

Article

Impact Assessment in the Process of Propagating Climate Change Uncertainties into Building Energy Use

Hamed Yassaghi  and Simi Hoque *

Department of Civil Architectural and Environmental Engineering, Drexel University, 3141 Chestnut Street, Philadelphia, PA 19104, USA; hy387@drexel.edu

* Correspondence: sth55@drexel.edu; Tel.: +1-(215)-895-2254

Abstract: Buildings are subject to significant stresses due to climate change and design strategies for climate resilient buildings are rife with uncertainties which could make interpreting energy use distributions difficult and questionable. This study intends to enhance a robust and credible estimate of the uncertainties and interpretations of building energy performance under climate change. A four-step climate uncertainty propagation approach which propagates downscaled future weather file uncertainties into building energy use is examined. The four-step approach integrates dynamic building simulation, fitting a distribution to average annual weather variables, regression model (between average annual weather variables and energy use) and random sampling. The impact of fitting different distributions to the weather variable (such as Normal, Beta, Weibull, etc.) and regression models (Multiple Linear and Principal Component Regression) of the uncertainty propagation method on cooling and heating energy use distribution for a sample reference office building is evaluated. Results show selecting a full principal component regression model following a best-fit distribution for each principal component of the weather variables can reduce the variation of the output energy distribution compared to simulated data. The results offer a way of understanding compound building energy use distributions and parsing the uncertain nature of climate projections.



Citation: Yassaghi, H.; Hoque, S. Impact Assessment in the Process of Propagating Climate Change Uncertainties into Building Energy Use. *Energies* **2021**, *14*, 367. <https://doi.org/10.3390/en14020367>

Received: 15 December 2020

Accepted: 7 January 2021

Published: 11 January 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: climate change; building energy use; propagating uncertainties; input distributions; regression model

1. Introduction

In the United States, commercial and residential buildings are major energy use consumers and contribute to carbon emissions. Studies have shown that efficient buildings are interlinked with sustainable built environment [1]. Building energy use is mainly dependent on exterior climate conditions, occupant behavior, envelope characteristics, and equipment efficiency which are also the main sources of uncertainty in quantifying building performance [2]. With the current trends of climate change, there is a need of evaluating buildings that ensure a resilient and sustainable design [3]. Building simulation tools are widely used to assess building energy performance and occupant comfort [4]. However, many require weather files with high spatial and temporal resolution and in the format of a Typical Meteorological Year (TMY). Additionally, due to climate change, existing TMY files (using historical data) should not be used to assess building energy performance for the future [5,6]. On the other hand, Global Climate Models (GCMs) and Regional Climate Models (RCMs) are coarse in resolution. Therefore, they cannot be used directly in building simulation tools, and need to be downscaled to hourly temporal resolutions. This is done using statistical downscaling techniques (e.g., stochastic weather generators) [7] or the morphing process [8]. Downscaling methods are useful to develop weather files that can be directly incorporated in building energy simulations in order to assess building performance under climate change scenarios [9]. However, weather files that are developed from downscaled climate models have many levels of uncertainty [10,11] and most building

simulation tools lack the ability to propagate various sources of uncertainties [12]. For example, the climate is changing and climate models which generate various projections based on emissions scenarios are uncertain in nature and their uncertainties need to be accounted for in building models. In recent years, developments in building standards and regulations have promoted improved building efficiency (e.g., ASHRAE 90.1-2016). Occupant behaviors and dynamic schedules have also been incorporated in building energy tools [13,14]. However, with global temperatures projected to increase by 4.8 °C by the end century (IPCC 2014 [15]), buildings will operate differently from their original design loads [16]. Given these various sources of uncertainty, there is a need to assess building energy performance using probabilistic approaches [17]. One method is the Monte Carlo technique which is commonly used to propagate input variations (e.g., climate, building operation, etc.) through the model to determine uncertainties to the output (e.g., energy consumption) [18,19]. A probabilistic approach to propagate climate uncertainties into building energy use has driven modeling efforts [20,21].

In this regard, fitting distributions to input variables and conducting an uncertainty analysis following random sampling in building simulations is a common practice [22,23]. Sun et al. (2014) looked into the uncertainties of microclimate variables in building energy models and presented a framework for uncertainty quantification which uses a detailed specification of urban form model to quantify microclimate conditions. They compared their results which included quantifying different sources of uncertainties using statistical methods [24]. Gang et al. (2015) examined the uncertainties of nine factors of building parameters, weather and indoor conditions and developed a probability distribution of the cooling energy use and capital cost [25]. In their study they followed three typical distributions of normal, uniform and triangular. De Wilde and Tian (2010) evaluated the impact of the selection of performance metrics and assumptions made in modeling building emissions, overheating and office work performance under climate change using the UKCIP02 climate scenarios following a uniform distribution in Monte Carlo sampling [26]. Wang et al. (2012) investigated uncertainties in energy consumption from different building operational practices and weather data. For the analysis of weather sources uncertainties, they compared the energy results from historical weather data to TMY3 weather files [27]. González et al. (2019) provided an extensive review on critical uncertainty indices [28]. Brohus et al. (2012) conducted a comprehensive uncertainty analysis of input distribution following different distributions [29]. They used a Kolmogorov-Smirnov goodness of fit test was followed as a determination of stochastic sampling.

Regression models are also used to predict building energy use [30] and are useful methods to generate correlations between thermophysical, weather and occupant factors to building output energy use. For example, Braun et al. (2014) used a multiple regression model with regressors of humidity ratio and relative humidity to predict gas and electricity use consumption for future climate periods in northern England [31]. Another study by Catalina et al. (2013) authors develop a regression model between multiple input variables and building energy demand [32]. They found a regression fit to predict heating energy demand to be simple and largely applicable to building simulation.

The process of conducting a Monte Carlo analysis can be complex and would require intense computational needs and recent studies are attempting to reduce the intensity of the uncertainty analysis. Yassaghi, Gurian and Hoque (2020) presented a four-step climate uncertainty propagation method which could reduce computational efforts and is applicable to regions where limited hourly weather data can be developed to evaluate climate change impacts on buildings energy performance [33]. However, probabilities are not always easy to assess. In many cases it is difficult to fit a distribution to the input (the distribution may not be known) or to develop an appropriate correlation between the model input and output, which could then add to the total uncertainties. In addition, the output distributions of energy use from uncertainty analysis can be confound and present a broad scale and therefore, could be difficult to interpret. Studies have attempted to investigate the impact of selecting various input variable distributions on propagating uncertainties to

building energy use following Monte Carlo [25,29]. This paper evaluates the simultaneous impact of selecting appropriate average annual weather variable distribution and regression models between the weather variables and energy use when propagating climate change uncertainties. This is an approach that has not been fully investigated and is not completely understood. The output results of the impact assessment aim to enhance a robust and credible estimate of the uncertainties and interpretations of building energy performance under climate change. The goal is to offer a way of understanding compound building energy use distributions and parsing the uncertain nature of climate projections.

The present work looks into the impact of selecting various distributions for the mean annual weather variables and following different regression models in a four-step uncertainty propagation method. A sample office reference building was selected as the case study for the climate conditions of Philadelphia, PA, USA. Downscaling methods were used to generate hourly weather files based on the Intergovernmental Panel on Climate Change (IPCC) emissions scenarios and develop a data base of current and future TMY files that were incorporated into EnergyPlus. The outcome of this study offers more nuanced estimate of the uncertainties and interpretations of building energy performance under climate change. Section 2 (Methodology) presents the steps taken to conduct the input weather distribution and regression model impact assessment.

2. Methodology

2.1. Study Approach

This study assesses the impact of various average annual weather variables distributions and selecting different regression model in building energy use distributions when conducting an uncertainty propagation analysis. We follow the four-step uncertainty propagation technique presented by Yassaghi, Gurian and Hoque (2020) [33] to conduct our impact assessment analysis. The four-step climate uncertainty propagation method consists of a development of regression model, fitting distribution to weather variables and random sampling (Figure 1).

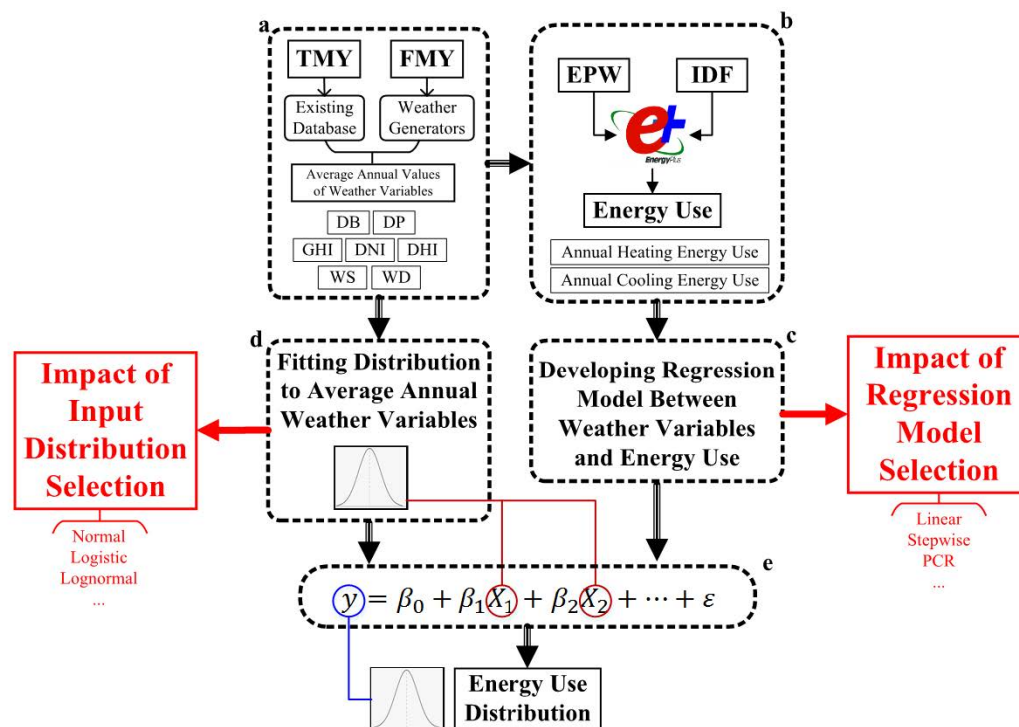


Figure 1. The Four-Step Uncertainty Propagation method followed in the impact assessment (adapted from method proposed by [33]).

In Figure 1, the method shows four major steps. First, multiple building simulations are conducted for all available weather files (current and future). From the simulation results (e.g., for energy use) a multiple linear regression model is fit between the annual energy use results and the annual average weather variables of the TMY files used in the simulation. The process would develop a model between energy outputs and weather variables. Next, a distribution is fit to the average annual weather variables following a Kolmogorov-Smirnov (KS) goodness of fit test and a random sampling is conducted. The best-fit distribution is the distribution with the highest p -value in the Kolmogorov-Smirnov goodness of fit test. Finally, by having a regression model (between the average annual weather variables and energy use) and input average annual weather variables distribution, the uncertainties can be propagated to the energy use distribution. In the uncertainty propagation method presented by [33] the use of different regression models or input average annual weather variable distributions could alter the output energy use distribution results.

To conduct the model and distribution impact assessment (presented in red in Figure 1), we first compare the use of full multiple linear regression (LR) model with a stepwise multiple linear regression (LRS) model in the four-step uncertainty propagation method for heating and cooling energy use distributions. The regression models are conducted between energy use (dependent) and their associated average annual weather variables (regressors) of the weather files used in a dynamic building simulation. Then, we fit various distributions to the average annual weather variables of 46 current and future weather files. The input weather variable distributions are used in the regression model to propagate uncertainties to energy use. Next, we conduct a Principal Component Analysis (PCA) and develop principal components for the average annual weather variables. Then we develop full Principal Component Regression (PCR), stepwise principal component regression (PCRS), and 2-factor principal component regression (PCR2) models between the principal components and heating and cooling energy use. Then various distributions are fit to the principal components which are then used in the principal component regression models to propagate the uncertainties. Results are then compared to assess the impact of the regression model selection and distribution fit on the output energy use of the uncertainty propagation method. Below is a summary of the steps taken:

1. Current and future weather files are created. Current weather files are obtained from existing resource presenting different historical periods. Future weather files are developed using weather generators (Step a in Figure 1).
2. Multiple dynamic building simulations for each weather file is conducted and annual heating and cooling energy use are determined (Step b in Figure 1).
3. The average annual values (arithmetic mean values) for each weather variable of the weather files are calculated and the corresponding principal components of the weather variables are developed.
4. Regression models between the average annual weather variables (and principal components) as regressors and annual heating and cooling energy use (as dependent variables) are developed (Step c in Figure 1).
5. Sample distributions are fit to the average annual weather variables of the weather files data sets and their corresponding principal components (Step d in Figure 1). The sample distribution is repeated for all weather variables and PCs except for the best fit scenario. For example, when the Erlang distribution is selected, all weather variables are assumed to follow an Erlang distribution.
6. Parameters of the distributions are obtained and a random sampling of a sample size of 100 is conducted for all weather variables and principal components.
7. The uncertainties of the input variables (weather variables and their associated principal components) are then propagated to the heating and cooling energy use using the regression model developed (Step e in Figure 1).

Details of the weather files developed and building case study is presented in Section 2.2. Sections 2.3 and 2.4 provide details of the regression models development and distribution fitting process respectively. Results are then presented in Section 3.

2.2. Case Study and Weather Files

Stochastic weather generators such as the AdvancedWEatherGENerator [34–36], the CCWorldWeatherGen [37] which adapts the morphing technique [8], and Meteonorm were used to downscale future weather files based on emissions scenarios. The downscaling methods were used to develop future TMY files that were incorporated into EnergyPlus. Appendix A shows a summary of the weather files, their sources and period of generation used here. In total, 46 weather files were developed (Appendix A), of which 10 are obtained from existing TMY sources and 36 are downscaled future weather files from weather generators and for different emissions scenarios into three future time slices. The weather variables that are used in the development of the assessment model in this study are Dry Bulb Temperature (DB), Dew Point Temperature (DP), Global Horizontal Irradiation (GHI), Direct Normal Irradiation (DNI), Diffuse Horizontal Irradiation (DHI), Wind Direction (WD) and Wind Speed (WS).

The DOE large office reference building in the 4A Philadelphia climate category was analyzed. The office building has 12 floors plus a basement with a total area of 46321.45 m² and window to wall ratio of 40%. The heating and cooling systems used in the building are a boiler and chiller which use gas and electricity respectively as their energy fuel. The walls, roof and windows U-value, lighting load density, window to wall ratio and equipment type following ASHRAE standards 90.1-2004 are presented in Appendix B.

2.3. Selection of Regression Model

A model showing the association between input weather files and output energy use along with input distributions are required to propagate the climate uncertainties. However, developing a suitable regression model can be difficult, or due to lack of data, there could be doubt about the appropriate distribution of the input parameters. In the first part of this study, the model is developed using a full multiple linear regression (LR) between the input weather variables and heating and cooling energy use. We also conduct a stepwise multiple linear regression (LRS) model to understand the most influential weather variables that impact the building energy use.

$$\text{Energy Use} = \beta_0 + \beta_1 \times \text{DB} + \beta_2 \times \text{DP} + \beta_3 \times \text{GHI} + \beta_4 \times \text{DNI} + \beta_5 \times \text{DHI} + \beta_6 \times \text{WD} + \beta_7 \times \text{WS} \quad (1)$$

The regression models are developed for both heating and cooling *energy use* and β'_i are the regression coefficients. In addition, a full, stepwise, and 2-factor principal component regression is also developed for further investigation on the impact of the model selection on the output of the weather uncertainty propagation method (Figure 1).

Principal Component Analysis (PCA) is a statistical analysis that transforms original variables with possible collinearity into new uncorrelated variables called principal components [38]. A Principal Component Regression (PCR) model is a linear regression using the Principal Components (PCs) of the data set as the independent variable. The PCR model has the capability to remove possible collinearity among the independent variable reflected in a linear regression [39]. Three different distributions are considered to analyze how the selection of the input weather distribution impacts the output results of the uncertainty propagation.

$$\text{Energy Use} = \beta'_0 + \beta'_1 \times F1 + \beta'_2 \times F2 + \beta'_3 \times F3 + \beta'_4 \times F4 + \beta'_5 \times F5 + \beta'_6 \times F6 + \beta'_7 \times F7 \quad (2)$$

where $F1, F2, F3, F4, F5, F6$ and $F7$ are principal components corresponding to DB, DP, GHI, DNI, DHI, WS and WD respectively and β'_i are the regression coefficients.

The literature shows a strong linear correlation between building energy use and exterior climate conditions, even though the main weather variables acting as the driving force for energy use can differ depending on the location and case study building. In this study we consider all seven weather variables to be essential to the case study. This is the key reason why we focused on developing linear regression models. If a different regression model were to be used (e.g., non-linear), the results could change but would also be in contradiction to our current physics-based understanding of building performance and exterior climate conditions. On the other hand, we have limited knowledge about the climate and the statistical distribution of weather variables. This is due to lack of sufficient climate data. To address this limitation, the regression models are not developed for a specific time period and all future weather files are used in the process of developing the regression model. Indeed, as our knowledge of our climate and climate change increases, the uncertainties due to the selection of weather variables distribution will likely be reduced. Details of the weather variable distribution selection is given hereafter.

2.4. Selection of Average Annual Weather Variables Distributions

The average annual weather variables and their associated principal components distributions were selected following a Kolmogorov-Smirnov (KS) goodness of fit test [40].

$$D_n = \sqrt{n} \sup_x |F_n(x) - F(x)| \quad (3)$$

In Equation (3), D_n is the test statistic, F is the theoretical cumulative distribution, and F_n is the empirical distribution for n number of observations. The KS test examines the theoretical Cumulative Distribution Function (CDF) difference to the Empirical Cumulative Distribution Function (ECDF) of the sample data and shows whether the sample data follow a certain distribution. The KS test was selected since it does not require normality assumptions and can be used when small sample sizes are available. The test ranks multiple distributions which can have a potential fit to the data (based on their p -value). When the distributions are fit, a random sampling was conducted to the data for a sample size of 100 and results are then incorporated into the regression models to propagate the uncertainties of the weather files to the building energy use.

Results of the output energy use distributions developed for each regression model case and following each sample input distributions are analyzed. We rely on scatter plots, box and whisker plots and mean absolute error values to assess the results.

3. Results

3.1. Summary of Current and Future Weather Data & Energy Use

Climate models are inherently uncertain. We can, however, generate snapshots of what it might look like. In addition, the process of propagating the climate uncertainties is not absolute and comes with uncertainties with regard to the selection of appropriate parameters and models. Figure 2 reflects the dry bulb temperature (DB) for each weather file to reflect the impact of climate change on the current and future weather files and shows the results of heating (HeatingGas) and cooling (CoolingElec) energy use of the sample office building when EnergyPlus simulation is conducted for each weather file.

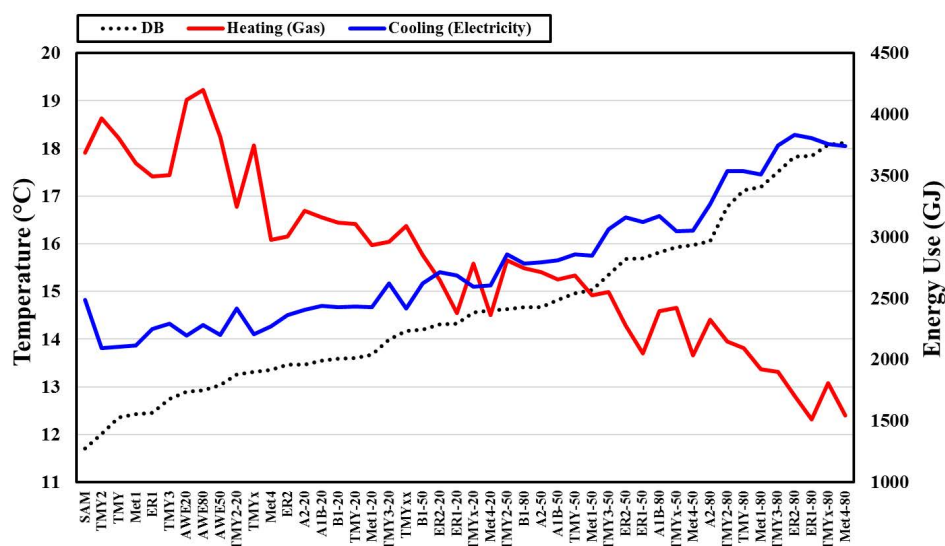


Figure 2. Results of dry bulb temperature and direct normal irradiation for each of the current and future TMY files and their associated heating and cooling energy use from EnergyPlus modeling of the case study.

The horizontal axis shows the weather files (current and future) used in this study and are sorted by temperature in an ascending order. Details of the weather files are presented in Appendices A and C. Air temperature in many building cases is predictably the most influential driving force for heating and cooling energy use. Although this could differ depending on the location of study and building. As temperature increases, heating energy consumption decreases and cooling energy use increases (Figure 2). Using the morphed TMY3 weather file, adapted with IPCC A2 scenario (which is the second most extreme scenario), there is an increase of 1454.9 GJ in cooling energy use by the end of the century compared to the current available TMY3 file for Philadelphia. Other variables that also impact energy performance and are considered in this study are direct normal irradiation, diffuse horizontal irradiation, global horizontal irradiation, dew point and wind speed/direction. The main weather variables influencing a building's energy use can vary from building to building and depending on the building's physical characteristics, ventilation type, location and purpose of use may change.

3.2. Developed Regression Models

Multiple regression models are developed after conducting dynamic simulations on the case study and for all weather file data base. The regression coefficient values of the average annual weather variables of the weather files and their corresponding principal components are presented in Table 1. The regression models are between the annual heating and cooling energy use and the average annual weather variables (and their associated principal components for PCR). One set of the regression models are full and stepwise multiple linear regression between annual heating and cooling energy use and the average annual weather variables of the weather files used in the simulations. Another set of the regression models are full, stepwise and 2-factor principal component regression models between annual heating and cooling energy use and the principal components associated to the average annual weather variables.

Table 1. Details of the regression coefficient values for heating and cooling energy use and following different regression models.

Cooling				Heating			
Variable	Coeff.	LR	LRS		LR	LRS	
Intercept	β_0	−294.329	−1260.9		9727.159	11147.75	
DB	β_1	202.714	211.805		−396.019	−344.807	
DP	β_2	145.703	121.489		58.934	0	
GHI	β_3	−11.689	0		16.153	0	
DNI	β_4	4.994	2.136		−11.924	−9.17	
DHI	β_5	12.34	0		−48.3	−24.446	
WD	β_6	−1.784	0		3.671	0	
WS	β_7	−88.557	−75.335		114.733	0	
Variable	Coeff.	PCR	PCRS	PCR2	PCR	PCRS	PCR2
Intercept	β'_0	2782.435	2782.435	2782.435	2796.132	2796.132	2796.132
F1	β'_1	−78.034	−78.034	−78.034	101.503	101.503	101.503
F2	β'_2	344.047	344.047	344.047	−430.372	−430.372	−430.372
F3	β'_3	−50.977	−50.977	0	−114.421	−114.421	0
F4	β'_4	95.46	95.46	0	−206.575	−206.575	0
F5	β'_5	31.77	0	0	−41.26	0	0
F6	β'_6	−185.519	−185.519	0	635.965	635.965	0
F7	β'_7	120.31	120.31	0	−157.836	0	0

Appendix E shows a more detailed summary of the regression models coefficient factors errors and p -values.

3.3. Input Variables Distributions

Results of the KS statistic test for development of the input weather variable and their corresponding principal components distribution are summarized in Appendix F. For the multiple linear regression models the distributions of Normal, Lognormal, Logistic, Gamma, Weibull, Fisher-Tippett, Erlang and BestFit were selected for the average annual weather variables. The BestFit is the selection of best distribution fit for each weather variables separately based on the highest p -value in the Kolmogorov-Smirnov goodness of fit test which then would be used in the regression model to propagate the uncertainties to the energy use. Appendix G provides details of the best fit characteristics. For the PCRs we fit distributions of Beta, Fisher-Tippett, GEV, Logistic, Normal and BestFit to the principal components associated to the average annual weather variables. Figure 3 shows summary of the input distributions selected for the impact assessment analysis. The distribution selection of the weather variables for the multiple linear regression model and the principal component regression model are slightly different. This is due to the parameter fitting test that produced distributions (with relatively reasonable p -values) that differed between the weather variables and their corresponding principal components.

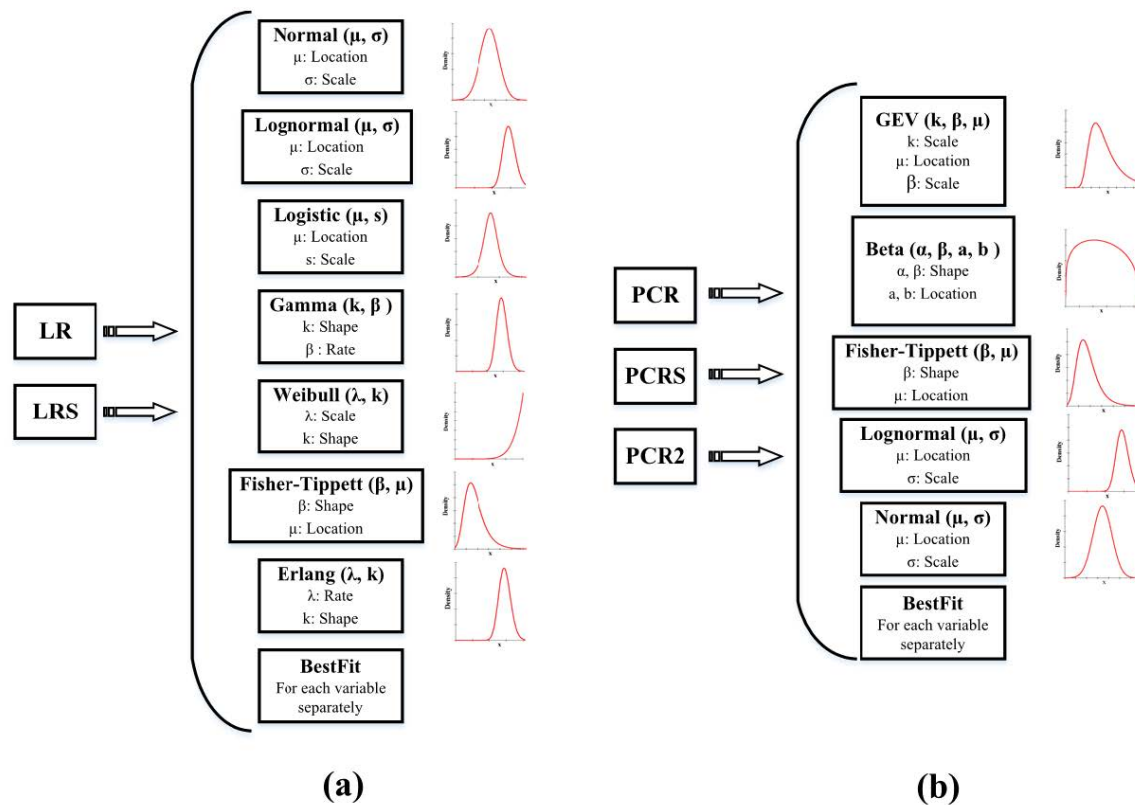


Figure 3. Summary of the average annual weather variables distributions selected for use in LR and LRS regression models (a) and their associated principal components distributions selected for use in PCR, PCRS and PCR2 regression models (b) of the impact assessment analysis.

3.4. Impact of Regression Model Selection on Energy Use

The heating and cooling energy use prediction developed from the regression models of interest (LR, LRS, PCR, PCRS and PCR2) are presented in Figure 4. The standardized coefficient factors of the regression models are summarized in Appendix D which also shows which variables were determined to be statistically less influential on the heating and cooling energy use in the stepwise regression models and therefore were left out of the regression model. Details of the regression factors values are presented in Appendix E. A scatter plot of the predicted energy use and actual energy use is given in Figure 4.

As it can be seen from Figure 4, the predicted cooling energy use of the regression models shows a relatively closer fit to the actual cooling energy use compared to predicted heating energy use vs actual heating energy use. In general, most regression models for heating and cooling energy use show a relatively close prediction to the actual data except for the 2-factor principal component regression model. The regression models developed in this step are used to propagate the uncertainties of the input parameters to the building energy use.

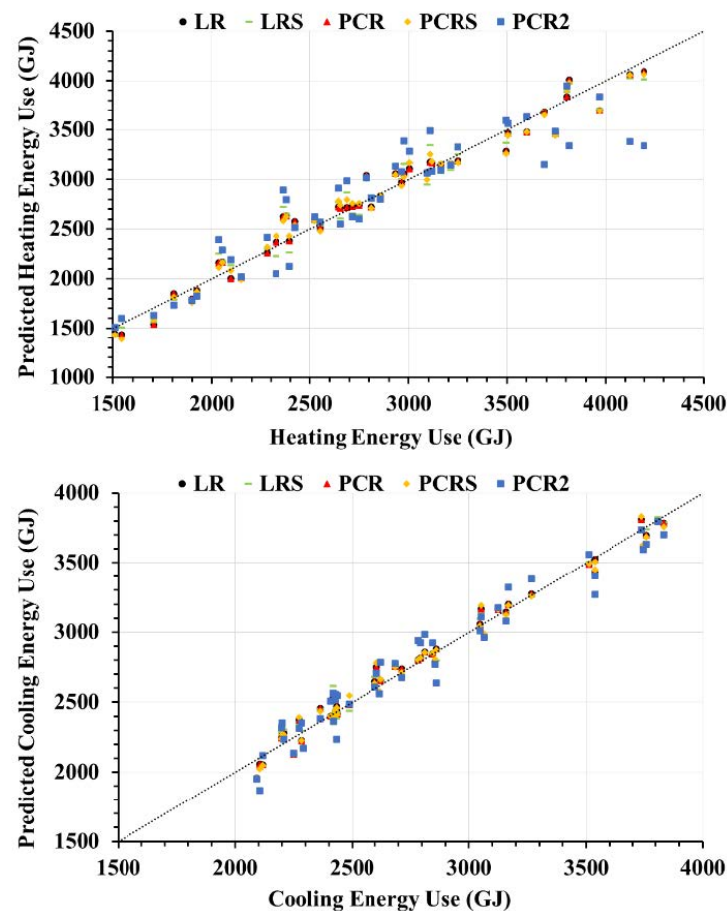


Figure 4. Results of the regression models of prediction for heating and cooling energy use over the simulated data.

3.5. Impact Assessment of Regression Model and Input Average Annual Variables Distributions on Energy Use Distributions

The sample distributions for the weather variables and their corresponding principal components are developed and random samples (100 samples per each variable) are generated and used in the regression models to propagate their uncertainties to the energy use. A summary of the parameters of each distribution is presented in Appendix F. We rely on box plots and mean absolute errors to assess the propagated energy use following selection of various regression models and input distributions of the uncertainty propagation method. Outliers were excluded from the box plots. Figure 5 is a consolidation of the graphs and is presented as a guideline to interpret the box plots. Note that the name of the distributions at the bottom of the graph represents the input variables distribution series selected to obtain the energy use probability density function in the uncertainty propagation method.

In the assessments hereafter we define the output energy distributions which were obtained following different input weather variable distributions in the uncertainty propagation method as “series”. For example, Normal distribution series for heating energy use refers to the distribution of heating energy use obtained from the uncertainty propagation method when normal distribution was fit to the average annual weather variables.

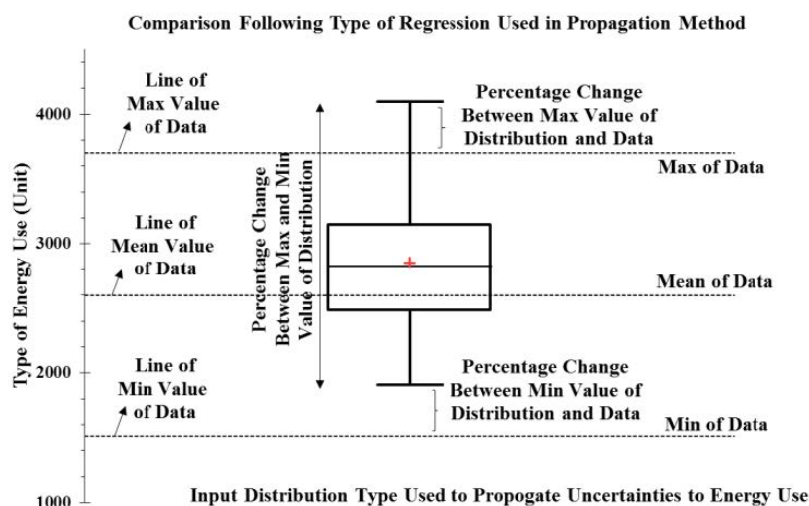


Figure 5. Legend of the figures to interpret output of the impact assessment.

3.5.1. Impact Assessment on Heating Energy Use Distribution

Figures 6 and 7 show the results of heating distributions following full multiple linear regression and stepwise multiple linear regression respectively and compare the selection of different input distributions in the uncertainty propagation method.

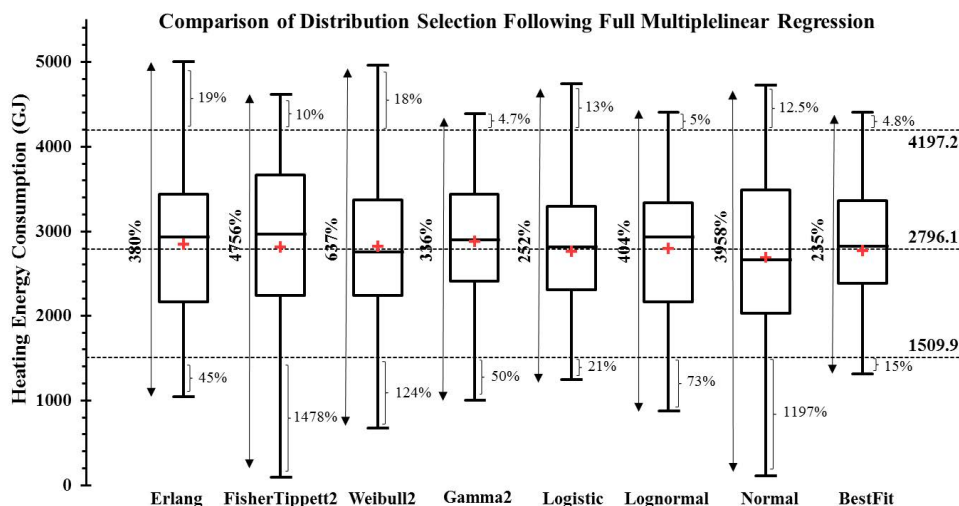


Figure 6. Heating energy use probability density functions when following a full multiple linear regression model and comparing selection of various input average annual weather variables distributions used in the uncertainty propagation method.

The values of 4197.2, 2796.1 and 1509.9 and their representative line in Figures 6 and 7 show the maximum, mean and minimum values of the EnergyPlus simulation output, respectively. From Figure 6, following a full multiple linear regression, the selection of different input distributions can result in high differences in the lower and upper bound of the output heating energy use distribution. In some cases (Normal, Weibull and Fisher-Tippet), the lower bound showed more than 100% difference compared to the minimum of the data obtained from EnergyPlus simulation. However, the average for all input distribution selections show to be close to the average of the simulated data. The heating distribution result following the BestFit distribution of the input values (BestFit series) showed smaller variations from the higher and lower bound of the simulated heating

energy use. In addition, among the distribution series, BestFit shows a smaller difference between the maximum and minimum value of the heating energy use distribution.

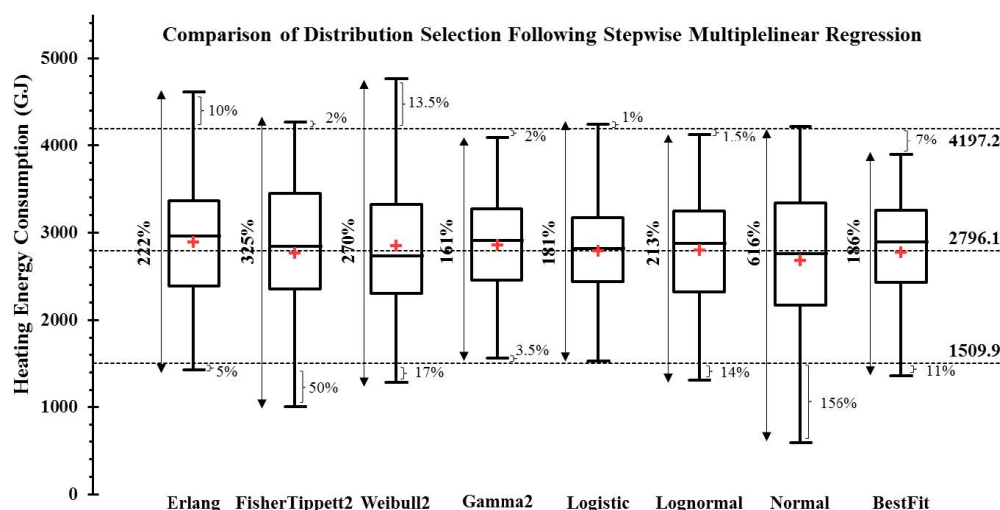


Figure 7. Heating energy use probability density functions when following a stepwise multiple linear regression model and comparing selection of various input average annual weather variables distributions used in the uncertainty propagation method.

Figure 7 shows the results of heating energy use when the stepwise multiple linear regression was selected in the uncertainty propagation method and compares the heating energy distributions obtained following various input distributions for the weather variables.

As it can be seen from Figure 7, the average values for each input distribution selection are similar to the average of the simulated data. However, the lower and upper bound for some distributions show relatively high difference to the simulated data. Yet the difference in the higher and lower bound of the heating energy use data and the simulated data are generally smaller when following a stepwise linear regression (Figure 7) compared to a full linear regression (Figure 6). In Figure 7, the heating energy distribution following the BestFit input distribution has relatively closer fit to the higher and lower bounds of the simulated data (compared to Erlang, FisherTippett, Weibull and Normal series) and shows a lower percentage change between the max and min of the distribution. However, Logistic, Gamma and Lognormal series showed closer values to the higher and lower bounds and smaller percentage difference between max and min of distribution compared to the BestFit. When the difference between higher and lower bound of a distribution are lower it could enhance the credibility of the output distribution when interpreting the data. This can also be seen in Table 2, which shows the mean absolute error for each series and following LR and LRS.

Table 2. Results of the mean absolute errors of heating energy use probability density functions stemming from each input average annual weather variable distribution series for LR and LRS regression models.

Heating	Erlang	FisherTippett2	Weibull2	Gamma2	Logistic	Lognormal	Normal	BestFit	Sum	Avg.
LR	720.2	818.4	687.8	612.9	648.8	665.2	823.0	672.8	5649.0	706.1
LRS	567.3	610.3	581.8	477.3	516.7	515.1	650.8	557.9	4477.3	559.7

As it can be seen from Table 2, the sum and average value of the mean absolute error of all the distribution series following a stepwise multiple linear regression showed lower values compared to the full multiple linear regression. This could reflect that following a stepwise regression model in the uncertainty propagation method could reduce the uncertainties in the selection of weather variable input distribution and reduce the variation of the output results between the lower and higher bound. This can be explained

by the fact that stepwise regression model has fewer weather variables compared to a full regression model and as a result fewer uncertainty associated to the variables are propagated to the output energy use. It should be noted that both the stepwise and full regression model had shown to be a good fit to the data.

Figures 8–10 show the heating energy use distribution when following a PCR, a stepwise PCR and a 2 factor PCR when selecting various input distributions of principal component variables in the uncertainty propagation method respectively.

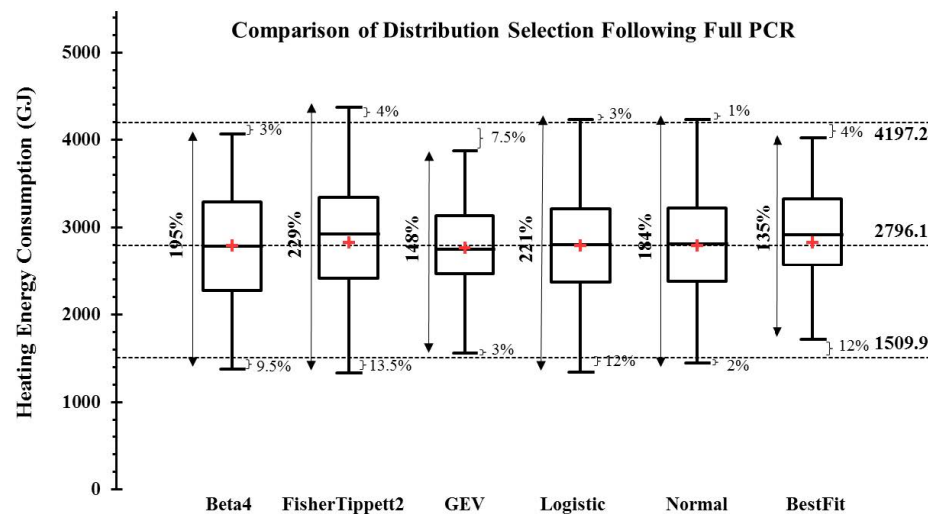


Figure 8. Heating energy use probability density functions when following a Full Principal Component Regression model and comparing selection of various input principal component distribution series used in the uncertainty propagation method.

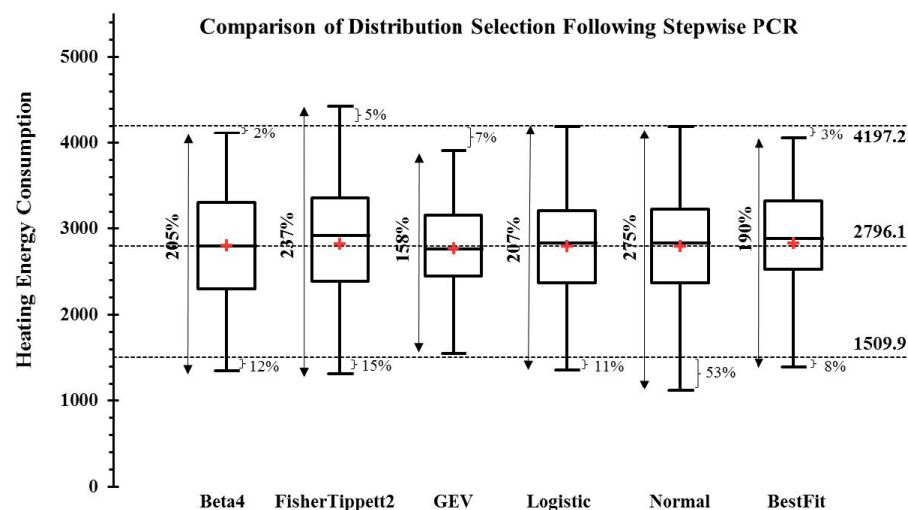


Figure 9. Heating energy use probability density functions when following a Stepwise Principal Component Regression model and comparing selection of various input principal component distribution series used in the uncertainty propagation method.

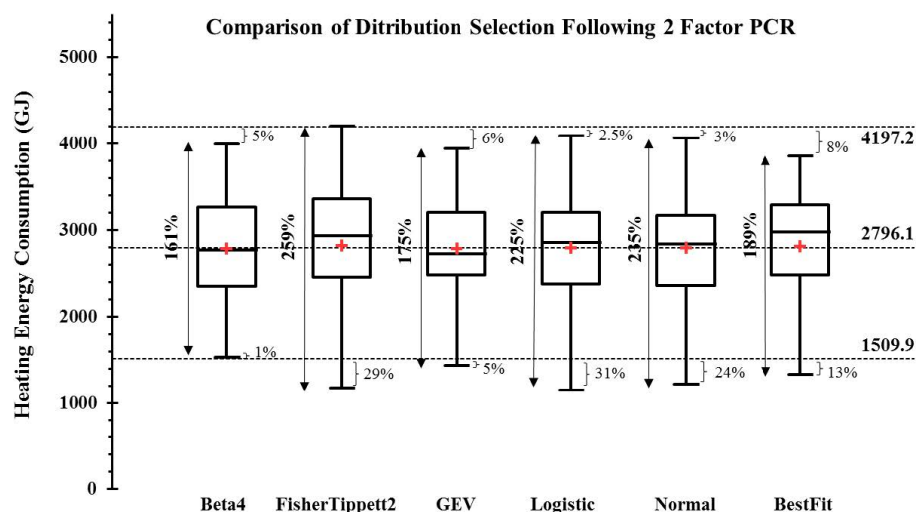


Figure 10. Heating energy use probability density functions when following a 2 Factor Principal Component Regression model and comparing selection of First and Second input principal component distribution series.

In Figures 8–10 for all distribution series the average values show a relatively good fit to the mean simulated data. As it can be seen from Figure 8, the series show relatively smaller variations between maximum and minimum of the heating energy use distribution compared to Figure 6 which followed a full multiple linear regression model on the weather variables. Although the same distribution to input weather variables were not fitted to the principal components but by comparing the similar ones (FisherTippett, Normal, Logistic and BestFit) from Figures 6 and 7 it is clear that not only the difference between the higher and lower bound to the simulated data are smaller in Figures 8–10 but the percentage change of the max and min of the output distribution is also small. This is also evident in the sum and average of the mean absolute errors of the principal component regression models approach to multiple linear regression models approach in Table 3.

Table 3. Results of the mean absolute errors of heating energy use probability density functions stemming from each input principal component variable distribution series for PCR, PCRS and PCR2 regression models.

Heating	Beta4	FisherTippett2	GEV	Logistic	Normal	BestFit	Sum	Average
PCR	538.4	579.8	452.2	532.3	513.0	536.2	3152.0	525.3
PCRS	541.2	584.1	456.2	530.7	511.7	529.8	3153.6	525.6
PCR2	504.3	540.8	444.8	503.0	494.8	517.1	3004.9	500.8

From Table 3, it can also be seen that following a two-factor principal component regression model in the uncertainty propagation showed lower average and sum of mean absolute errors of all series and compared to all other Figures. In Figure 8, the BestFit series showed smallest percentage change between max and min of the distributions compare to all other series and relatively smaller difference to most other series in Figures 9 and 10 which was also seen in Figure 6 and for most series of Figure 7.

3.5.2. Impact Assessment on Cooling Energy Use Distribution

Figures 11 and 12 shows the results of cooling energy use when propagating uncertainties following the regression models on various input weather file distribution.

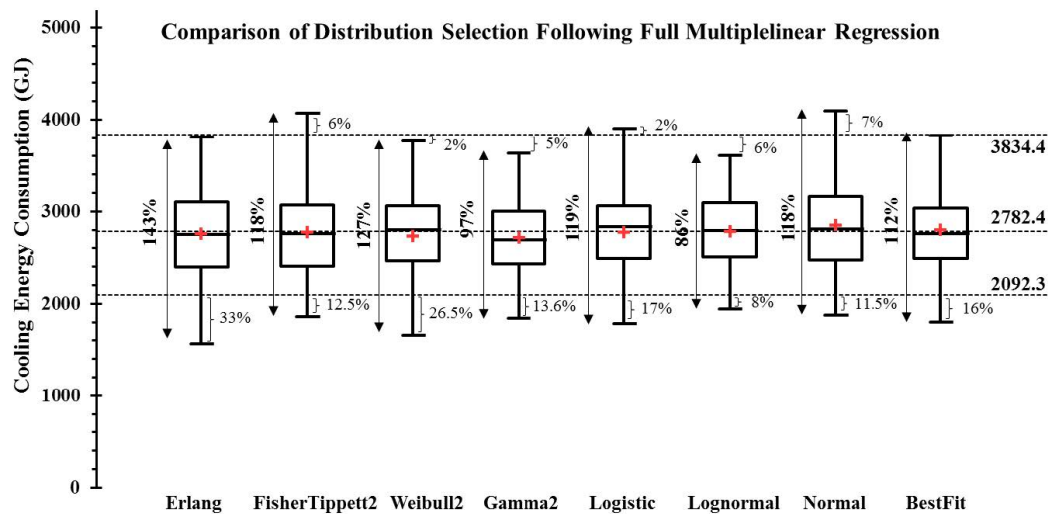


Figure 11. Cooling energy use probability density functions when following a full multiple linear regression model and comparing selection of various input average annual weather variables distributions used in the uncertainty propagation method.

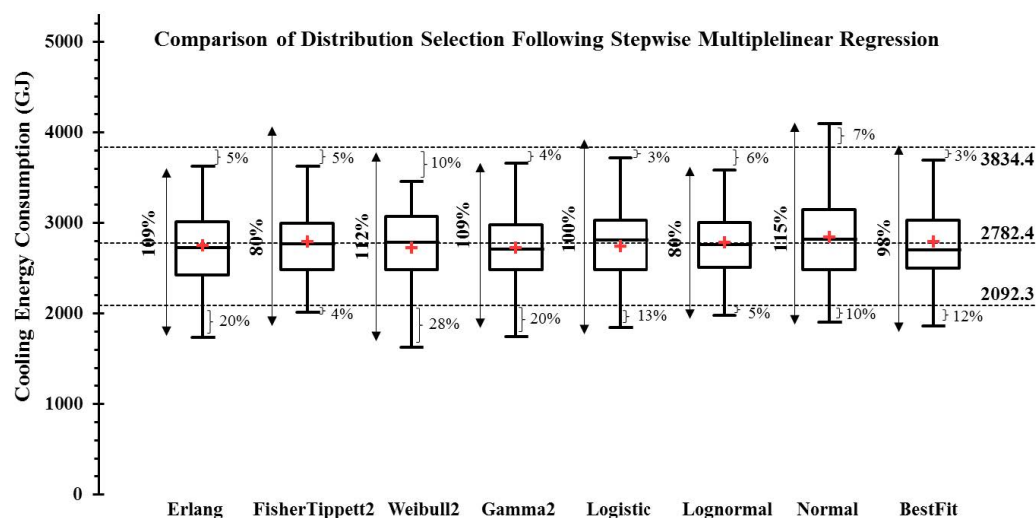


Figure 12. Cooling energy use probability density functions when following a stepwise multiple linear regression model and comparing selection of various input average annual weather variables distributions used in the uncertainty propagation method.

From Figures 11 and 12 the average cooling energy use distributions show a close fit to the mean simulated data. In addition, the percentage changes between the max and min of the cooling energy distributions are relatively smaller than what was shown for heating energy consumption (Figures 6 and 7). This is also apparent in the average and sum of all mean absolute errors of the series for both regression models shown in Table 4. Figures 11 and 12 smaller variations can be seen in the lower and upper bound of the cooling energy use distribution from actual simulated data compared to heating energy use results presented in Figures 6 and 7. The BestFit distribution, compared to most series of both regression models show relatively smaller percentage change of max and min of the distribution and smaller variations between the lower and higher bound compared to the simulated data.

Table 4. Results of the mean absolute errors of cooling energy use probability density functions stemming from each input weather variable distribution series for LR and LRS regression models.

Cooling	Erlang	FisherTippett2	Weibull2	Gamma2	Logistic	Lognormal	Normal	BestFit	Sum	Avg.
LR	390.3	405.3	355.4	306.8	365.0	321.2	402.1	355.4	2901.5	362.7
LRS	340.0	339.0	348.4	275.0	338.2	277.6	365.3	352.0	2635.5	329.4

Figures 13–15 show the results of cooling energy use when propagating uncertainties following the principal component regression models on various input principal components distributions.

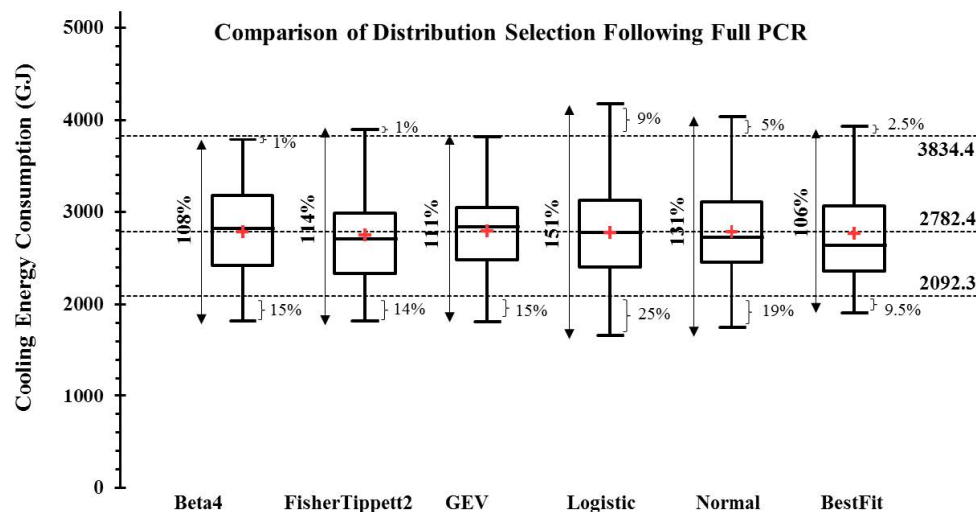


Figure 13. Cooling energy use probability density functions when following a Full Principal Component Regression model and comparing selection of various input principal component distribution series used in the uncertainty propagation method.

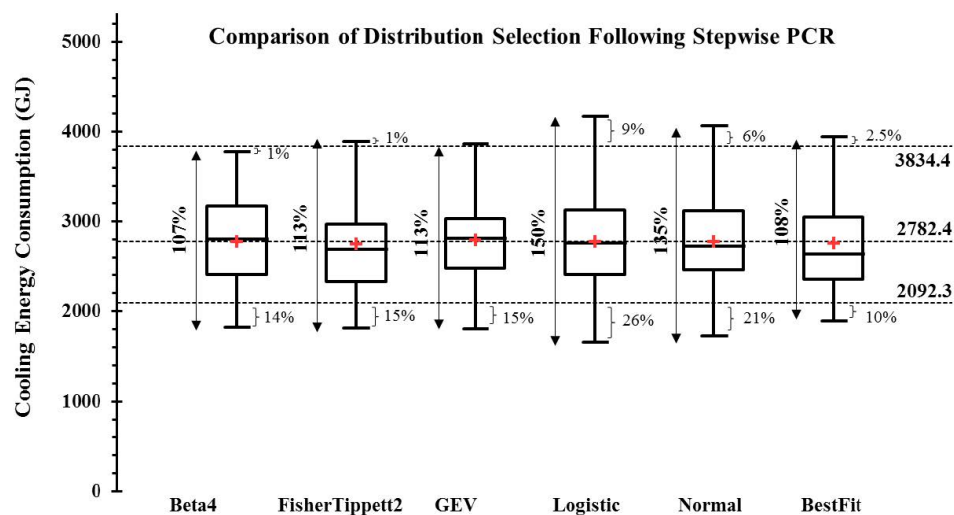


Figure 14. Cooling energy use probability density functions when following a Stepwise Principal Component Regression model and comparing selection of various input principal component distribution series used in the uncertainty propagation method.

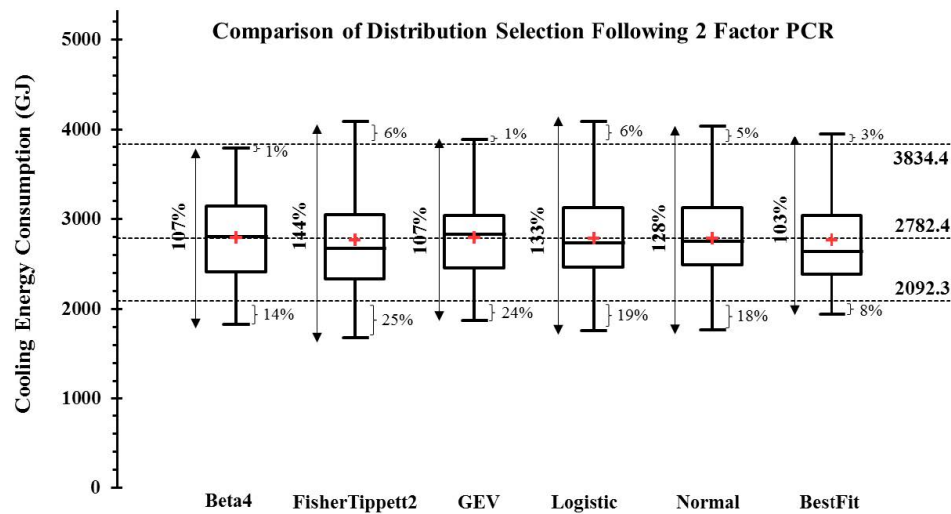


Figure 15. Cooling energy use probability density functions when following a 2 Factor Principal Component Regression model and comparing selection of First and Second input principal component distribution series used in the uncertainty propagation method.

Similar to previous Assessments (Figures 6–12), the results of the cooling energy use distributions for all principal component regression model cases and all distribution series show a close fit to the average and relatively low variations in lower and higher bound of distributions compared to simulated data (Figures 13–15). The BestFit series for all regression cases show lowest percentage change between min and max of the distribution compared to other series of distributions. The sum of the mean absolute error of the series for the principal component regression cases (Figures 13–15) show lower values compared to the linear regression models (Figures 11 and 12) but the average of the mean absolute errors of the series were higher (Table 5).

Table 5. Results of the mean absolute errors of cooling energy use probability density functions stemming from each input principal component variable distribution series for PCR, PCRS and PCR2 regression models.

Cooling	Beta4	FisherTippett2	GEV	Logistic	Normal	BestFit	Sum	Average
PCR	417.0	420.3	364.8	408.1	415.9	426.2	2452.4	408.7
PCRS	418.4	420.5	363.5	410.0	413.7	425.7	2451.8	408.6
PCR2	402.1	429.5	354.4	401.1	394.2	412.2	2393.6	398.9

4. Discussion

In general, the variations in the cooling energy use distributions were relatively lower compared to heating energy use distributions for all regression cases and all distribution series. One reason for the small variations among the distributions for cooling energy use and following the full and stepwise multiple linear regression can be explain by the regression fit in the scatter plots of Figure 4 which shows the regression models had a relatively better fit to the cooling energy use compared to heating energy use. In addition, when propagating input uncertainties into building energy use following a regression model and a parameter fitting process, the selection of regression model can significantly impact the uncertainties of the higher and lower bound of the output energy use when following different input distributions and could change the variations in percentage change of the maximum and minimum of the energy use distribution.

The use of principal component regression model also showed small variation between lower and higher bounds of the results especially for heating energy use. The drawback of the principal component regression model is that principal components cannot be attributed to any specific weather variable. In other words, we cannot determine which

weather variables have the highest impact on the energy results. In addition, unlike a multiple linear regression model where the developed correlation can be used for other cases; here, the principal component regression model is only valid for the case under examination.

5. Conclusions

Given the current trends of climate change, a probabilistic approach to assessing building energy performance is necessary. Propagating the climate uncertainties into building energy performance would result in a building energy use distribution into the future. This offers the opportunity to assess the risk of variations in building cooling and heating energy use during the design phase. However, understanding the new uncertainties introduced during the uncertainty propagation method is of high importance. For example, the process of conducting an uncertainty propagation could introduce new uncertainties and without their consideration could yield unreliable energy results. We may not have a complete understanding of weather variables distributions, and in many cases fitting a proper regression model to the weather variables and energy use is a challenge. In this study the simultaneous impact of selecting appropriate average annual weather variable distribution and regression model between the weather variables and energy use when propagating climate change uncertainties has been investigated.

To propagate climate uncertainties into building energy use we followed a four-step propagation method which consists of a regression model development, input distribution fitting and random generation. We then assessed the impact of the selection of input average annual weather variable distributions and regression models on the heating and cooling energy distribution. The impact assessment addresses the uncertainties stemming from the selection of input parameters and the development of regression models.

Results show regardless of regression model selection or the input variables distributions, the average energy use showed small variations across the series and relatively small difference to the actual simulated data. This would mean we could have higher confidence in interpreting the mean values when considering climate change uncertainties in building energy use. However, in many cases, extreme conditions are of interest (higher and lower bounds) and finding the appropriate distribution to fit the input variables is not trivial. Our findings show the selection of the BestFit distributions for weather variables and principal components, for most cases could significantly reduce the higher and lower bound variations of the heating and cooling energy use distributions compared the simulated data. In addition, the BestFit distribution showed a relatively smaller percentage change in maximum and minimum of the heating and cooling energy use compared to most distribution series. Although in some cases for heating energy use distribution, the logistic distribution series showed to be a better fit. Our findings suggest that when the modeler is uncertain about input distributions, selecting the appropriate regression model can reduce the variations between high and low bound of output energy. For heating energy use distributions, results showed that a two-factor principal component regression model would result in less average and sum of mean absolute error for the distribution series of weather variables and their corresponding principal components respectively. For cooling energy use distribution however, results show following a stepwise multiple linear regression yields a lower average mean absolute error for all input distribution series but following a two-factor principal component in the uncertainty propagation method results in a lower sum of mean absolute errors of the output cooling energy use distribution series. The purpose of presenting the results are to show the impact the selection of different regression models and input distribution fit on the output energy distribution. The selection of the most appropriate combination can be very complex and would require a large amount of input data. Furthermore, it would necessitate developing simulations with high certainty, which is difficult. The results presented in this study offer a way of understanding how to parse the uncertain nature of climate projections and the limited number of future hourly weather generator methods available.

Author Contributions: Conceptualization, H.Y. and S.H.; methodology, H.Y.; validation, H.Y. and S.H.; formal analysis, H.Y.; investigation, H.Y.; resources, H.Y. and S.H.; data curation, H.Y.; writing—original draft preparation, H.Y.; writing—review and editing, S.H.; visualization, H.Y.; supervision, S.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not Applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data will be available upon reasonable requests.

Acknowledgments: Authors are thankful to Jin Wen and Patrick Gurian for their technical guidance in this study.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. Summary of current weather file sources and future weather files method of generation.

Current Weather Files			Future Weather Files	
File	Source	Period	Method	Scenario(s)
TMY	National Solar Radiation Data Base	1952–1975	Advanced Weather Generator (AWE-GEN)	-
TMY2	National Solar Radiation Data Base	1961–1990	Morphed TMY	A2
TMY3	National Solar Radiation Data Base	1991–2005	Morphed TMY2	A2
TMYx	Climate.OneBuilding.Org	2003–2017	Morphed TMY3	A2
TMYxx	Climate.OneBuilding.Org	2004–2018	Meteonorm	A1B, B1, A2
ER1	European Commission of Energy Efficiency Research	2004–2014	Morphed Met1	A2
ER2	European Commission of Energy Efficiency Research	2005–2015	Morphed Met4	A2
Met1	Meteonorm Base	1961–1990	Morphed TMYx	A2
Met4	Meteonorm Base	2000–2009	Morphed ER1	A2
SAM	System Advisor Model	2010	Morphed ER2	A2

Appendix B

Table A2. Summary of the sample office reference building.

Parameter		Unit	Large Office
Wall	U-factor	W/m ² -K	0.857
Roof	Type	-	IEAD ¹
	U-factor	W/m ² -K	0.365
Floor	Type	-	Slab
	U-factor	W/m ² -K	2.193
Window	U-value	W/m ² -K	3.241
	WWR ²	%	40.00
	SHGC ³	-	0.385
Infiltration	-	m ³ /s-m ²	0.302 × 10 ⁻⁴
Outdoor Air	Requirement	m ³ /s-person	0.0125
Occupancy	Density	m ² /person	18.58
PLD ² —LPD ³	Interior	W/m ²	10.76
Fan	-	Control Type	VAV ⁴
	Efficiency	%	60
Heating	Type	-	Boiler
	Efficiency	%	78
Cooling	Type	-	Chiller
	Efficiency	COP	5.5

¹ Insulation Entirely Above Deck ² Light Power Density ³ Plug Load Density ⁴ Variable Air Volume with reheat.

Appendix C

Table A3. Details of the current and future weather files and their associated principal components.

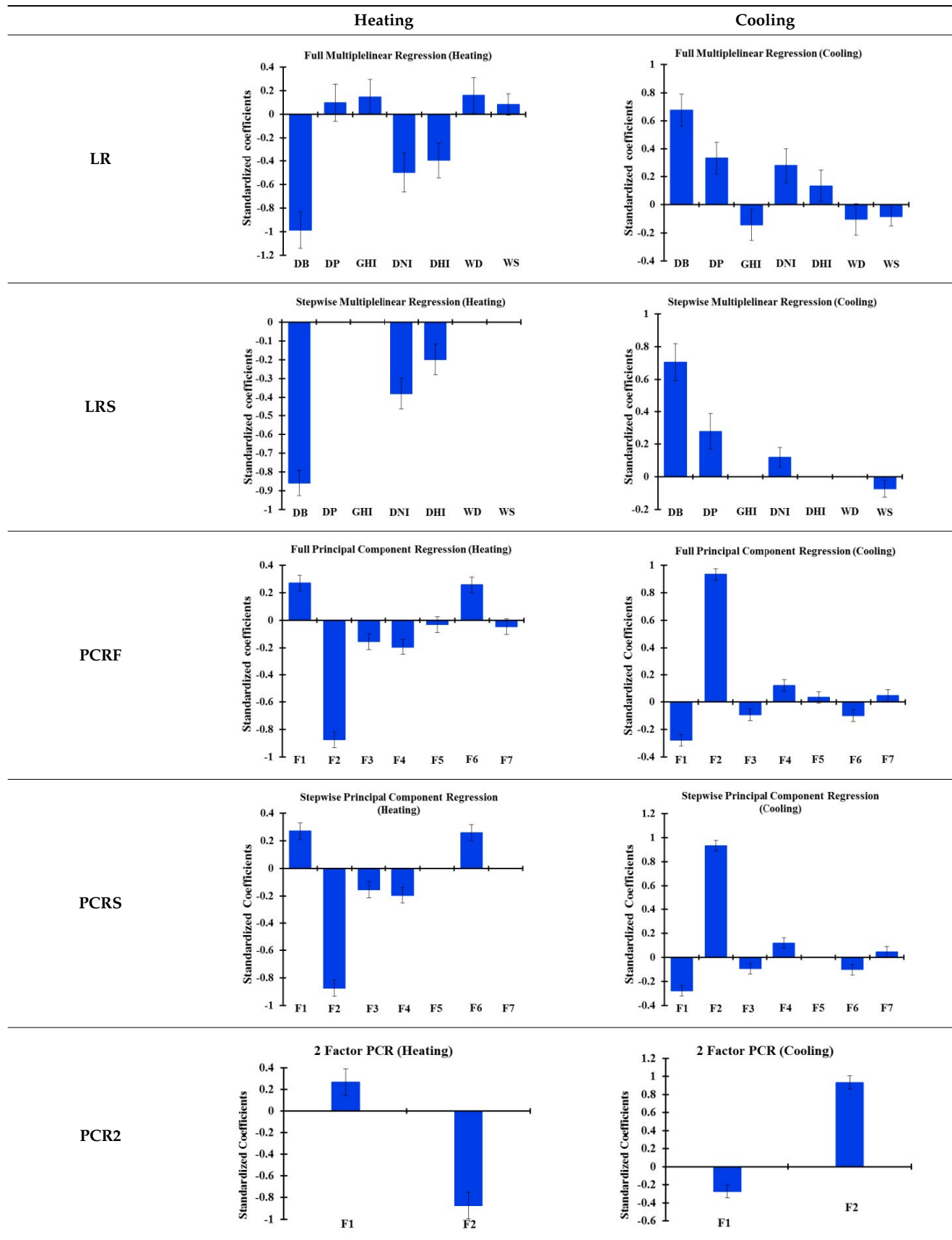
Scenario	Weather Variables							Principal Components							Energy Use	
	DB	DP	GHI	DNI	DHI	WD	WS	F1	F2	F3	F4	F5	F6	F7	Heat.	Cool.
A1B-20	13.5	6.8	171.9	164.1	78.0	214.4	4.8	0.96	−0.47	1.22	−0.32	0.11	0.21	−0.02	3161.5	2436.4
A1B-50	14.8	8.1	172.3	166.1	77.9	215.7	4.8	0.79	0.76	0.93	−0.34	0.12	0.36	0.14	2653.8	2810.8
A1B-80	15.8	9.0	172.7	166.9	77.5	214.4	4.8	0.60	1.70	0.66	−0.39	0.19	0.47	0.25	2397.3	3169.5
A2-20	13.5	6.8	171.2	162.1	78.8	214.4	4.8	1.11	−0.54	1.16	−0.32	0.06	0.14	0.04	3212.6	2407.8
A2-50	14.7	7.9	171.9	163.4	79.1	215.7	4.8	0.98	0.63	0.93	−0.38	0.01	0.25	0.17	2714.9	2794.3
A2-80	16.1	9.3	171.4	167.8	75.5	214.4	4.8	0.48	1.85	0.45	−0.21	0.49	0.61	0.26	2324.0	3266.7
AWE20	12.9	7.7	166.4	91.6	74.8	147.2	4.1	0.59	−1.22	−2.25	−1.70	0.08	0.10	−0.06	4121.6	2197.0
AWE50	13.0	7.7	167.7	92.9	74.8	147.2	4.0	0.41	−1.16	−2.21	−1.70	−0.07	0.06	−0.16	3815.0	2199.3
AWE80	12.9	7.7	168.5	94.5	74.0	147.2	4.1	0.33	−1.18	−2.07	−1.78	0.07	0.19	−0.21	4197.2	2283.9
B1-20	13.6	6.9	171.2	165.3	76.9	214.4	4.8	0.88	−0.46	1.14	−0.21	0.27	0.29	−0.01	3116.8	2428.6
B1-50	14.2	7.5	173.3	162.3	81.1	215.7	4.8	1.11	0.26	1.21	−0.57	−0.30	0.09	0.11	2855.1	2622.4
B1-80	14.7	7.9	173.2	166.5	78.4	214.4	4.8	0.76	0.63	1.05	−0.44	0.02	0.30	0.11	2747.6	2784.8

Table A3. Conts.

Scenario	Weather Variables							Principal Components							Energy Use	
	DB	DP	GHI	DNI	DHI	WD	WS	F1	F2	F3	F4	F5	F6	F7	Heat.	Cool.
ER1	12.5	5.2	179.2	195.7	69.1	147.1	3.7	−2.57	−2.46	0.88	−0.32	−0.15	−0.47	−0.12	3496.6	2249.6
ER1-20	14.3	6.7	181.7	211.3	66.1	147.6	3.6	−3.44	−0.80	0.75	−0.14	0.04	−0.17	−0.03	2381.4	2685.3
ER1-50	15.7	7.7	182.9	213.5	65.4	147.6	3.6	−3.78	0.28	0.50	−0.11	0.06	−0.15	−0.05	2051.0	3124.8
ER1-80	17.8	9.2	184.5	216.4	64.3	147.6	3.5	−4.26	1.98	0.12	−0.09	0.16	−0.13	−0.09	1509.9	3807.8
ER2	13.5	6.2	173.4	188.6	65.8	145.8	3.9	−2.33	−1.68	0.04	−0.05	0.74	−0.20	0.16	3003.6	2362.7
ER2-20	14.3	6.7	176.9	202.4	64.7	140.3	3.8	−3.06	−0.99	0.27	−0.13	0.73	−0.20	0.19	2645.9	2714.2
ER2-50	15.7	7.6	178.0	204.3	64.0	140.3	3.7	−3.39	0.09	0.01	−0.10	0.75	−0.18	0.16	2278.9	3158.8
ER2-80	17.8	9.1	179.6	207.4	62.8	140.3	3.7	−3.88	1.79	−0.37	−0.07	0.85	−0.16	0.12	1705.3	3834.4
Met1	12.4	5.7	169.0	157.5	79.9	215.7	4.8	1.54	−1.57	1.27	−0.20	0.02	−0.07	0.01	3601.9	2117.9
Met1-20	13.7	6.8	171.0	154.2	81.0	215.7	4.7	1.42	−0.45	1.05	−0.44	−0.20	−0.09	−0.02	2933.4	2428.5
Met1-50	15.0	7.8	172.1	156.6	80.1	215.7	4.7	1.05	0.65	0.77	−0.39	−0.19	−0.05	−0.04	2526.1	2845.7
Met1-80	17.2	9.3	173.7	160.2	78.7	215.7	4.6	0.52	2.37	0.37	−0.34	−0.08	0.02	−0.07	1923.8	3510.9
Met4	13.4	6.2	167.6	156.4	79.9	215.7	4.2	1.07	−1.12	0.44	0.42	−0.60	−0.35	−0.03	2974.3	2273.3
Met4-20	14.6	7.2	169.6	152.7	81.4	215.7	4.1	0.99	0.00	0.23	0.16	−0.85	−0.40	−0.05	2365.1	2606.3
Met4-50	16.0	8.2	170.8	154.9	80.6	215.7	4.1	0.64	1.09	−0.04	0.20	−0.84	−0.38	−0.07	2033.0	3053.2
Met4-80	18.1	9.7	172.4	158.0	79.5	215.7	4.0	0.15	2.81	−0.43	0.23	−0.75	−0.34	−0.10	1542.9	3739.7
SAM	11.7	7.7	177.6	202.5	63.6	147.2	1.9	−4.60	−1.91	−1.50	1.22	−2.07	0.91	0.32	3690.7	2487.9
TMY	12.4	5.8	154.2	133.6	72.6	212.8	4.3	1.86	−2.24	−0.99	1.17	1.02	0.14	0.17	3804.9	2104.6
TMY2	12.0	5.5	166.1	154.0	76.7	209.6	4.3	1.07	−2.17	0.48	0.36	−0.04	−0.03	−0.06	3968.1	2092.3
TMY-20	13.6	6.7	156.2	128.4	79.3	212.8	4.3	2.29	−1.07	−1.00	0.67	0.26	−0.34	0.41	3106.9	2431.7
TMY2-20	13.3	6.6	168.1	146.9	80.9	209.6	4.3	1.27	−0.93	0.26	−0.07	−0.61	−0.22	0.07	3248.3	2418.8
TMY2-50	14.6	7.6	169.2	149.0	80.1	209.6	4.2	0.92	0.17	−0.02	−0.04	−0.60	−0.18	0.05	2810.3	2855.9
TMY2-80	16.8	9.2	170.8	152.1	78.8	209.6	4.1	0.42	1.91	−0.42	−0.01	−0.50	−0.13	0.02	2146.8	3538.3
TMY3	12.7	5.9	167.7	161.0	74.3	206.6	4.2	0.44	−1.68	0.37	0.49	0.04	0.03	−0.15	3506.2	2291.0
TMY3-20	14.0	6.9	169.7	154.6	77.8	206.6	4.1	0.57	−0.52	0.17	0.12	−0.43	−0.15	−0.10	2963.0	2618.9
TMY3-50	15.4	7.9	170.8	156.7	76.9	206.6	4.1	0.21	0.57	−0.11	0.16	−0.41	−0.12	−0.13	2552.0	3064.6
TMY3-80	17.5	9.4	172.4	159.5	75.7	206.6	4.0	−0.28	2.29	−0.50	0.18	−0.31	−0.07	−0.16	1897.9	3745.8
TMY-50	15.0	7.7	157.3	130.4	78.5	212.8	4.3	1.97	0.03	−1.24	0.68	0.32	−0.30	0.39	2688.4	2860.8
TMY-80	17.1	9.2	158.9	133.4	77.2	212.8	4.2	1.48	1.75	−1.63	0.70	0.44	−0.24	0.36	2094.4	3538.1
TMYx	13.3	6.3	162.2	142.2	74.4	221.4	4.3	1.40	−1.27	−0.24	0.79	0.43	0.15	−0.25	3746.5	2208.2
TMYx-20	14.6	7.2	163.9	143.2	74.4	221.4	4.3	1.14	−0.25	−0.40	0.70	0.39	0.14	−0.29	2785.3	2596.7
TMYx-50	15.9	8.2	165.1	145.1	73.7	221.4	4.2	0.80	0.84	−0.67	0.73	0.39	0.16	−0.32	2423.9	3047.4
TMYx-80	18.1	9.7	166.6	147.0	72.7	221.4	4.2	0.35	2.55	−1.08	0.74	0.46	0.17	−0.36	1807.1	3759.1
TMYxx	14.2	6.4	171.0	182.4	67.0	210.1	3.9	−0.99	−0.86	0.43	1.16	0.51	0.35	−0.60	3091.0	2417.6

Appendix D

Table A4. Details of the regression models standardized coefficients.



Appendix E

Table A5. Details of the regression models coefficients and standard errors.

	Variable	Coeff.	Cooling			Heating		
			Value	Std. error	Pr > t	Value	Std. error	Pr > t
LR	Intercept	β_0	−294.329	564.078	0.605	9727.159	1020.073	<0.0001
	DB	β_1	202.714	16.830	<0.0001	−396.019	30.435	<0.0001
	DP	β_2	145.703	25.119	<0.0001	58.934	45.424	0.202
	GHI	β_3	−11.689	4.437	0.012	16.153	8.024	0.051
	DNI	β_4	4.994	1.092	<0.0001	−11.924	1.975	<0.0001
	DHI	β_5	12.340	5.042	0.019	−48.300	9.117	<0.0001
	WD	β_6	−1.784	0.945	0.067	3.671	1.710	0.038
	WS	β_7	−88.557	34.102	0.013	114.733	61.671	0.071
LRS	Intercept	β_0	−1260.904	170.992	<0.0001	11147.753	497.251	<0.0001
	DB	β_1	211.805	16.863	<0.0001	−344.807	13.134	<0.0001
	DP	β_2	121.489	23.818	<0.0001	0.000	0.000	-
	GHI	β_3	0.000	0.000	-	0.000	0.000	-
	DNI	β_4	2.136	0.534	0.00	−9.170	0.998	<0.0001
	DHI	β_5	0.00	0.000	-	−24.446	4.990	<0.0001
	WD	β_6	0.00	0.000	-	0.000	0.000	-
	WS	β_7	−75.335	26.990	0.008	0.000	0.000	-
PCR	Intercept	β'_0	2782.435	10.854	<0.0001	2796.132	19.629	<0.0001
	F1	β'_1	−78.034	5.944	<0.0001	101.503	10.748	<0.0001
	F2	β'_2	344.047	7.763	<0.0001	−430.372	14.039	<0.0001
	F3	β'_3	−50.977	11.580	<0.0001	−114.421	20.941	<0.0001
	F4	β'_4	95.460	16.694	<0.0001	−206.575	30.189	<0.0001
	F5	β'_5	31.770	20.103	0.122	−41.260	36.354	0.263
	F6	β'_6	−185.519	39.064	<0.0001	635.965	70.642	<0.0001
	F7	β'_7	120.310	54.266	0.033	−157.836	98.134	0.116
PCRS	Intercept	β'_0	2782.435	11.061	<0.0001	2796.132	20.084	<0.0001
	F1	β'_1	−78.034	6.057	<0.0001	101.503	10.997	<0.0001
	F2	β'_2	344.047	7.911	<0.0001	−430.372	14.364	<0.0001
	F3	β'_3	−50.977	11.800	0.000	−114.421	21.426	<0.0001
	F4	β'_4	95.460	17.012	<0.0001	−206.575	30.889	<0.0001
	F5	β'_5	0.000	0.000	-	0.000	0.000	-
	F6	β'_6	−185.519	39.806	<0.0001	635.965	72.279	<0.0001
	F7	β'_7	120.310	55.298	0.036	0.000	0.000	-
PCR2	Intercept	β'_0	2782.435	18.136	<0.0001	2796.132	42.290	<0.0001
	F1	β'_1	−78.034	9.931	<0.0001	101.503	23.157	<0.0001
	F2	β'_2	344.047	12.971	<0.0001	−430.372	30.246	<0.0001

Appendix F

Table A6. Details of the average annual weather variables distributions and their associated principal component distributions used in this study.

Distribution	Var.	Parameter(s) [Standard Error]	Var.	Parameter(s) [Standard Error]
Normal (μ, σ) μ : Location σ : Scale	DB	(14.65 [0.18], 1.1716 [0.188])	F1	(0.00 [0.265], 1.826 [0.188])
	DP	(7.53 [0.171], 1.17 [0.125])	F2	(0.00 [0.210], 1.397 [0.146])
	GHI	(170.48 [-], 6.34 [-])	F3	(0.00 [0.138], 0.937 [0.097])
	DNI	(160.33 [0.265], 28.606 [0.42])	F4	(0.00 [0.096], 0.65 [0.068])
	DHI	(74.88 [0.42], 5.59 [0.34])	F5	(0.00 [0.08], 0.54 [0.056])
	WD	(196.001 [0.198], 30.27 [0.59])	F6	(0.00 [0.041], 0.278 [0.029])
	WS	(4.197 [0.074], 0.501 [0.05])	F7	(0.00 [0.029], 0.199 [0.021])
Lognormal (μ, σ) μ : Location σ : Scale	DB	(2.68 [0.016], 0.115 [0.012])		
	DP	(2.006 [0.023], 0.158 [0.016])		
	GHI	(5.138 [0.006], 0.038 [0.004])		
	DNI	(5.06 [0.028], 0.19 [0.02])		
	DHI	(4.313 [0.01], 0.078 [0.008])		
	WD	(5.26 [0.025], 0.17 [0.018])		
	WS	(1.125 [0.021], 0.145 [0.015])		
Logistic (μ, s) μ : Location s : Scale	DB	(14.53 [0.26], 1.002 [0.13])	F1	(0.362 [0.205], 0.946 [0.126])
	DP	(7.49 [0.18], 0.69 [0.08])	F2	(−0.062 [0.253], 0.831 [0.103])
	GHI	(170.65 [-], 3.40 [1.18])	F3	(0.091 [0.142], 0.531 [0.066])
	DNI	(159.88 [0.61], 15.39 [0.61])	F4	Fitted Normal
	DHI	(75.69 [0.42], 3.16 [0.27])	F5	Fitted Normal
	WD	(202.12 [0.26], 16.94 [0.22])	F6	(−0.019 [0.04], 0.155 [0.019])
	WS	Fitted Normal	F7	(0.000 [0.028], 0.111 [0.014])
Gamma (k, β) k : Shape β : Rate	DB	(74.59 [15.55], 0.19 [0.04])		
	DP	(40.49 [8.44], 0.18 [0.04])		
	GHI	(705.91 [147.19], 0.24 [0.05])		
	DNI	(29.09 [6.068], 5.51 [1.16])		
	DHI	(170.58 [35.64], 0.44 [0.09])		
	WD	(36.9 [7.68], 5.31 [1.11])		
	WS	(55.26 [11.51], 0.07 [0.016])		
Weibull (λ, k) λ : Scale k : Shape	DB	(8.79 [0.00], 15.43 [1.162])		
	DP	(7.016 [0.61], 8.04 [0.17])		
	GHI	(28.72 [0.34], 173.47 [0.11])		
	DNI	(6.19 [-], 172.14 [0.19])		
	DHI	(18.97 [0.59], 77.26 [0.13])		
	WD	(9.88 [0.59], 207.9 [0.59])		
	WS	11.86 [0.94], 4.39 [0.05])		

Table A6. Conts.

Distribution	Var.	Parameter(s) [Standard Error]	Var.	Parameter(s) [Standard Error]
Fisher-Tippett (β , μ) μ : Location β : Shape	DB	(1.42 [0.22], 13.83 [0.16])	F1	(2.156 [0.34], −1.013 [0.223])
	DP	(1.07 [0.17], 6.95 [0.12])	F2	(1.212 [0.189], −0.677 [0.139])
	GHI	(6.85 [1.07], 162.2 [0.70])	F3	(1.039 [0.163], −0.501 [0.108])
	DNI	(30.26 [7.74], 145.76 [3.09])	F4	(0.762 [0.12], −0.346 [0.075])
	DHI	(6.12 [0.96], 71.87 [0.65])	F5	(0.426 [0.071], −0.246 [0.017])
	WD	(32.93 [5.17], 179.52 [3.58])	F6	(0.226 [0.035], −0.129 [0.026])
	WS	(0.76 [0.12], 3.90 [0.67])	F7	(0.218 [0.034], −0.102 [0.022])
Erlang (k , λ) k : Shape λ : Rate	DB	(71.0 [0.14], 4.863 [0.08])	-	-
	DP	(39.0 [0.14], 5.312 [0.12])	-	-
	GHI	(705.0 [0.15], 4.141 [0.02])	-	-
	DNI	(30.0 [0.14], 0.192 [0.005])	-	-
	DHI	(175.0 [0.15], 2.344 [0.02])	-	-
	WD	(40.0 [0.14], 0.209 [0.005])	-	-
	WS	(68.0 [0.14], 16.373 [0.29])	-	-
Chi-square (DF) DF: Degree of Freedom	DB	(15.538 [0.245])	-	-
	DP	(8.414 [0.305])	-	-
	GHI	(171.346 [0.211])	-	-
	DNI	(158.583 [0.211])	-	-
	DHI	(75.66 [0.214])	-	-
	WD	(194.35 [0.211])	-	-
	WS	(5.12 [0.944])	-	-
GEV (k , β , μ) k : Scale β : Shape μ : Location	F1	(−0.74 [0.001], 1.88 [0.004], −0.22 [0.005])		
	F2	(−0.197 [0.15], 1.293 [0.18], −0.54 [0.22])		
	F3	(−0.77 [0.003], 1.07 [0.003], −0.1 [0.004])		
	F4	(−0.43 [0.04], 0.7 [0.036], −0.18 [0.05])		
	F5	(−0.47 [0.02], 0.58 [0.017], −0.14 [0.24])		
	F6	(−0.04 [0.102], 0.23 [0.026], −0.12 [0.036])		
	F7	(−0.35 [0.12], 0.208 [0.03], −0.06 [0.046])		
Beta (α , β , a , b) α , β : Shape a , b : Location	F1	(337.92 [35.76], 1.845 [0.026], −438.636 [3.51], 2.37 [-])		
	F2	(1.205 [0.026], 1.329 [0.028], −2.51 [-], 2.862 [-])		
	F3	(4.006 [0.35], 1.21 [0.02], −4.11 [0.26], 1.306 [-])		
	F4	(44.65 [1.67], 9.41 [0.068], −10.51 [0.214], 2.214 [0.136])		
	F5	(220.62 [7.66], 7.43 [0.058], −44.27 [0.43], 1.49 [0.095])		
	F6	(2.88 [0.044], 12.13 [0.122], −0.54 [0.023], 2.282 [0.21])		
	F7	(26.37 [0.46], 15.89 [0.075], −1.7 [0.04], 1.024 [0.035])		

Appendix G

Table A7. Summary of the BestFit distribution used in this study.

Var.	Best Fit Selection	p-value	Next Three Fits
F1	Beta4	0.036	Gumbel, Logistic, GEV
F2	Fisher-Tippett 2	0.851	GEV, Logistic, Normal
F3	Logistic	0.736	Fisher-Tippett2, Gumbel, Normal
F4	Beta4	0.233	Fisher-Tippett2, GEV, Normal
F5	Beta4	0.860	GEV, Normal
F6	Beta4	0.982	Fisher-Tippett2, Logistic, GEV
F7	Normal	0.914	Beta4, Fisher-Tippett2, Logistic
DB	Fisher-Tippett 2	0.984	Erlang, Lognormal, Logistic
DP	Lognormal	0.592	Logistic, Normal, Gamma2
GHI	Logistic	0.55	Erlang, Lognormal, Normal
DNI	Chisquare	0.299	Lognormal, Logistic, Fisher-Tippett2
DHI	Weibull2	0.376	Logistic, Lognormal, Normal
WD	Normal	0.00027	-
WS	Normal	0.396	Erlang, Lognormal, Gamma2

References

- Šujanová, P.; Rychtáriková, M.; Mayor, T.S.; Hyder, A. A Healthy, Energy-Efficient and Comfortable Indoor Environment, a Review. *Energies* **2019**, *12*, 1414. [\[CrossRef\]](#)
- Tian, W.; Heo, Y.; De Wilde, P.; Li, Z.; Yan, D.; Park, C.S.; Feng, X.; Augenbroe, G. A review of uncertainty analysis in building energy assessment. *Renew. Sustain. Energy Rev.* **2018**, *93*, 285–301. [\[CrossRef\]](#)
- Meacham, B.J. Sustainability and resiliency objectives in performance building regulations. *Build. Res. Inf.* **2016**, *44*, 474–489. [\[CrossRef\]](#)
- Escandón, R.; Suárez, R.; Sendra, J.J.; Ascione, F.; Bianco, N.; Mauro, G.M. Predicting the Impact of Climate Change on Thermal Comfort in A Building Category: The Case of Linear-type Social Housing Stock in Southern Spain. *Energies* **2019**, *12*, 2238. [\[CrossRef\]](#)
- Hallegatte, S. Strategies to adapt to an uncertain climate change. *Glob. Environ. Chang.* **2009**, *19*, 240–247. [\[CrossRef\]](#)
- Yassaghi, H.; Mostafavi, N.; Hoque, S. Evaluation of current and future hourly weather data intended for building designs: A Philadelphia case study. *Energy Build.* **2019**, *199*, 491–511. [\[CrossRef\]](#)
- Bueno, B.; Norford, L.; Hidalgo, J.; Pigeon, G. The urban weather generator. *J. Build. Perform. Simul.* **2013**, *6*, 269–281. [\[CrossRef\]](#)
- Belcher, S.; Hacker, J.; Powell, D. Constructing design weather data for future climates. *Build. Serv. Eng. Res. Technol.* **2005**, *26*, 49–61. [\[CrossRef\]](#)
- Khan, M.S.; Coulibaly, P.; Dibike, Y. Uncertainty analysis of statistical downscaling methods. *J. Hydrol.* **2006**, *319*, 357–382. [\[CrossRef\]](#)
- Goodess, C.; Hanson, C.; Hulme, M.; Osborn, T. Representing Climate and Extreme Weather Events in Integrated Assessment Models: A Review of Existing Methods and Options for Development. *Integr. Assess.* **2003**, *4*, 145–171. [\[CrossRef\]](#)
- Nik, V.M.; Kalagasidis, A.S. Impact study of the climate change on the energy performance of the building stock in Stockholm considering four climate uncertainties. *Build. Environ.* **2013**, *60*, 291–304. [\[CrossRef\]](#)
- Yassaghi, H.; Hoque, S. An Overview of Climate Change and Building Energy: Performance, Responses and Uncertainties. *Buildings* **2019**, *9*, 166. [\[CrossRef\]](#)
- Belazi, W.; Ouldboukhithine, S.-E.; Chateauneuf, A.; Bouchair, A.; Bouchair, H. Uncertainty analysis of occupant behavior and building envelope materials in office building performance simulation. *J. Build. Eng.* **2018**, *19*, 434–448. [\[CrossRef\]](#)
- Gaetani, I.I.; Hoes, P.P.-J.; Hensen, J.J. On the sensitivity to different aspects of occupant behaviour for selecting the appropriate modelling complexity in building performance predictions. *J. Build. Perform. Simul.* **2016**, *10*, 601–611. [\[CrossRef\]](#)
- IPCC. 2014: *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*; Core Writing Team, Pachauri, R.K., Meyer, L.A., Eds.; IPCC: Geneva, Switzerland, 2014; p. 151.
- Martinez-Soto, A.; Jentsch, M.F. A transferable energy model for determining the future energy demand and its uncertainty in a country's residential sector. *Build. Res. Inf.* **2019**, *48*, 587–612. [\[CrossRef\]](#)
- Bortolini, R.; Forcada, N. A probabilistic performance evaluation for buildings and constructed assets. *Build. Res. Inf.* **2020**, *48*, 838–855. [\[CrossRef\]](#)
- Eisenhower, B.; O'Neill, Z.; Fonoberov, V.A.; Mezić, I. Uncertainty and sensitivity decomposition of building energy models. *J. Build. Perform. Simul.* **2012**, *5*, 171–184. [\[CrossRef\]](#)
- Cecconi, F.R.; Manfren, M.; Tagliabue, L.C.; Ciribini, A.L.C.; De Angelis, E. Probabilistic behavioral modeling in building performance simulation: A Monte Carlo approach. *Energy Build.* **2017**, *148*, 128–141. [\[CrossRef\]](#)

20. Nik, V.M. Making energy simulation easier for future climate—Synthesizing typical and extreme weather data sets out of regional climate models (RCMs). *Appl. Energy* **2016**, *177*, 204–226. [[CrossRef](#)]
21. Shen, P. Impacts of climate change on U.S. building energy use by using downscaled hourly future weather data. *Energy Build.* **2017**, *134*, 61–70. [[CrossRef](#)]
22. Rodríguez, G.C.; Andrés, A.C.; Muñoz, F.D.; López, J.M.C.; Zhang, Y. Uncertainties and sensitivity analysis in building energy simulation using macroparameters. *Energy Build.* **2013**, *67*, 79–87. [[CrossRef](#)]
23. Tian, W.; de Wilde, P. Uncertainty and sensitivity analysis of building performance using probabilistic climate projections: A UK case study. *Autom. Constr.* **2011**, *20*, 1096–1109. [[CrossRef](#)]
24. Sun, Y.; Heo, Y.; Tan, M.; Xie, H.; Wu, C.J.; Augenbroe, G. Uncertainty quantification of microclimate variables in building energy models. *J. Build. Perform. Simul.* **2013**, *7*, 17–32. [[CrossRef](#)]
25. Gang, W.; Wang, S.; Shan, K.; Gao, D. Impacts of cooling load calculation uncertainties on the design optimization of building cooling systems. *Energy Build.* **2015**, *94*, 1–9. [[CrossRef](#)]
26. De Wilde, P.; Tian, W. Predicting the performance of an office under climate change: A study of metrics, sensitivity and zonal resolution. *Energy Build.* **2010**, *42*, 1674–1684. [[CrossRef](#)]
27. Wang, L.; Mathew, P.; Pang, X. Uncertainties in energy consumption introduced by building operations and weather for a medium-size office building. *Energy Build.* **2012**, *53*, 152–158. [[CrossRef](#)]
28. González, V.G.; Colmenares, L.Á.; López-Fidalgo, J.; Ruiz, G.R.; Bandera, C.F. Uncertainty's Indices Assessment for Calibrated Energy Models. *Energies* **2019**, *12*, 2096. [[CrossRef](#)]
29. Brohus, H.; Frier, C.; Heiselberg, P.; Haghighat, F. Quantification of uncertainty in predicting building energy consumption: A stochastic approach. *Energy Build.* **2012**, *55*, 127–140. [[CrossRef](#)]
30. Fabrizio, E.; Monetti, V. Methodologies and Advancements in the Calibration of Building Energy Models. *Energies* **2015**, *8*, 2548–2574. [[CrossRef](#)]
31. Braun, M.; Altan, H.; Beck, S. Using regression analysis to predict the future energy consumption of a supermarket in the UK. *Appl. Energy* **2014**, *130*, 305–313. [[CrossRef](#)]
32. Catalina, T.; Iordache, V.; Caracaleanu, B. Multiple regression model for fast prediction of the heating energy demand. *Energy Build.* **2013**, *57*, 302–312. [[CrossRef](#)]
33. Yassaghi, H.; Gurian, P.L.; Hoque, S. Propagating downscaled future weather file uncertainties into building energy use. *Appl. Energy* **2020**, *278*, 115655. [[CrossRef](#)]
34. Ivanov, V.; Bras, R.L.; Curtis, D.C. A weather generator for hydrological, ecological, and agricultural applications. *Water Resour. Res.* **2007**, *43*. [[CrossRef](#)]
35. Fatichi, S.; Ivanov, V.; Caporali, E. Simulation of future climate scenarios with a weather generator. *Adv. Water Resour.* **2011**, *34*, 448–467. [[CrossRef](#)]
36. Peleg, N.; Fatichi, S.; Paschalis, A.; Molnar, P.; Burlando, P. An advanced stochastic weather generator for simulating 2-D high-resolution climate variables. *J. Adv. Model. Earth Syst.* **2017**, *9*, 1595–1627. [[CrossRef](#)]
37. Jentsch, M.F.; James, P.A.B.; Bourikas, L.; Bahaj, A.S. Transforming existing weather data for worldwide locations to enable energy and building performance simulation under future climates. *Renew. Energy* **2013**, *55*, 514–524. [[CrossRef](#)]
38. Chen, Y.; Wen, J. *Whole Building System Fault Detection Based on Weather Pattern Matching and PCA Method*; Institute of Electrical and Electronics Engineers (IEEE); China Agricultural University: Beijing, China, 2017; pp. 728–732.
39. Wang, E. Benchmarking whole-building energy performance with multi-criteria technique for order preference by similarity to ideal solution using a selective objective-weighting approach. *Appl. Energy* **2015**, *146*, 92–103. [[CrossRef](#)]
40. Devore, J.L.; Farnum, N.R.; Doi, J. *Applied Statistics for Engineers and Scientists*; Nelson Education: Scarborough, ON, Canada, 2013.