensen , A methodology for performance robustness assessment of low-energy buildings using scenario analysis, *Applied Energy*, 212, 428–442, 2018
- [28] G. Taguchi, “Taguchi on Robust Technology Development Bring Quality Engineering Upstream, 1993, The American Society of Mechanical Engineers, Yew York, NY.

- [29] A. Moazami, S. Carlucci, V. Nik, S. Geving, Towards climate robust buildings: An innovative method for designing buildings with robust energy performance under climate change, *Energy & Buildings*, 202, 109378, 2019
- [30] N.P. Suh, *Axiomatic Design: Advances and Applications*, Oxford University Press, 2001, ISBN 0-19-513466-4
- [31] M. Karatas, Hydrogen energy storage method selection using fuzzy axiomatic design and analytic hierarchy process, *International Journal of Hydrogen Energy*, 45, 32, 16227-16238, 2020
- [32] A. Verma, J. Maiti, G. Boustras, Analysis of categorical incident data and design for safety interventions using axiomatic design framework, *Safety Science*, 123, 104557, 2020
- [33] V.R. Palleti, J.V. Joseph. A. Silva, A contribution of axiomatic design principles to the analysis and impact of attacks on critical infrastructures, *International Journal of Critical Infrastructure Protection*, 23, 21-32, 2018
- [34] M. Cavique, A.M. Goncalves-Coelho, Axiomatic design and HVAC systems: An efficient design decision-making criterion, *Energy and Buildings*, 41, 146–153, 2009
- [35] City of Toronto, Toronto Green Standard Version 3, Energy/GHG & Resilience for Mid to High-Rise Residential & all Non-Residential Development, retrieved on September 2020:
<https://www.toronto.ca/city-government/planning-development/official-plan-guidelines/toronto-green-standard/toronto-green-standard-version-3/mid-to-high-rise-residential-all-non-residential-version-3/energy-ghg-resilience-for-mid-to-high-rise-residential-all-non-residential-development/>
- [36] D.C. Montgomery, G.C. Runger. “Applied Statistics and Probability for Engineers”, Second Edition, John Wiley & Sons, Inc, New York, Chikester, Brisban, Toronto, Singapore, 1999, ISBN 0-471-17027-5
- [37] D.C. Montgomery, “Design and Analysis of Experiments”, 5th ed, John Wiley & Sons, Inc, New York, Chikester, Weinheim, Brisban, Toronto, Singapore, 1997& 2001, ISBN 0-471-31649-0
- [38] B. Lee, N. Pourmousavian, J.L.M. Hensen, “Full-factorial design space exploration approach for multi-criteria decision making of the design of industrial halls” *Energy and Buildings*, 117, 352–361, 2016
- [39] Z. Yu, F. Haghghat, B.C. Fung, H. Yoshino, A decision tree method for building energy demand modeling, *Energy and Building*, 42, 1637–1646, 2010
- [40] G.K. Tso, K.K. Yau, Predicting electricity energy consumption: a comparison of regression analysis, decision tree and neural networks, *Energy*, 32, 1761–1768, 2007

[41] S. Seyedzadeh, F. Pour Rahimian, P. Rastogic, I. Glesk, Tuning machine learning models for prediction of building energy loads, *Sustainable Cities and Society*, 24, 33–41, 2016

Chapter 9. CONCLUSION AND FUTURE WORK

9.1 Conclusion

For evaluating the building energy performance, the current practice is building simulation with a single typical meteorological year statistically selected from the actual historical year. The single representative year's main problem is that it cannot be used for extreme condition analysis and varying weather conditions. Knowing that building energy performance is considerably influenced by weather conditions, building energy performance can be significantly degraded under climate change. For example, one actual problem for cold climate buildings is the risk of overheating, which can endanger the occupant's health during extreme heatwaves. The main reason is that the buildings were initially designed for the heating-dominant condition a couple of decades ago. Under the effect of climate change, a major challenge is that the decision-makers including designers, investors, and engineers of the project cannot select the best design solution because they cannot foresee the consequence of uncertain events that affect the payoff of each design solution.

This study first introduced a novel machine learning approach to systematically reduce the deviation of the simulation with a typical meteorological year from the simulation with actual weather year results. The systematic approach replaces the universal expert's judgments-based with non-universal data-analytic-based using machine learning algorithms. The approach considerably improved the applicability of a single typical meteorological year.

Furthermore, a novel machine learning approach is introduced to downscale climate change to construct multiple year-by-year future weather files under different climate change scenarios. Unlike the simple so-called 'morphing' method, the method captures future climate data's intensity and frequency and constructs future weather years. This feature helps create logical future years for future building performance prediction and optimizing the performance considering natural weather variation under multiple climate change scenarios at the building's design stage. The constructed weather files can also be used for extreme conditions analysis.

Finally, a novel method is introduced to select the design variable that meets the requirements with a high probability of success under different climate change scenarios. The workflow helps architects and designers find the best design solutions under uncertain climate change conditions

at the buildings' design stage. Besides, policymakers can use the workflow to observe the energy performance of multiple building characteristics under climate change to make appropriate national, provincial, and municipality codes for building envelop and different buildings' energy performance in different climate zones.

9.2 Future work

The current study used a single GCM model containing all the four climate change scenarios RCPs. The current study is conducted with the assumption that the GCM used correctly predict the future climate condition. However, different GCMs might differently predict future climate due to the difference in algorithms and mathematical equations. Therefore as a first future work, an ensemble of multiple GCMs or regional climate models will be used for more reliable results.

Furthermore, the generated future weather files work best for building energy simulation with regards to energy consumption. For other energy systems such as renewables, the importance of weather parameters on outputs is different. Therefore, a new set of weather files can be constructed with updated weather parameters importance depending on the application as a second future task.