The University of Nottingham

Doctoral Thesis

Emulation based Uncertainty and Sensitivity Analysis for Complex Building Performance Simulation

Parag Wate

Supervisors:
Prof. Darren Robinson
Dr. Marco Iglesias
Prof. Volker Coors

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

in the

Faculty of Engineering
Department of Architecture and Built Environment

March 2020

To My Parents and loving brother Harshad.

Abstract

Uncertainties in building performance simulation arise from the propagation of both epistemic (e.g. due to poorly or un- observed input parameters) and aleatory (e.g. due to occupants' stochastic interactions) uncertainties. A good framework for the quantification and decomposition of uncertainties in dynamic building performance simulation should: (i) simulate the principle deterministic processes influencing heat flows in buildings and the stochastic perturbations to them, (ii) quantify and decompose the total uncertainty into its respective sources, and the interactions between them, and (iii) achieve this in a computationally efficient manner. In this thesis, a new framework, named as Emulation based Uncertainty and Sensitivity Analysis (EmUSA) framework, is introduced which does just that. Two case studies, a monozone office building for the frameworks' proof of principle and the other multizone residential building for it's extension are being investigated to perform Uncertainty and Sensitivity Analysis for the Complex case of Building Performance Simulation (BPS) i.e. the Stochastic BPS (S-BPS). The detailed development of this new framework for emulating both the mean and the variance in the response of a stochastic building performance simulator (EnergyPlus co-simulated with a multi-agent stochastic simulator called No-MASS) is presented, for heating and cooling load predictions.

The effectiveness of these emulators, applied to a monozone office building, has been demonstrated and evaluated to quantify and decompose prediction uncertainties. With a range of 25-50 kWh/m^2, the epistemic uncertainty due to envelope parameters dominates over aleatory uncertainty relating to occupants' interactions, which ranges from 6-8 kWh/m^2, for heating load predictions. However, the converse is observed for cooling loads, which vary by just 3 kWh/m^2 for envelope parameters, compared with 8-22 kWh/m^2 for their aleatory counterparts. This is due to the correspondingly larger stimuli provoking occupants' interactions. Sensitivity indices corroborate this result, with wall insulation thickness (0.97) and occupants' behaviours (0.83) having the highest impacts on heating and cooling load predictions respectively. This new emulator framework (including training and subsequent deployment), for

the case of monozone office, achieves a factor of c.30 reduction in the total computational budget, whilst overwhelmingly maintaining predictions within a 95% confidence interval, and successfully decomposing prediction uncertainties.

The second case of a multizone residential building demonstrates the applicability and readiness of this new framework for the uncertainty problems (pertaining to the building scale) of increasing scope and complexity. An increased scope is analysed by investigating the impact of uncertainties in detailed thermophysical (e.g. conductivity, density, solar transmittance, etc.) properties of opaque and transparent building envelope materials, along with the mostly inexact infiltration rate. For the increased complexity, a greater degree of stochastic occupants' interaction freedom (e.g. use of lights) along with the stochastic phenomena of occupants' behaviours (e.g. presence, windows and shades interactions) was enabled. To this end, the framework has shown a ready extensibility and applicability by incorporating a parameter screening step as a pre-step to the Uncertainty Quantification (UQ) workflow for the purpose of segregating the most influential input parameters from the non-influential ones. For the extended case, the framework is now overall (training and subsequent deployment) moderately computationally efficient as compared to the proof of principle case with the costs reduced to a low fraction (0.25) of its equivalent classical counterpart, with predictivity coefficient $Q_2 > 0.95$. As the occupants' behaviours are relatively constrained during heating season, the effects of uncertainties in wall insulation conductivity (sensitivity index = 0.8428) and thickness (0.0771) dominate the heating load predictions than the stochasticity due to occupants' interactions (0.0010). The similar insignificance of stochastic phenomena (0.0130) as compared to brickwork solar absorptance (0.5032) and wall insulation conductivity (0.3714) has been observed during cooling season, due to the acceleration of short- and long-wave radiation phenomena which then dampens the occupants' interaction to restore their thermal comfort.

Keywords : Gaussian Process Emulator, building performance, stochasticity, uncertainty quantification and decomposition.

Acknowledgements

I would like to thank my supervisor trio Prof. Darren Robinson (The University of Sheffield, UK), Dr. Marco Iglesias (The University of Nottingham, UK) and Prof. Volker Coors (Stuttgart University of Applied Sciences, Germany) for giving a scientific rigour to my abstract ideas during this remarkable journey of five years that has now lead to this interdisciplinary doctoral thesis. A grateful acknowledgement to the European Commission for providing financial support during the conduct of this research under the FP7-PEOPLE-2013 Marie Curie Initial Training Network CI-NERGY project with Grant Agreement Number 606851. I would also like to acknowledge the technical support received from the Baden-Wrttemberg High Performance Computing (bwHPC) team and also from the University of Nottingham High Performance Computing Facility.

I also take this opportunity to thank my internal and external examiners, namely Dr. Simon Preston (The University of Nottingham, UK) and Prof. Godfried Augenbroe (Georgia Institute of Technology, Atlanta, USA) for taking time to read through my Thesis and examining me. I also thank Prof. Richard Wilkinson (The University of Sheffield, UK) for an initial guidance meeting on discussing the potential applications of Gaussian Process Regression. Also, a special thank to Dr. Jacob Chapman (Data Scientist, Kindred Group plc, London, UK), who developed Multi-agent stochastic simulation platform during his time at The University of Nottingham as a doctoral researcher, for his support and addressing my queries regarding the platform while incorporating his work during my development process of Uncertainty and Sensitivity Analysis framework in this thesis. I also thank my Journal and Conference manuscripts reviewers for their helpful comments and suggestions for further improvements.

A special thanks to my HfT, Stuttgart colleagues (Ruben Pesch, Habib Dastageeri, Eric Duminil, Sven Schneider, Preston Rodriguez) for taking me out of my PhD bubble from time to time and motivating me to learn the German language. Without your encouragement, it would have been next to impossible for me to know and understand the German multi-culture.

My CI-NERGY fellow researchers, you were simply great! Many thanks to

you all (Ana, Pilar, Nils, Sebastien, Mario, Nina, Kunpeng, Giovanni, Solene, Jean-Loup, Giorgio, Lisa and Dilay) for introducing Europe and it's culture to me. I have learnt a lot from you all.

My Nottingham colleagues (Hannah, Renoy, Phil, Ben, Philip and Gustavo), I really enjoyed our lunch time frisbee-breaks, summer outings and evenings in the UK, thank you for everything. A tons of thanks to my friends (Rabi, Hristijan and Athanasios) for those cycle rides through Stuttgart wine yards, your company and a lot of talks in leisure times.

Also, my new Sheffield colleagues (Dr. Isaiah Durosaiye, Stuart McKenzie, Dr. Daniel Paredes-Soto, Reena Sayani and Dr. Sayantan Bhattacharya), thank you for your really warm welcome and friendship when I just arrived and started my new position as a Lecturer in Architectural Science there at the University of Sheffield.

And finally my parents, my relatives, my old college friends (Unmesh, Sushrut and Vineet). Thank you for all your blessings and throughout support from my distant home in India and from elsewhere abroad.

And brother Harshad, what to say. My brother Harshad, a "Thank you" is a very small word for what I want to say here. You were and are holding one end of our partnership on the pitch of life so tight come what may, that only because of this I could take on this endeavour! Tons of love and a HUGE hug!

Contents

List of Figures xi

List of Tables xv

1 Introduction 1

1.1 Motivation . . . 2

1.1.1 The *performance gap* . . . 2

1.1.2 The *uncertainty* . . . 4

1.1.3 The *computational cost* . . . 4

1.2 Research Hypothesis and Objectives . . . 6

1.3 Knowledge Contribution . . . 9

1.4 Thesis structure . . . 10

1.4.1 Thesis outline . . . 14

1.4.2 Publications . . . 16

2 Stochastic Building Performance Simulation (S-BPS) 17

2.1 Deterministic phenomena . . . 17

2.2 Stochastic phenomena . . . 24

2.2.1 Stochastic models of Occupant behaviours . . . 24

2.3 S-BPS Platform : A coupling of phenomena . . . 31

3 Methods for Uncertainty and Sensitivity Analysis 35

3.1 Measures for Uncertainty and Sensitivity . . . 36

3.2 Sources of uncertainty . . . 37

3.3 Uncertainty Treatment : The process flow . . . 39

3.3.1 Qualitative methods . . . 40

3.3.2 Quantitative methods . . . 45

3.4 Uncertainty assessment approaches in BPS 54
3.5 Summary 61

4 Development of Mathematical and Computational Framework : EmUSA 65

4.1 Metamodeling Techniques 65
4.1.1 Selection of a candidate metamodelling technique . . . 70
4.1.2 The proposal 73
4.2 Standard Gaussian Process (GP) Metamodeling for deterministic simulators 75
4.2.1 Validation measures 77
4.3 Proposed methodology for the emulation of S-BPS 79
4.3.1 Gaussian process fitted to the log-variance $\log V(\mathbf{X})$. . 79
4.3.2 Gaussian process fitted to the mean response $f_d(\mathbf{X})$. . . 82
4.3.3 Emulation of the stochastic response 85
4.4 Global sensitivity analysis through metamodeling 86
4.4.1 The deterministic case 86
4.4.2 The stochastic case 87
4.5 Computational aspects 93
4.6 EmUSA methodology workflow : A summary 93

5 Application I - A proof of principle study 97

5.1 Case Study 98
5.1.1 Experiment set up 98
5.1.2 Parameters studied 101
5.2 Results and Discussion 102
5.2.1 Design points vs repetitions 102
5.2.2 Gaussian Process emulation of the mean $f_d(\mathbf{X})$ and log-variance $\log V(\mathbf{X})$ of the S-BPS outputs 105
5.2.3 Uncertainty quantification using the GP emulator . . . 106
5.2.4 Uncertainty decomposition using the GP emulator : Sensitivity analysis 110
5.3 Conclusions 116

6 Application II - An extended study **119**
6.1 Case Study . . . 119
6.1.1 Experiment set up . . . 119
6.1.2 Parameters studied . . . 120
6.2 Results and Discussions . . . 124
6.2.1 Parameter screening . . . 124
6.2.2 Gaussian Process Emulation of the S-BPS outputs . . . 127
6.2.3 Uncertainty Quantification using the GP emulator . . . 131
6.2.4 Uncertainty decomposition using the GP emulator : Sensitivity analysis . . . 135
6.2.5 Computational savings in Sensitvity Analysis under EmUSA framework . . . 142
6.3 Conclusions . . . 142

7 Conclusions, Future works and Recommendations **145**
7.1 Answers to the Research Questions . . . 145
7.2 Future Works . . . 148
7.3 Recommendations . . . 149

Appendix **150**

Bibliography **153**

List of Figures

2.1 Thermal Energy flow paths (deterministic phenomena) 18
2.2 Systemic illustration of thermal energy flow paths and their interactions: Deterministic and Stochastic sub-systems of BPS 23
2.3 Occupants' window interaction algorithm workflow after hybrid model by Haldi and Robinson [1] 28
2.4 Schematic of EnergyPlus and No-MASS Co-simulation platform [2]: the Stochastic Building Performance Simulator (S-BPS) . 33

3.1 Sources of uncertainty at different stages of a simulation study [3] 38
3.2 Illustration of trajectory in input space when number of inputs $k = 2$ 44
3.3 Single loop Uncertainty propagation scheme 48
3.4 Double loop Uncertainty propagation scheme 49
3.5 Classification system of Uncertainty and Sensitivity Analysis (USA) methods against the computational expense (Adapted from Pianosi et al. [4], Tian [5], Nguyen and Reiter [6] and Tian et al. [7] for the methods explored in earlier sections 3.3.1, 3.3.2 and applied in Chapters 5, 6) (Note: k - number of inputs) 62

4.1 Monte Carlo Uncertainty Propagation Scheme for S-BPS . . . 80
4.2 Flow diagram for Algorithm 1 84
4.3 Flow diagram for Algorithm 2 92
4.4 Model Parameterisation Workflow 94
4.5 Computational Workflow 95

5.1 Building geometry of hypothetical office 99
5.2 Comparison of prediction accuracies of mean and log-variance GPs 106

5.3 GP regression surface fitting to sampled mean (left) and log-variance (right) of heating demand (top) and cooling demand (bottom). The black asterisks are the training points and every point on the mesh is a test data point [not used for training]. The predictivity coefficients values are computed based on the evaluation of fitted GP surfaces at 60 randomly selected test data points, in Figure 5.4 subsection 5.2.3. 107
5.4 Comparison of S-BPS simulated random values within GP predicted credible intervals at test data for heating (a) and cooling (b) demand 108
5.5 Uncertainty propagation comparison between GP predicted (blue) and S-BPS simulated PDFs (red) for heating demand 109
5.6 Uncertainty propagation comparison between GP predicted (blue) and S-BPS simulated PDFs (red) for cooling demand 110
5.7 Comparison of GP predicted (left) and S-BPS simulated (right) CDFs for heating (top) and cooling (bottom) demand 111
5.8 Comparison between S-BPS Simulated and GP predicted means of heating (left) and cooling (right) demand replicates in Figures 5.7 111
5.9 Sensitivity Tornado plots for heating (a) and cooling (b) demand 113
5.10 Convergence of main (left) and total (right) effect sensitivity indices of design variables for heating demand at distinct sample sets of increasing size. The top row shows the convergence of the width of index credible interval and bottom row shows the actual credible interval (Eq. 4.49). 114
5.11 Convergence of main (left) and total (right) effect sensitivity indices of design variables for cooling demand at distinct sample sets of increasing size. The top row shows the convergence of the width of index credible interval and bottom row shows the actual credible interval (Eq. 4.49). 115

6.1 Building geometry of multizone house 120
6.2 Building geometry of multizone house : First floor 121
6.3 Building geometry of multizone house : Ground floor 121
6.4 mu-sigma plot : Heating demand 125

6.5 mu-sigma plot : Cooling demand 125
6.6 Main / Individual effect plots : Heating demand 129
6.7 Main / Individual effect plots : Cooling demand 130
6.8 Comparison of S-BPS simulated random values within GP predicted credible intervals at test data for heating (a) and cooling (b) demand 131
6.9 Uncertainty propagation comparison between GP predicted (blue) and S-BPS simulated PDFs (red) for heating demand 132
6.10 Uncertainty propagation comparison between GP predicted (blue) and S-BPS simulated PDFs (red) for cooling demand 133
6.11 Comparison of GP predicted (left) and S-BPS simulated (right) CDFs for heating (top) and cooling (bottom) demand 134
6.12 Comparison between S-BPS Simulated and GP predicted means of heating (left) and cooling (right) demand replicates in Figures 6.11 134
6.13 Sensitivity Tornado plots for heating (a) and cooling (b) demand 138
6.14 Convergence of main (left) and total (right) effect sensitivity indices of design and physical properties input variables for heating demand at distinct sample sets of increasing size. The top row shows the convergence of the width of index credible interval and bottom row shows the actual credible interval (Eq. 4.49). 140
6.15 Convergence of main (left) and total (right) effect sensitivity indices of design and physical properties input variables for cooling demand at distinct sample sets of increasing size. The top row shows the convergence of the width of index credible interval and bottom row shows the actual credible interval (Eq. 4.49). 141

List of Tables

2.1 Exemplar time-dependent probabilities for the occupants' activities during a typical evening hour in residential living room . 26

3.1 Comparison between FAST and Sobol methods 53

3.2 Comparison between sensitivity analysis methods. (Note: k and N are number of inputs and samples respectively.) 53

3.3 Summary of applications uncertainty and sensitivity analysis approaches in BPS. (Note: ISO 13790:2008 Quasi-steady simplified monthly method, EnergyPlus and ESP-r Dynamic heat conduction and balance simulation program, Simplified MAV - Simplified Monthly Average Values) 60

4.1 Comparison between metamodeling techniques against the performance criteria (Note: fully - 5 points, partially - 3 points, not-at-all - zero. The numbers in the brackets are the points scored by the technique based on the subjective judgement of fully / partially / not-at-all.) 71

5.1 Thermophysical properties of construction layers : thickness, thermal conductivity, specific heat capacity, density and U-value 100

5.2 Comparison of quality of prediction Q of fitted GP response surfaces at test data given a fixed computation budget (of 3000 simulation runs) for training. Here the item entry 5×600 represents 600 simulation repetitions at 5 design points and so on. 104

5.3 Comparison of emulator uncertainty (Eq. 4.10) of fitted GPs in terms of proportion of test points lying inside 95% uncertainty interval 104

5.4 Sensitivity indices quantifying the impacts of thickness, transmittance and occupants' behaviour on the heating and cooling energy demand predictions. S_1 and S_2 are individual effects due to thickness and transmittance and S_{T_w} total effect due to occupants' behaviours. 112
5.5 Comparison of computational runs between Classical and EmUSA framework 116
5.6 Comparison of computational expense between Classical and EmUSA framework (One simulation run benchmark = 20 seconds) 116

6.1 Thermophysical properties of construction layers : thickness, thermal conductivity, specific heat capacity and density 122
6.2 Thermophysical properties of construction layers : thickness t(m), thermal conductivity λ(W/(mK)), specific heat capacity C(J/(kgK)), density ρ(kg/m^3), thermal emittance ϵ and solar absorptance α. In brackets of Surface and Material columns, a number of parameters (in total 59) considered to be uncertain. "N" denotes Normal distribution with mean and standard deviation and "U" denotes Uniform distribution with lower and upper bounds. 123
6.3 The uncertain parameters' annotations and their descriptions 126
6.4 Sensitivity indices quantifying the impacts of input parameters and occupants' behaviour on the heating and cooling energy demand predictions. S_1, S_2, S_3, S_4, S_5, S_6, S_7 and S_8 are the individual effects due to BrickC, BrickSA, PolyT, PolyC, ConcC, FoamC, WindowST and InfilRate respectively and S_{T_w} total effect due to occupants' behaviours 137
6.5 Comparison of computational runs between Classical and EmUSA framework for the extended study 139
6.6 Comparison of computational expense on HPC between Classical and EmUSA framework for the extended study (One simulation run benchmark = 90 seconds) 140

7.1 GP prior distribution settings for hyperparameters 152
7.2 GP prior value settings for hyperparameters 152

Chapter 1

Introduction

Cities consume nearly three-fourths of global primary energy resource and are responsible for about half of the world's total greenhouse gasses (GHG) emissions [8]. Transport, Households and Industry - trinity of sectors account for majority of the energy consumption in cities. In Europe, where about three-fourths of its total population live in urban areas, the Household sector stands second highest end user of energy in the EU-28 [9], with a share of renewables in an energy mix at EU level for heating, cooling and electricity end-use in the residential sector at mere 15 percent [10] [11]. Cities aim to reach energy efficiency and sustainability targets by increasing renewable energy use and promoting low GHG emissions development.

Under Europe 2020 Strategy, the EU has put in place "20/20/20" climate change and energy targets. The strategy aims at increasing the share of renewable energy in final energy consumption and reducing GHG emission and energy consumption through energy efficiency measures, all the three by 20 percent compared to 1990 levels [12]. In the UK, the Climate Change Act 2008 legislates for an 80% reduction in GHG emissions with respect to 1990 levels by 2050 [13]. Energy Performance of Buildings [14] and Energy Efficiency [15] Directives legislate the calculation of amount of energy consumed in buildings for their heating and cooling needs and also indoor thermal comfort conditions using Building Performance Simulation (BPS) approach. In addition, the utility of BPS approach has also been encouraged by agencies and organisations such as International Energy Agency (IEA) and Department of Energy in the US to support the strategic decisions that acknowledge new energy saving po-

tentials in low GHG emissions development [16]. BPS approach thus acts as an instrument and implementing measure for the decision making bodies to aid in assessing the energy performance of buildings [17].

1.1 Motivation

1.1.1 The *performance gap*

Building performance simulation tools have evolved considerably in recent decades, with significant efforts having been invested to improve the scope and validity of their underlying algorithms and the usability of interfaces to these algorithms, to the extent that they have now entered into the mainstream. Indeed they are now commonly required to demonstrate compliance with national regulations, as required for example by the European Energy Performance of Buildings Directive [18]. Yet predictions from standard Building Performance Simulation tools - whether in relation to existing or proposed buildings - continue to deviate significantly from those that are observed. This is commonly referred to as the *performance gap*. This deviation is problematic for two main reasons. Firstly, it undermines confidence in the accuracy of energy use estimates, as required for example by energy performance certifications schemes. Secondly, it may negatively impact on the design decisions made, potentially leading to suboptimal design decisions, or design features that lack robustness.

The causes of the energy performance gap are many and complex [19, 20], including: (1) errors or omissions in the simulation of non-trivial phenomena [21] such as coupled heat and moisture flow and its dependence on changing thermo-physical properties or of occupants' stochastic interactions [22], (2) use of standardised external boundary conditions and building systems' control settings for thermostat and operation hours [23], (3) practical issues, such as the inability to fully predict a building's future functions during design, or to account for plausible future variations in function during the design process, (4) failure to meet specified insulation and airtightness standards due to poor workmanship, (5) unavailability of data regarding e.g. internal gain and plug loads that are representative of the building's post-occupancy operation [24].

Efforts to improve the robustness of BPS tools, for instance by develop-

ing 'reference simulations' as part of model calibration processes [25, 26], and through feedback from post-occupancy evaluation exercises [27], are ongoing. But a particularly promising avenue of current exploration is through the characterisation and propagation of aleatory (e.g. due to occupants' stochastic interactions) and epistemic (e.g. due to poorly or un- observed input parameters) uncertainties, so that these can be quantified and decomposed. This is because for the range of model improvement strategies addressing the energy *performance gap* to be well targeted, it is important to account for, well in advance, the uncertainties that are arising in energy performance predictions. The challenges lie in - systematic identification of the sources of uncertainties, characterisation of these uncertainties mathematically, their [computationally efficient] propagation from the sources through the model till the predictions, numerical quantification and analysis of uncertainties at source and prediction end and decomposition of the prediction uncertainty into the respective sources that cause it. Addressing these challenges, enable the numerical quantification of the importance of a given uncertain input/process relative to others and as well as their respective impact on the model predictions (so that the uncertainties arising from those highly influential inputs/processes can be subsequently addressed if possible).

It is self evident that buildings' energy demands for heating and cooling are dependent upon the heat transfers across the building envelope: on conductive heat transfers across opaque and transparent surfaces; on radiative transfers through the transparent surfaces; on advective transfers through accidental (imperfections) or deliberate (windows) openings in the envelope. These latter pathways may also be influenced by occupants' interactions with windows and shading devices. Haldi and Robinson [28] have shown that these aleatory (behavioural) uncertainties in energy use can be equivalent, even greater, in magnitude to their epistemic (envelope) counterparts. They also argue [29] that as the performance of the envelope improves, thus better conserving energy, so the consequences of occupants' actions to regulate the envelope will be exaggerated; that it will be more and more important to accurately model occupants' behaviours and the uncertainties arising from them, simultaneously with their epistemic counterparts.

1.1.2 The *uncertainty*

The uncertainties in building energy use prediction are basically two prominent kinds of uncertainties namely, epistemic and aleatory that affect plausibility of the energy predictions. Chapman et al. [22] categorise the causes for these uncertainties as : (type I) errors in the modelling of deterministic phenomena or indeed the neglect of these phenomena, (type II) errors in the inputs to these deterministic models, (type III) errors in the modelling of stochastic phenomena or indeed the neglect of these phenomena, (type IV) errors in the inputs to these stochastic models. Types I and III essentially relate to epistemic uncertainties, whereas types II and IV represent a combination of epistemic and aleatory uncertainties.

More specifically, epistemic uncertainty includes the phenomena assumptions such as considerations of only uni-directional heat diffusions through constructions, time-invariant thermophysical properties and ignorance of thermal storage in the modelling of heat diffusions. It also relates to the input data assumptions such as lack of precise knowledge of building architectural details, construction layer compositions, material thermophysical properties, household compositions and the type of installed heating systems.

Aleatory uncertainty consists of two kinds of stochastic sub-components: outdoor weather and indoor human behaviour. Stochastic nature of empirical quantities such as wind pressure across the building envelope affect convective heat transfers occurring due to ventilation and infiltration or cloud cover obstructing the transmitted shortwave irradiation. On the other hand, stochastic occupant presence and their interactions with the envelope devices such as windows and shades, lights and electrical appliances result in a random internal heat gain quantity, which adds a layer of stochastic perturbations to the energy predictions and also affects heat transmission across the envelope.

1.1.3 The *computational cost*

The relationship between building design / retrofit input parameters, $\mathbf{X} = (X_1, \ldots, X_p)$, (e.g. wall thickness, window transmittance) and the corresponding BPS outputs (e.g. annual heating and cooling demand), Y, can be expressed via $Y = f(\mathbf{X})$, where f is typically a deterministic function

with values defined by running the simulator for a given choice of inputs $\mathbf{X}$. When the inputs are uncertain, only their probability density function, $p(\mathbf{X})$, may be available to us. Since the functional form of f is generally unknown, propagating uncertainty from BPS inputs through to the corresponding predictions/outputs is only feasible via Monte Carlo (MC) sampling as applied in [30] [31] for uncertainty and sensitivity analysis and demonstrated using an uncertainty analysis workbench [32]. More specifically, samples from $p(\mathbf{X})$ are mapped onto the corresponding outputs and then used to construct an empirical distribution that approximates the distribution of model predictions $p(Y)$.

Even when optimal MC sampling techniques such as Latin Hypercube Sampling (LHS) [33] and quasi-random sequences [34] are used, accurate approximations of $p(Y)$ via sampling often requires thousands of samples (of the order of 10^3 to 10^5 simulation runs) depending upon model complexity and the number of input parameters under consideration. Furthermore, if we wish to decompose prediction uncertainty (i.e. to know which amongst k inputs are dominant) via a variance decomposition technique, the total computation cost is $N(k+2)$ simulations where N is the number of samples (as per theorem 1 stated in [35] on the extension of Sobol [36, 37] sensitivity analysis method).

For the simulation of complex phenomena, as is the case in building simulation (likewise also the case of environmental models elsewhere in [4]), each individual run entails a non-trivial computational cost. For instance, in tens of seconds even for a relatively simple shoebox building case study model in Chapter 5 (and extended case study in Chapter 6). Moreover, the number of uncertain inputs can be relatively large (e.g. order of 10^2 in [31] on uncertainty analysis in building simulation). Burhenne (Chapter 5 in [38]) reported the cost to be already of order 10^3 to 10^4 for less than 10 (actually just $6-7$) input parameters for the variance-based sensitivity analysis of energy performance (and related cost-effectiveness) to different design options of a typical German residential building using MC sampling techniques. Comprehensive uncertainty quantification studies thus tend to be highly computationally intensive, resource and time consuming; at the limits, simply intractable.

Furthermore, a conduct of sensitivity indices convergence study [39], which is a trial and error approach involving model evaluations at increasing sample

sizes N, allows to evaluate the reliability in indices values which is more desirable than a single random estimate of it (which is the practice till now with the applications of classical Monte Carlo methods in building simulation). This is important to appreciate the uncertainty in sensitivity indices values itself when Uncertainty Analysis (UA) goes mainstream for industrial applications.

The number of model simulations to be performed depends upon the complexity of building energy model under study, the number of input parameters involved and as well as how quickly i.e. at what sampling size the indices convergence criterion is met. This computational load [in terms of both computation time and number of computations] can be immense, as shown in the case studies (in Chapter 5 and 6), for to be performed under a classical Monte Carlo settings that employ a computationally expensive simulator and also that industrial UA applications in building energy modelling cannot afford to invest in such computational expense. This thesis therefore explores a semi-parametric Gaussian Process Regression as an emulation approach approximating the simulator response for the purposes of achieving a high computational efficiency in such Uncertainty and Sensitivity Analysis studies.

The challenge of effectively accounting for the factors that lead to *performance gap* by quantifying the *uncertainty* in building energy predictions with the frugal *computational cost* points out the clear need for a new Uncertainty and Sensitivity Analysis framework in dynamic building performance simulation that: (i) proposes a plausible equivalent emulator efficiently approximating a building performance simulator that rigorously encapsulates stochastic processes arising, in the first instance, from occupants' behaviours along with the principle deterministic processes influencing heat flows in buildings, (ii) quantifies and decomposes the total uncertainty into its respective sources, and the interactions between them, and (iii) achieves this in a computationally efficient manner. This thesis introduces a new Emulation based Uncertainty and Sensitivity Analysis (EmUSA) framework which addresses this exact challenge.

1.2 Research Hypothesis and Objectives

Coupling a framework for simulating occupants' stochastic behaviours with BPS tools can considerably increase the computational complexity in the sim-

ulation of buildings' energy performance; particularly if repeated simulations are required to adequately quantify these stochastic influences. For a fixed input $\mathbf{X}$, repeated runs of such a *stochastic* simulator will produce different outputs - yielding an ensemble of prediction values instead of a single fixed value, as with a conventional *deterministic* BPS *simulator*. These prediction ensembles now represent a probability distribution of the simulator output, associated with each input $\mathbf{X}$. A standard approach to describe the stochastic response (for each fixed input $\mathbf{X}$) is via

$$Y = f(\mathbf{X}, \omega) \tag{1.1}$$

where ω is a variable that we introduced to denote the intrinsic aleatory uncertainty in the simulator. This new variable can be thought as an uncontrollable parameter (or seed variable) encoded within the stochastic simulator. The distribution of this uncontrollable variable is often unknown and, consequently, standard metamodeling techniques, based on sampling the (known) input space can no longer be applied in a straightforward fashion to emulate/infer the functional form of f. Along with that, absence of computationally efficient approach for the propagation of uncertainties through complex [dynamic + stochastic] building performance simulator and lack of readiness of the global uncertainty and sensitivity analysis methods for the decomposition of both epistemic and aleatory uncertainties constitute the basis for this thesis research hypothesis.

Hypothesis

> Efforts to address the performance gap are severely constrained by the limitations in characterisation and propagation of the uncertainty in building performance predictions arising from the joint consideration of deterministic and stochastic phenomena, in a computationally tractable manner.

Objectives

To develop and test this hypothesis the following objectives are formulated:

Objective I. Establish a selection strategy for an equivalent metamodel / emulator efficiently approximating the simulations of stochastic energy per-

formance of buildings. This requires a range of metamodelling techniques to be reviewed selecting the most promising one. The primary criteria for the selection are that:

a. the emulator should be parsimoniously structured: it should be as simple as possible, but no simpler.

b. the emulator should be plausible enough in conveying underlying non-linear behaviour of the dynamic BPS program in a transparent way by allowing a flexibility in fitting semi or non-parametric prior to the training dataset exhibiting model behaviour.

Objective II. Develop and validate an emulator, that accounts for occupant-related stochastic variability, against the test dataset using a predictivity coefficient measure. To achieve this objective, it is necessary to:

a. Represent occupants' variability i.e. the stochastic perturbations in energy predictions into emulator estimation procedure.

b. Construct a first emulator to fit empirical mean of observations representing mean component and

c. a second emulator to fit empirical variance of observations representing variance component.

d. Validate the joint metamodel (above pair of emulators) at new test points.

Objective III. Demonstrate the plausibility and robustness of emulator formulation in emulating detailed deterministic and stochastic thermal phenomena occurring in a Building. This objective constitutes a proof of principle for the demonstration of Emulation based Uncertainty and Sensitivity Analysis (EmUSA) framework methodology considering the constrained scope, in terms of the number of design variables and the range of occupants' behaviours investigated as well as the complexity of the adopted use case. To accomplish

this objective, the fitted emulator is required to:

a. Provide uncertainty due to emulation itself at training and test points.

b. Predict and capture the mean and random response of stochastic BP simulator respectively.

c. Propagate input parameter and stochastic uncertainties.

d. Quantify and decompose the total uncertainty in energy predictions into respective uncertain sources - expressed in terms of global sensitivity indices.

Objective IV. Extend the scope of applicability of new EmUSA framework, to account for additional and more complex physical variables (e.g. the composition and thermophysical properties of construction materials) and greater degrees of occupants' interaction freedom (e.g. use of lights). To handle this enlarged scope, the framework is required to:

a. Incorporate a parameter screening mechanism to rank the parameters in order of their influence exhibited by them on the energy performance predictions.

b. Conduct uncertainty quantification and decomposition study on the most influential parameters, obtained via previous parameter screening step, in computationally efficient manner.

1.3 Knowledge Contribution

The primary contribution of this thesis is the development of a novel and foundational framework, for the first time in building simulation studies, to perform uncertainty quantification of a stochastic building energy performance. This new framework for the quantification and decomposition of uncertainties in complex [stochastic and dynamic] building performance simulation is called Emulation based Uncertainty and Sensitivity Analysis (EmUSA) framework.

The EmUSA framework facilitates : (i) simulation of principle deterministic processes influencing heat flows in buildings and the stochastic perturbations to them, (ii) quantification and decomposition of the total uncertainty into its respective sources, and the interactions between them, and (iii) achieves this in a computationally efficient manner.

The current and future practical applications of building performance simulation exercises are required to provide probabilistic estimations of output quantity of interest, since a single deterministic value is often implausible in predicting the actual building performance. For that reason, there is a clear present-day need experienced in research and industry community for a shift in paradigm from deterministic to probabilistic building performance simulations. In presenting probabilistic predictions, the simulation exercises are expected to be: 1) robust in accounting for the diverse sources (parameter and model form) and nature (deterministic and stochastic) of uncertainties that lead to probabilistic estimations, 2) rigorous in mathematical characterisation and handling of entire ranges of uncertainties simultaneously, given that there may be interaction effects on output variables and 3) complete with building performance uncertainty and sensitivity analytical tool routinely integrated into them. To this end, the knowledge contribution in this thesis is aimed at providing computationally efficient means to effectively perform Uncertainty Analysis of building performance so that it may further augment the rigour and completeness of present and future practical applications of building simulations [1].

1.4 Thesis structure

In addressing the objectives the remainder of the thesis is structured as follows.

Chapter 2 describes different deterministic and stochastic phenomena in-

[1] It is important to place this thesis contribution in the wider context. A parallel relating to metamodelling / emulation of stochastic models can be drawn with the applications in other fields of research on oil reservoir ([40] demonstrated in [41]) and groundwater flow modelling ([42]), for the purpose of computationally efficient estimation of model outputs for uncertainty and sensitivity analysis. However, so far until now, the UQ computational challenge for an instance of stochastic BPS model has not been addressed. The addressing of the challenge in this thesis therefore enables building simulation research community to perform uncertainty and sensitivity analysis of building performance in a computationally efficient manner.

teracting within the heated building envelope and presents their coupling as Stochastic Building Performance Simulation (S-BPS) platform - a co-simulation platform coupling EnergyPlus with a multi-agent stochastic simulator called No-MASS, developed in Chapman [43].

Chapter 3 begins with a literature review on Uncertainty and Sensitivity analysis methods (both qualitative and quantitative) and their applications in BPS studies. The chapter also presents an overview of standard uncertainty treatment methodology in BPS - from sampling and uncertainty propagation through the simulation model and quantification of the total uncertainty in energy predictions and finally its decomposition into the respective sources that caused it. The later part of the chapter discusses about the computational expense in the propagation of uncertainties, the type of impact (main and / or interaction sensitivity indices) of uncertain parameters assessed and the way (local or global sensitivity analysis) they are computed. Based on the literature review and its synthesis, the chapter further summarizes and highlights the required improvements in the standard methodology in order to be able to perform UQ study for Complex [Stochastic + Dynamic] BPS.

Chapter 4 first reviews the alternative approaches for estimating metamodels of S-BPS, selecting Gaussian Process Regression (GPR) since it stands out against the performance criteria (accuracy, robustness, efficiency, transparency and simplicity) as discussed therein (addresses **Objective I.** of the thesis). The second part presents the development of a mathematical framework based on GPR accounting occupants' related stochastic variance in an emulator / metamodel estimation procedure. The framework is validated on a proof of principle case constituting a monozone office model enabled with stochastic occupant related phenomena adopted from Haldi and Robinson's 2011 study [28] (addresses **Objective II.** of the thesis). The third part explains the computational framework behind the new EmUSA framework with respect to sampling design, parametrization of EnergyPlus Input Data Files (IDFs), automation in running S-BPS (EnergyPlus + No-MASS) on High Performance Computing (HPC) machine, construction of GP emulator of S-BPS and computation of Sobol sensitivity indices.

Chapter 5 is the first application chapter on demonstration of the viability of an EmUSA framework for the computationally efficient quantification and

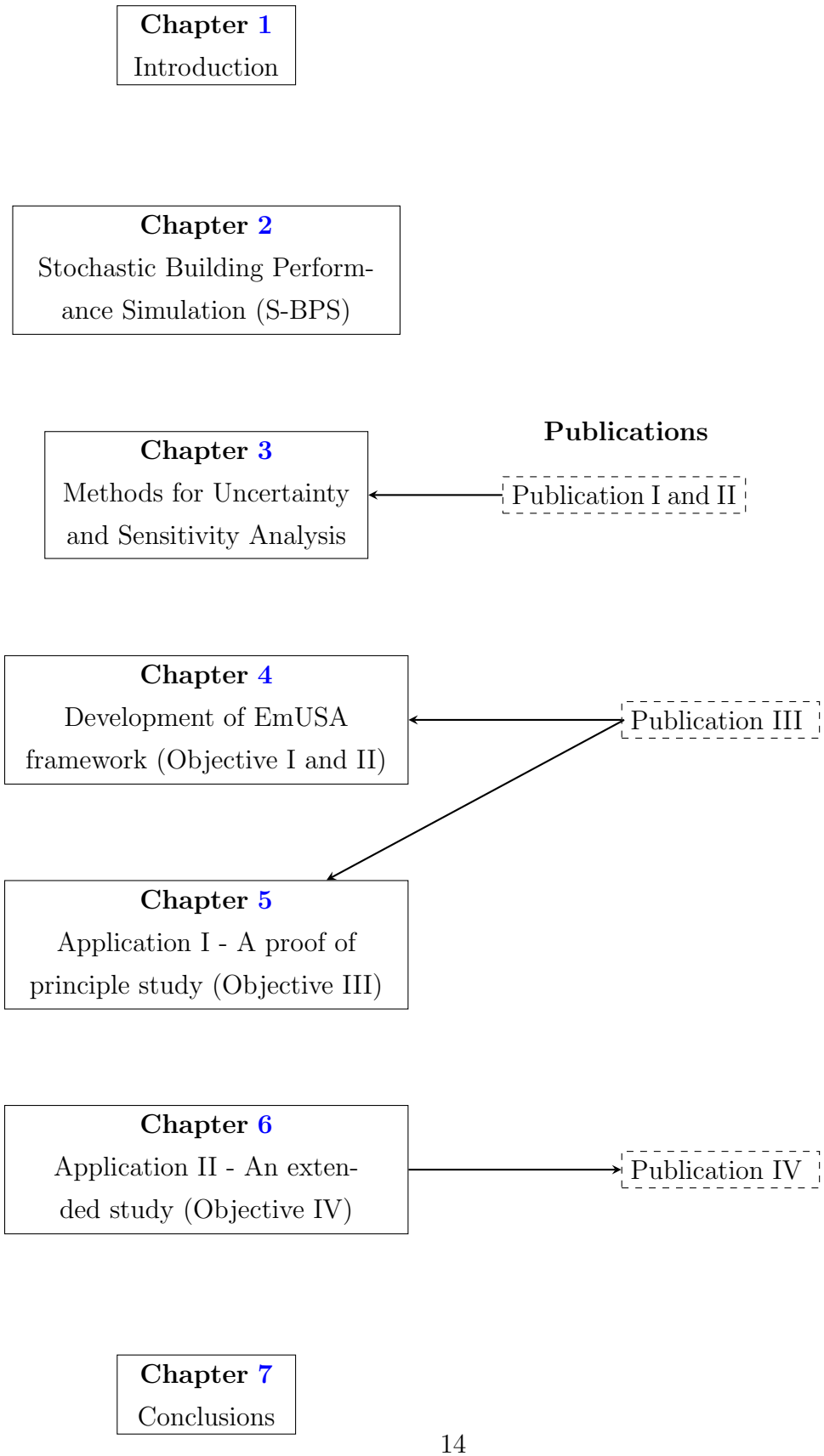
decomposition of both epistemic and aleatory uncertainty, albeit with a limited scope using a monozone hypothetical office building case study (addresses **Objective III.** of the thesis).

Chapter 6 is the second application chapter on proving the fit-for-purpose of new EmUSA framework for the extended scope and complexity due to additional and more complex physical variables (e.g. thermophysical properties of building envelope materials) and greater degrees of occupants' interaction freedom (e.g. use of lights) using a multizone residential case study building (addresses **Objective IV.** of the thesis).

Chapter 7 concludes this thesis by discussing the potential complementary applications of this work and as well as a foundational role it may play in future contributions on UQ studies in BPS.

The thesis outline is given as below along with its links to the publications that are arising from the thesis.

1.4.1 Thesis outline



1.4.2 Publications

I Wate, P., Coors, V., Robinson, D. and Iglesias, M., 2016. Qualitative Screening Method for Impact Assessment of Uncertain Building Geometry on Thermal Energy Demand Predictions. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 42,* p.127-134.

II Wate, P., Coors, V., Iglesias, M. and Robinson, D., 2019. Uncertainty assessment of building performance simulation: An insight into suitability of methods and their applications. In *Urban Energy Systems for Low-Carbon Cities* p.257-287. Academic Press.

III Wate, P., Iglesias, M., Coors, V. and Robinson, D., 2020. Framework for emulation and uncertainty quantification of a stochastic building performance simulator. *Applied Energy, 258,* 113759.

IV Wate, P. and Robinson, D. Emulation-based Uncertainty and Sensitivity Analysis framework for Dynamic Simulation: recent developments (*Abstract accepted for full paper submission in Building Simulation and Optimization IBPSA-England Conference, Loughborough University, UK (September 2020)*).

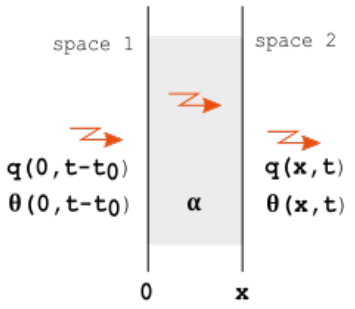
Chapter 2

Stochastic Building Performance Simulation (S-BPS)

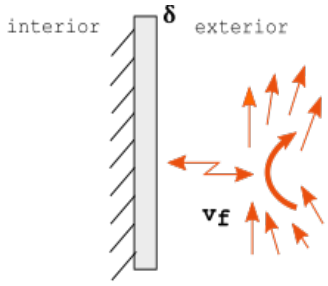
Building consists of different heat transfer and storage (e.g. opaque solid wall and transparent window) components exhibiting interactive and dynamic behaviour with other such components. Clarke [44] defines the thermal behaviour of building as a mathematical system of numerous intra- and inter-thermodynamic processes (e.g. from inside radiative components due to occupants to outside / adjacent zone radiation) which evolve at varying rates during the course of simulation when subjected to the external weather conditions. These processes represent different thermal energy flow paths through heat carrying components and are governed by the law of conservation of thermal air mass, momentum and energy. In principle, the heat transfer and exchange between the building components occurs due to the conduction, convection and radiation of thermal energy. Figure 2.1 illustrates these heat transfer and exchange mechanisms occurring inside the building envelope as extensively described in Clarke [44]. The mechanisms are briefly summarised as below:

2.1 Deterministic phenomena

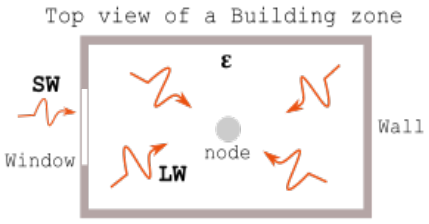
The deterministic phenomena are represented by the system of differential algebraic equations as described in Clarke [44]. **Transient conduction** of



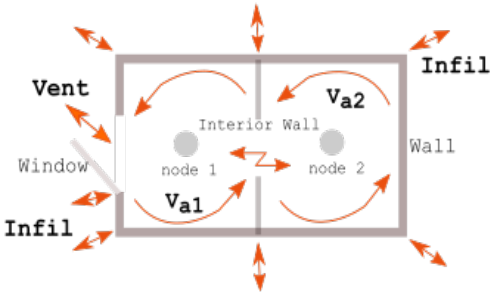
Transient conduction through construction composition



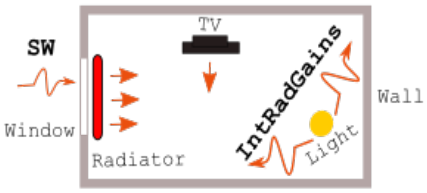
Surface convection due to flow fluid (air)



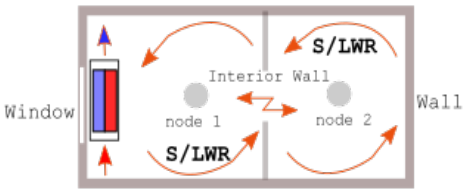
Short- and Longwave (SW & LW) Radiation exchange inside the building envelope



Air mass exchanges between the zones and advection due to ventilation and infiltration



Internal convective and radiative heat gains due to Lighting, Electrical equipments and heating systems



Interzonal short-,longwave radiation transfers and air circulation in dynamic phase change material

Figure 2.1: Thermal Energy flow paths (deterministic phenomena)

heat through an opaque solid wall is a function of applied and diffused heat excitations q and temperatures θ at its surface. The amount of heat conduction depends upon the thermal inertia of the construction composition layers which are characterized by hygro-thermophysical properties such as conductivity $\lambda(W/(mK))$, density $\rho(kg/m^3)$, specific heat capacity $C(J/(kgK))$ and the layer thickness $x(m)$. The two inter-relationships i.e. the change of temperature and flux with distance are :

$$\frac{\partial\theta(x,t)}{\partial x} = \frac{-1}{\lambda}q(x,t) \tag{2.1}$$

$$\frac{\partial q(x,t)}{\partial x} = -\rho C\frac{\partial\theta(x,t)}{\partial t} \tag{2.2}$$

Combining these gives the underlying heat conduction equation, the Fourier heat equation:

$$\frac{\partial^2\theta(x,t)}{\partial x^2} = \frac{1}{\alpha}\frac{\partial\theta(x,t)}{\partial t} \tag{2.3}$$

Alternatively, the derived parameters of thermal diffusivity $\alpha = \lambda/\rho C(m^2/s)$ and effusivity $\epsilon = \sqrt{\lambda\rho C}(Jm^{-2}K^{-1}s^{-1/2})$ can also be used to assess the dynamic variation of heat fluxes at conducting surfaces. The Fourier heat equation can be solved to obtain heat flux and temperature at time t and distance x through the construction component using response function (analytical method) and numerical modelling approaches.

Surface convection of heat flux due to the flow of air (characterized by Nusselt and Reynolds number for natural and forced convection) over an exposed surface occurs at both the internal and external surfaces of envelope. At the external surface, the convective heat flux is induced due to the striking wind depending upon its speed and direction. While at the internal surface, it is because of the natural or forced ventilation, where it depends on the speed and direction of the fluid vector. The amount of heat exchange is computed based on the time and position dependent convection coefficient $h_c(Wm^{-2}K^{-1})$, surface-air temperature difference $\theta_A - \theta_S(K)$, surface roughness and its area $A_s(m^2)$:

$$q_c = h_c A_s(\theta_A - \theta_S) \tag{2.4}$$

In case of natural convection where the conducted heat is carried in a zone due to buoyancy forces caused by density variations due to fluid temperature variations, the heat convection coefficient is estimated from the dimensional groupings of Nusselt Number (Nu), surface-to-air temperature difference and the empirical data obtained from the experiments involving natural convection heat transfer from the heated plates. The dynamic surface conditions evaluated from the Computational fluid dynamics (CFD) simulations can be used to determine the time varying heat convection coefficients and hence, subsequently the convective heat exchanges in a zone. On similar lines, in case of forced convection where the fluid motion is caused by fan or wind, the heat convection coefficient (for internal and external surfaces) is determined based on the Reynolds Number (Re), the empirical data, fluid velocity (V) and the temperature differences ($\delta\theta$). The dynamic fan operations and wind speed and directions data can be used to establish forced convection heat exchanges in and outside of the zone.

Incoming **shortwave solar radiation** incident on the building envelope is partly absorbed and reflected by the opaque surfaces and in addition transmitted by the transparent surfaces depending upon their optical properties of absorptivity, reflectivity and transmissivity. The heat gains contributed by the absorbed shortwave radiation complements the transient conduction in surfaces resulting into increased surface body temperatures which subsequently affects the surface convection and results into the longwave radiation exchanges. The estimation of incident shortwave radiation depends on the modelling of three components: (1) surface geometry in the context of its spatial location; (2) shadowing pattern caused by the surrounding buildings and (3) hourly/monthly computation of direct beam, ground reflected and sky diffuse insolation incident on the exposed surfaces. The reflection, absorption and transmission of shortwave solar radiation within transparent media and the intra-zone transfer modulated by the surface absorptivity and shading factors of opaque media dynamically redistribute the radiation within the zone.

A part of the shortwave radiation absorbed by the building facades is emitted in the form of **longwave radiation** (wavelength between 0.28 and 2.8μm) which contributes to a significant variation in the internal and external surface

temperatures. The longwave radiation exchange with the indoor environment by opaque and transparent surface elements depends upon their emissivities, temperatures, reflection coefficients and geometric form factor constituted by interior surfaces. The longwave radiation is responsible for introducing a non-linear behaviour in zone's energy balance due to it's higher order variation with respect to body's temperature.

$$q = \epsilon \sigma A \theta^4 \tag{2.5}$$

where, q is the grey body i.e. building facades' radiation flux (W), ϵ the temperature dependent surface emissivity, σ the Stefan-Boltzman constant $(1.3 \times 10^{-23} W m^{-2} K^{-4})$, A the surface area and θ the building facades' body temperature (K). Likewise, the external longwave radiation exchange depends upon the temperature profiles of sky, ground and surroundings and their respective view factors for the external surface. The net longwave exchange between the buildings' external facades and surroundings is given as the difference between the emitted and received flux:

$$q = A_s \epsilon \sigma (\theta_e^4 - \theta_s^4) \tag{2.6}$$

where, A_s is the surface area (m^2), θ_s the absolute temperature of the surface (K) and θ_e the equivalent temperature function of the temperatures of the sky, ground and surroundings:

$$\theta_e^4 = f_s \theta_{sky}^4 + f_g \theta_{grd}^4 + f_u \theta_{sur}^4 \tag{2.7}$$

where, f_s, f_g and f_u are the view factors to the sky, ground and surroundings respectively. The temperatures are estimated from the given weather data.

Another phenomenon of **advective heat exchange** between the air volumes at different temperatures dominates when there is a frequent intentional openings of windows by the occupants (referred as **natural ventilation**) and deliberate supply of air for space conditioning (referred as **mechanical ventilation**). The air flows caused are the result of pressure variations and buoyancy forces due to density differences associated with the temperatures of the air volumes. The **infiltration** of outside air due to fabric leakages and small cracks around window and door openings also affects the heat flux transfers

along with the latent heat contributed by the external moist air. As these phenomena are been partly triggered by the occupants' interaction as well, an another layer of stochastic perturbations due to models of occupants' behaviours is added to the energy demand predictions. Figure 2.2 gives a systemic illustration of the deterministic and stochastic energy phenomena occurring inside the buildings and the parameters on which they depend upon.

EnergyPlus, a thermal building simulation program is a numerical solver that presents modelling techniques (such as transfer function or numerical finite difference methods) to obtain the solutions of system of differential equations [45]. The schematic (Figure 2.2) shows the interactions between different deterministic phenomena via their interconnections and their specific input data requirements in order to perform energy load computations.

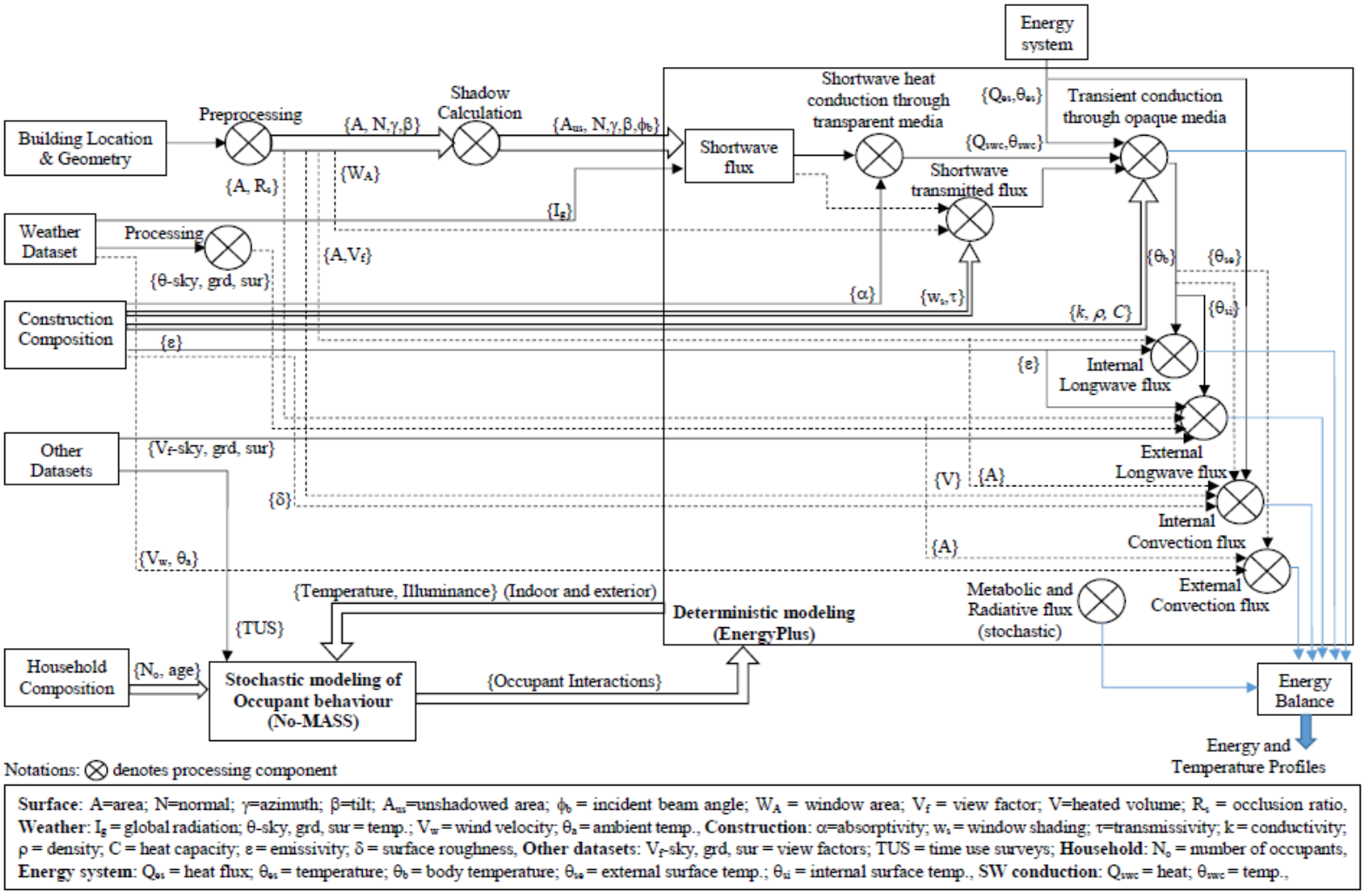


Figure 2.2: Systemic illustration of thermal energy flow paths and their interactions: Deterministic and Stochastic sub-systems of BPS

2.2 Stochastic phenomena

The presence of occupants and their interactions with the building envelope (such as window opening / closing) contributes to the internal heat gains in the form of metabolic heat gains and other casual gains (also due to use of electrical equipment and lighting). Such form of heat gains cannot be predicted in that straightforward way as they, in principle, are governed by the stochastic nature of humans and their behaviours. Significant progress has been made in recent years in the modelling of occupants' behaviours, employing a range of techniques. A large number of data-driven stochastic models have been developed to model occupants' presence (or activities and dependent presence in the case of residences) and dependent interactions with the building envelope and its systems. The underlying phenomena are typically represented as a Bernoulli process [predicting the probability that a particular state will be observed, such as a window being open], as a discrete-time random or Markov process [predicting the probability with which a state transition will take place], as a continuous-time random or survival time process [predicting the duration for which a particular state will survive], or hybrids of them [46].

2.2.1 Stochastic models of Occupant behaviours

Occupant presence and metabolic gains

More specifically, Chapman [2] employs two separate methods for estimating occupant presence and corresponding metabolic gains in case of non-residential and residential buildings distinctively. For non-residential buildings, the presence model [47] (used for case study 1 on monozone office building) is based on an inhomogeneous Markov chain predicting when an occupant is present within the building using a mobility parameter μ and a time-dependent profile of the probability of presence $P(t)$. The sequence of presence and absence is generated by calculating the transition probability at each time step,

from absent to present:

$$T_{01}(t) = \frac{\mu - 1}{\mu + 1} P(t) + P(t + 1) \tag{2.8}$$

where μ the mobility parameter is by default held constant at an empirically

calculated value of 0.11.

present to present:

$$T_{11} = \frac{P(t)-1}{P(t)}[\frac{\mu-1}{\mu+1}P(t) + P(t+1)] + \frac{P(t+1)}{P(t)} \tag{2.9}$$

the other transitions simply are,

present to absent:

$$T_{10} = 1 - T_{11} \tag{2.10}$$

absent to absent:

$$T_{00} = 1 - T_{01} \tag{2.11}$$

In case of residential buildings, the activity model [22] (used for case study 2 on multizone residential building) predicts an activity to be performed from a number of home activities (sleeping, passive, audio/visual, IT, cooking, cleaning, metabolic, washing appliance use, personal washing and absence) as a time-dependent Bernoulli process based on multinomial logistic regression.

$$P(x,t) = \frac{\exp(A_j(x,t))}{\sum_{j=1}^{N}\exp(A_j(x,t))}, j = 1,\ldots,N \tag{2.12}$$

and

$$A_j(x,t) = \alpha_j + \sum_{k=1}^{n}\beta_{jk}x_{jk} \tag{2.13}$$

where, t is time, N the total number of activities, n the number of predictors, and x the predictors, each associated with a slope coefficient β_{jk} and α_j the intercept.

Equation 2.12 has been used to calculate the time-dependent probabilities for the number of presence-dependent occupants' activities. An activity is assigned to the occupant / agent by applying Inverse Function Method (IFM) on the probability distribution function of time-dependent probabilities corresponding to the activities. In IFM, a random number between 0 and 1 is drawn and is evaluated against the likely activity to it corresponds in a cumulative distribution function. For instance, for the residential living room during a typical evening hour, equation 2.12 has estimated following sequence of probabilities (Table 2.1) based on the time use survey dataset. It is more likely

Table 2.1: Exemplar time-dependent probabilities for the occupants' activities during a typical evening hour in residential living room

Activity	Probability	Cumulative Distribution Function (CDF)
sleeping	0.05	0.05
passive	0.2	0.25
audio / visual	0.3	0.55
IT	0.1	0.65
cooking	0	0.65
cleaning	0.1	0.75
metabolic	0.25	1
washing appliance use	0	1
personal washing	0	1
absence	0	1

that the occupant will be spending this evening hour either passively (reading books), enjoying music or TV programs or any other light / moderate metabolic activity such as talking while seated / standing (perhaps gossiping with partner about office colleagues!).

Let us say that three random numbers 0.13, 0.52 and 0.78 are drawn between 0 and 1 for the activity random variable. Using IFM, as 0.13 falls in 'passive' bin (0.05 and 0.25), the 'passive' activity is selected. Similarly, 0.52 corresponds to 'audio/visual' and 0.78 to 'metabolic' activities. Similar IFM procedure has been employed for transition probabilities in case of non-residential buildings where there are only two states / activities, occupants' presence and absence are under consideration. The metabolic gain is calculated using Fanger's PMV model (described in ISO 7730 [48]) given the standard physical (air and radiant temperatures, relative air velocity and humidity supplied by EnergyPlus) and personal (clothing level and metabolic rate values from ISO 7730 state-dependent activities) parameters. The total metabolic heat gain is then dependent upon above stochastically estimated zone activity schedule, corresponding metabolic rate values and a number of occupants present.

Window Interactions

The hybrid model by Haldi and Robinson [1] employs a combination of Markov chain (for the predictions in transition to window opening) and Weibull distri-

bution (for the prediction of duration for which a window remains open). The probability of transition to window opening state is estimated using a logistic regression model:

$$P_{ij}(x_1, ..., x_n) = \frac{\exp(\alpha + \sum_{k=1}^{n} \beta_k x_k)}{1 + \exp(\alpha + \sum_{k=1}^{n} \beta_k x_k)} \tag{2.14}$$

At occupants' arrival, the predictors x are indoor and outdoor temperatures, absence duration and rain presence. When zone is occupied, the absence duration predictor is replaced by length of presence. And at occupants' departure, the length of departure replaces the length of presence and the rain presence is replaced with a dummy parameter for the window if it is on the ground floor (to account for the probability of windows being open during night). EnergyPlus provides indoor and outdoor temperatures while No-MASS calculates temporal stochastic occupants' presence schedule and the ground floor parameter for window is a boolean input (true for the zone at ground floor) to No-MASS during runtime defined in a simulation configuration file. The exchanges of data values for these variables happen during simulation run time at each time-step via Functional Mockup Interface (FMI). The values of regression coefficients α, β_k are calculated from the aggregated population datasets. The Weibull distribution estimating the duration for which a window remains open is given by the probability density function as:

$$\lambda\alpha(\lambda t)^{\alpha-1} \exp(-(\lambda t)^{\alpha}) \tag{2.15}$$

where the parameter $\alpha = \exp(-0.871)$ is called shape and $\lambda = \exp(-(a + b\theta_{out}))$ the scale, $a = 2.213, b = 0.173$ and θ_{out} is the outdoor temperature given by EnergyPlus. Following process flow (Figure 2.3) describes the algorithm implemented in No-MASS to determine the occupants' presence-dependent window opening (1) and closing (0) at each time step. The window interaction schedule is then an external schedule for respective EnergyPlus variable set by No-MASS predicted stochastic schedule via FMI.

External Shading Interactions

A similar hybridisation of occupants' shading interactions has been modelled by Haldi and Robinson [49]. Upon occupants' arrival in a zone, the model

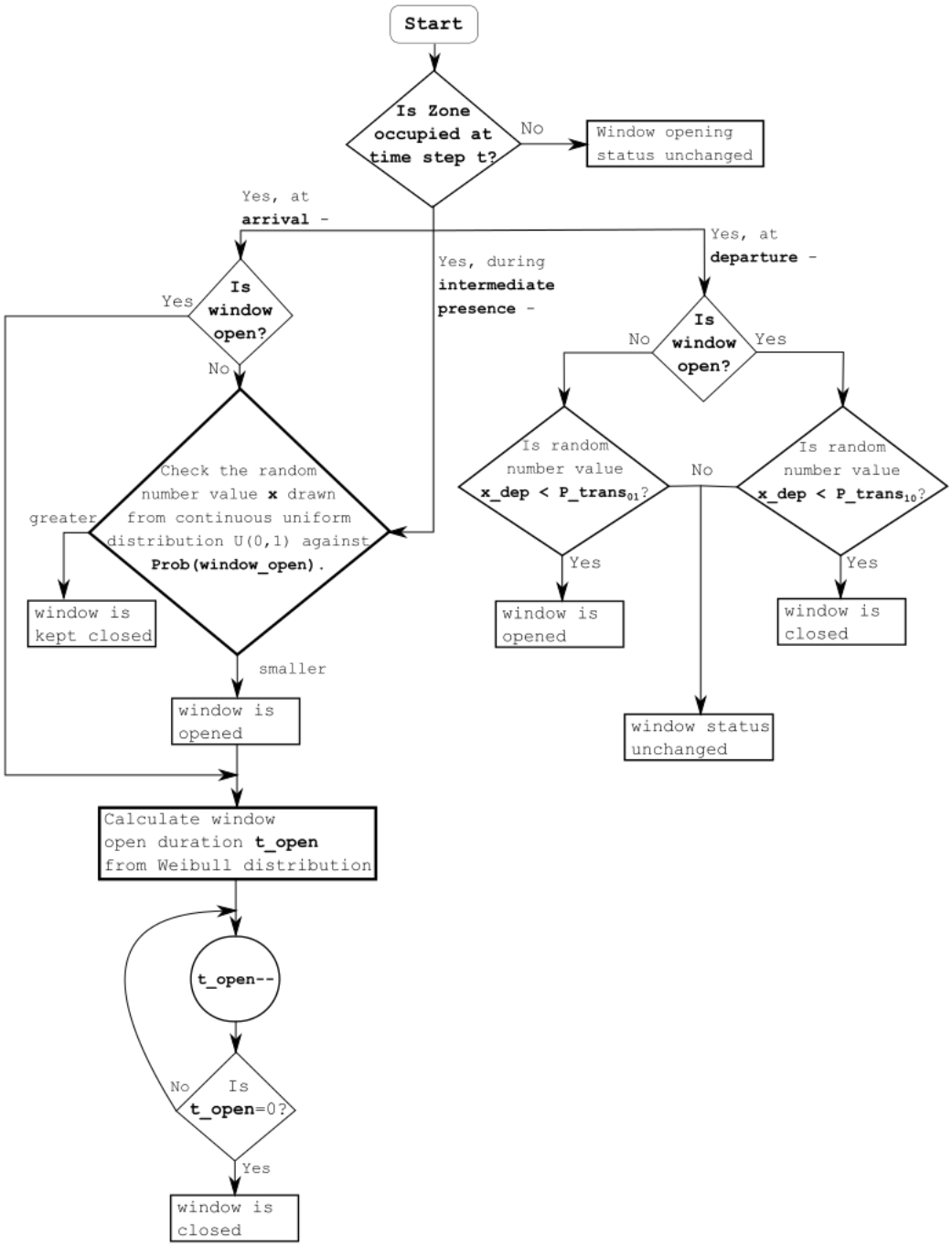


Figure 2.3: Occupants' window interaction algorithm workflow after hybrid model by Haldi and Robinson [1]

predicts lowering or raising of shades given the local visual stimuli i.e. indoor illuminance and the amount of unshaded fraction of window at the previous time-step. The probability with which a lowering or raising of shade will take place at next time step is given as:

$$P_{act}(E_{in}, B_L) = \frac{\exp(a + b_{in}E_{in} + b_L B_L)}{1 + \exp(a + b_{in}E_{in} + b_L B_L)} \tag{2.16}$$

where E_{in} is the indoor illuminance at an appropriate daylight reference point within the zone provided by EnergyPlus and B_L the unshaded window fraction at the previous time-step. At each time-step, a random number is drawn from uniform distribution and checked against this probability to predict the (raising or lowering) action that will take place. Upon predicting the action, whether the shade is fully raised or lowered is determined as:

$$P_{fullact}(E_{gl,hor}, B_L) = \frac{\exp(a + b_{out}E_{gl,hor} + b_L B_L)}{1 + \exp(a + b_{out}E_{gl,hor} + b_L B_L)} \tag{2.17}$$

where $E_{gl,hor}$ is the outdoor global horizontal illuminance. The regression parameters b_{in}, b_L and b_{out} are calculated using aggregated population datasets (see Haldi and Robinson [49]). When the shades are partly raised or lowered, their fractions are predicted using a Weibull distribution:

$$f(\Delta B) = \lambda\alpha(\lambda\Delta B)^{a-1} exp(-(\lambda\Delta B)^{\alpha}) \tag{2.18}$$

where the shape parameter $\alpha = 1.708$ and scale

$$\lambda = \frac{1}{\exp(-2.294 + 1.522B_{L,init})} \tag{2.19}$$

The shading fractions from these models provide then the schedule that to be set in EnergyPlus externally. Chapman's [43] modified EnergyPlus source code was used to account for the fraction of shade that is raised / lowered in order to proportionately reduce the amount of radiation transmitted through the window (see Appendix B in [43] for the corresponding changes in the source code).

Lighting actions

No-MASS employs Lightswitch-2002 algorithm by Reinhart [50] to predict the use of lights. The action of turning the lights on depends on the amount of indoor illuminance E whilst an occupant is present within the zone. For each time-step, the switch on probability is given as:

$$P = \frac{a + c}{1 + \exp[-b(\log_{10}(E) - m)]} \tag{2.20}$$

where E is an indoor illuminance of the zone provided by EnergyPlus. At arrival, $a = -0.0175, b = 4.0835, c = 1.0361, m = 1.8223$ and whilst present $a = 0.0027, b = 0.017, c = -64.19, m = 2.41$. At the time of departure, the probability of turning lights off is a function of the anticipated duration of occupants' absence in hours D in the range 0.5 and 12 (after Eilers and Reed, [51]), given as:

$$P = 0.268 \times \ln(D) + 0.259 \tag{2.21}$$

When an occupant is going to be absent for more than 12 hours, the lights are assumed to be turned off. And for the absences less than 0.5 hours, the lights remain on. This No-MASS predicted lighting status at each time-step is set as a lighting schedule within EnergyPlus for each building zone.

Other techniques are needed to model phenomena for which data is not abundant. These include belief-desire-intention (BDI) and agent-learning frameworks, both applied to agent-based representations of occupants [52]. In the former case, agents evaluate their beliefs about the state of the environment in which they are immersed, comparing these with their desired states and effect plans intended to achieve their desires. In the latter case, agents progressively learn, through repeated simulations, behavioural interactions that maximise rewards. BDI rules are suited to relatively simple interactions, whereas agent learning is able to accommodate considerably more complex behaviours, though this can come with a penalty in terms of the number of computations needed to learn reward-optimal behaviours. Finally, and as noted by Robinson et al. [46], it is also desirable to represent social interactions amongst occupants' sharing spaces, emulating their negotiation processes and the outcomes (agreed interactions) arising from them. Chapman et al. [52] describe a vote casting and processing mechanism to achieve this in the case of data-driven

models. Similar mechanisms are encodable in the case of BDI rules.

2.3 S-BPS Platform : A coupling of phenomena

Several computational platforms have been developed to couple stochastic models of occupants' behaviours within BPS tools, including those of Langevin et al. [53], Hong et al. [54] and Chapman et al. [22]. Langevin et al's platform integrates rules (akin to BDI rules) that allow agents to modify their clothing and the use of windows, fans and heaters (and corresponding setpoints) to achieve their comfort desires. This is integrated with the EnergyPlus solver using the Building Controls Virtual Test Bed (BCVTB). Hong et al's platform is in principle more general in character, employing an XML schema to structure the modelling approaches to be employed for a particular application, populated with the parameters to be employed by this structure, which is then solved using obFMU. The Functional Mockup Interface (FMI) co-simulation standard is employed to couple obFMU with EnergyPlus, with EnergyPlus acting as the master algorithm. To the author's knowledge, only Chapman et al's platform (No-MASS) incorporates all of the above mechanisms; in common with obFMU, utilising FMI to co-simulate with EnergyPlus.

As noted in the introduction chapter 1, the intention of the proof-of-concept and its extension demonstrated later in the thesis is that how can we effectively emulate a building performance simulator incorporating aleatory uncertainties arising from occupants' stochastic interactions, and to employ this simulator to quantify and decompose uncertainties. For this purpose we use EnergyPlus co-simulated with No-MASS, constituting our energy simulation platform called Stochastic Building Performance Simulation (S-BPS) (Figure 2.4). For the proof-of-concept (5) and its extended study (6), we restrict our scope to uncertainties relating to the physical properties (e.g. conductivity, density, specific heat capacity) of the building envelope materials and the stochastic behaviours of occupants, limited to their presence and associated use of windows and shading devices. For presence, we employ the time-inhomogeneous Markov chain model of Page et al. [47]; for the use of windows we employ the hybrid discrete- and continuous- time model of Haldi and Robinson [1]; while for

shading devices we employ a similar hybridisation, due to Haldi and Robinson [49], as described in the previous section. For further information regarding behavioural models and associated modelling techniques, the interested reader is referred to the reviews of Dong *et. al.* [55], Wagner *et. al.* [56], Gunay *et. al.* [57] and Robinson *et. al.* [46].

The stochastic models of occupants' behaviour involve Inverse Function Method (IFM) to predict a particular action / behaviour that would eventually take place given the probabilities for the occurrence of different behaviours. The random variable in IFM procedure allows to capture the stochastic nature of occupants' behaviours. This introduces a random bias i.e. an aleatory type of uncertainty due to occupants' behavioural randomness in the predictions of building energy performance. From a system modelling perspective of S-BPS platform and also as per formulation in section 1.2, the random variable representing occupants' behaviour cannot be accessed and controlled by the model user / analyst, unlike the case of material physical properties variables. For the purposes of Uncertainty and Sensitivity Analysis of energy performance in this thesis, the S-BPS platform has been considered as a systemic black box (Figure 2.2) having weather, building information modelling inputs, interacting deterministic thermodynamic heat transfer and stochastic occupants' behavioural phenomena resulting into the stochastic building performance predictions.

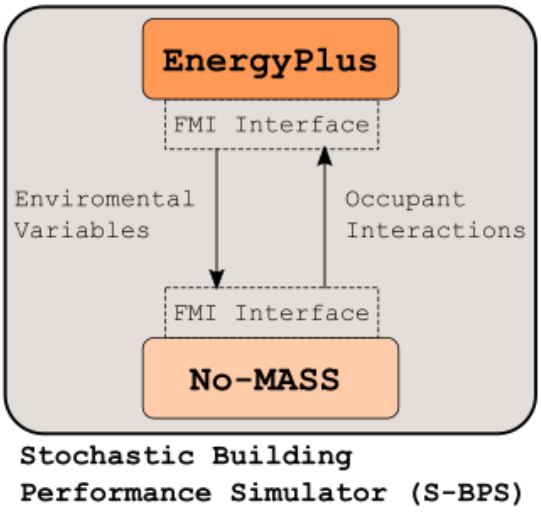


Figure 2.4: Schematic of EnergyPlus and No-MASS Co-simulation platform [2]: the Stochastic Building Performance Simulator (S-BPS)

Chapter 3

Methods for Uncertainty and Sensitivity Analysis

All models are wrong but some are useful.

George E. P. Box, Statistician
"Robustness in Statistics"
Workshop, 1978

As Box [58] pointed out, *since model designers / scientists cannot obtain a "correct" one, they tend to devise simple but evocative models.* He further argued that this not only applies to the statistical models but as well as to the scientific / physical models. Embarking on Box's thoughts, this thesis adopts an *uncertain* view about the modelling and simulation practices in evaluating Building Energy Performances and prepares the foundational step of assessing uncertainties in BPS for the related and relevant future contributions.

Complex prognostic models involve an integration of deterministic and stochastic processes such as the case of complex BPS involving interactions between heat transfer phenomena through the building envelope and stochastic internal gains due to occupants' behaviours. A resulting combined behaviour due to the involved complexity is thus unknown (unless and until the input values are provided and simulation is conducted) and perhaps uncertain - adding a layer of inherent model uncertainty to the predictions of quantity of interest. In addition to that, depending upon the nature and their availability, the model input data sources prone to measurement errors and scaling

issues also contribute to a systematic and random uncertainty to the model predictions. As a consequence, the prediction veracity and the confidence in predicting capability of the models is affected adversely.

Following this introduction, this chapter 3 presents the measures for uncertainty and sensitivity, the uncertainty treatment process flow, the qualitative and quantitative methods proposed in the literature and their applications in BPS. Finally, this chapter ends with a summary which forms the basis and motivation for a proposal of new framework in order to conduct Uncertainty and Sensitivity Analysis for Complex Building Performance Simulation.

3.1 Measures for Uncertainty and Sensitivity

Uncertainty in the model inputs and predictions is characterized as a mathematical structure of probability distributions represented by the respective input and output random variables. The statistical parameters (such as mean μ and standard deviation σ) of distributions determine the shape and value ranges of the probability distribution structure. The amount of uncertainty can be expressed as a variance measure [1]. Variance V measures the variability of the data values from the mean μ of the probability distribution. It is defined as the expected value $\mathbb{E}$ of squares of difference between the data values taken by random variable X and μ, given as -

$$V = \mathbb{E}((X - \mu)^2) \tag{3.1}$$

Sensitivity of an output quantity of interest to the given input variable i.e. a degree to which an output is sensitive to the input is defined as the ratio of change in model output to the corresponding change in model input. Mathematically, it can be expressed as the derivative of model output y with

[1] Alternatively, Shannon [59] (for discrete random variable) or differential (for continuous one) Entropy H can also be used to express the randomness or uncertainty in random variable. For our analysis in this thesis, we adopt *variance* as an uncertainty measure for the purposes of consistency of our approach with the related studies in literature. Also, due to the promising features of variance-based analysis i.e. model independent measure etc. stated in [60], pages 157-158.

respect to its input x_i [60].

$$S_{x_i} = \frac{\partial y}{\partial x_i} \tag{3.2}$$

The sensitivity of a variable / an input parameter is estimated in the form of sensitivity index in two ways: an analytical method of Taylor Series Expansion [61] and numerical methods based on Monte Carlo approach [62]. An analytical method is typically employed when the uncertainties in given inputs are assumed to be small and lie close to the mean and the sensitivity index is calculated at a local point (mostly at local mean). The conducted analysis is termed as **Local Sensitivity Analysis** (LSA). On the other hand, when the inputs are highly uncertain and the model has unknown linearity, the model independent numerical methods are employed which are also resilient towards highly uncertain inputs. These methods, although computationally expensive requiring large number of model executions, allow to explore the entire uncertainty range of input parameters so that the **Global Sensitivity Analysis** (GSA) is possible.

3.2 Sources of uncertainty

Figure 3.1 presents the generic overview of sources of possible uncertainties arising during different stages when conducting a simulation study. These sources relate to model, data inputs and algorithm (numerical approximation) uncertainties [63] [3]. A computer model of a reality of interest allows to abstract and represent the physical phenomena in order to better study and analyse the underlying behaviour, specifically for the cases where it is practically less feasible to conduct physical measurements / experiments. However, at the same time, depending upon the modelling expertise available, the process from defining a problem till experiment-ready design involves modelling assumptions and imposing of operating boundary conditions that - leads to simplifications and hence, discrepancy from the nature of actual underlying phenomenon being modelled. The input data preparation tasks constitute the second stage of uncertainty as the data requirements vary largely depending upon the type (detailed or simplified) of simulation model of underlying processes at hand. Finally, the process of translation of model formulation

to a computer code leads to approximation and which results in a deviating behaviour from the envisaged model functionality. The subsection 1.1.2 elaborates on the sources of uncertainty in the context of a building performance simulation study with respect to Figure 3.1.

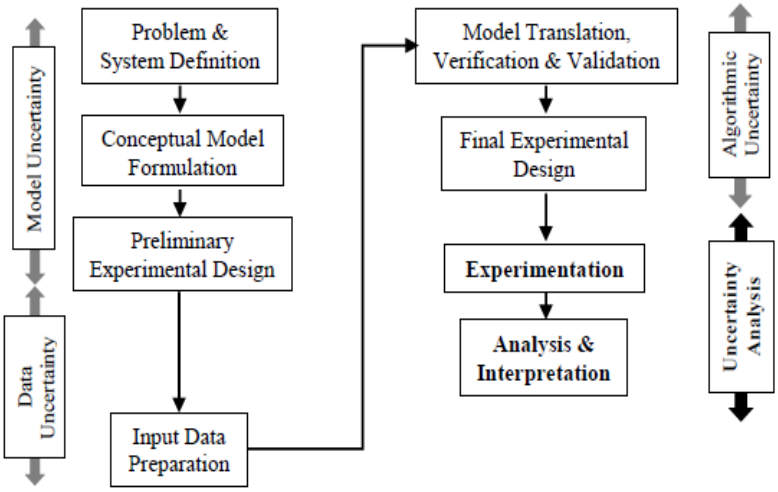


Figure 3.1: Sources of uncertainty at different stages of a simulation study [3]

Kennedy and O'Hagan [64] have proposed one of the pragmatic ways of classifying the various sources of uncertainty in experimental simulation studies involving complex mathematical models. The classification is based on the fundamental notion that - there exists a lack of knowledge about the exact values of some of the context-specific input parameters and initial conditions termed as *parameter uncertainty*. A similar notion applies to a computer model, since *all models are wrong* [58], when there exists a discrepancy between the model predictions and the observations of underlying real world phenomenon, called as *model inadequacy*. The other form of variation arises referred as *residual variability*, where the model takes on different prediction values every time when evaluated with same input values and conditions, due to inherent uncontrollable and stochastic nature of a particular phenomenon being modelled. Finally, in practice, the output of the computer code is unknown for given set of inputs until the code is run and observed what actually the output is. The complex models such as the case of stochastic BPS (S-BPS) are computation intensive and it is also not realistic to conduct model runs for every input configuration in order to know the nature of underlying function (which is governing the building performance simulation model) in an entirety.

Since the relationship expressed via mathematical model is complex (and it therefore necessitated the model implementation in a computer code), the uncertainty about the code output, termed as *code* or *algorithmic uncertainty*, requires consideration in simulation studies.

In the context of this thesis, random model discrepancy in the energy performance predictions i.e. an aleatory type of uncertainty due to inherent randomness in stochastic models of occupants' interactions and an epistemic type of uncertainty due to lack of precise knowledge about the data values for design and physical input variables have been considered in the case study applications.

3.3 Uncertainty Treatment : The process flow

Uncertainty analysis quantifies the amount of uncertainty in the model output due to the uncertainties in its inputs. **Sensitivity analysis** provides measures to decompose the total uncertainty in the model output into the respective sources of uncertainties that caused it i.e. the model inputs [65]. Uncertainty and sensitivity analysis techniques offer a pragmatic treatment to the uncertainty in the model predictions by providing better understanding of the model behaviour. The model input-output relationship can be investigated to reveal the additivity / non-additivity, linearity / non-linearity and monotonic / non-monotonic behaviour of the underlying model.

The treatment aids modellers in different ways [60]:

I to re-assure the robustness of model against assumptions made,

II to provide guidance on the prioritization of model improvement efforts by knowing the most influential inputs [2] called as *Factor Prioritization* setting (**FP**) and subsequent model simplifications by identifying least significant inputs called as *Factor Fixing* setting (**FF**),

III to identify critical regions in the input space which are responsible for the desired / target behaviour in the output space defined as *Factor Mapping* setting (**FM**) and

[2] which are resulting into greatest possible variance / uncertainty in the output Quantity of Interest (QoI) and when they are fixed to their true values, would reduce the variance / uncertainty in the output QoI the most.

IV to assess reliable ranges for total uncertainty in the predictions defined as *Variance Cutting* setting (**VC**) and as well as to know the justification behind the extreme model behaviour in order to bring the output uncertainty within certain tolerance range.

In the building simulation context, a prior and rigorous treatment of the uncertainty problem could allow to gain interesting insights, for instance to ascertain the assumptions made during early building design stages so that the impact of ignoring / neglecting a particular parameter / phenomena can be assessed. With **FP** setting, a building energy modeller would focus only on the most influential parameters for the optimisation of model fitness to test its applicability given the aim of reducing uncertainty i.e. achieving plausibility in low energy building designs, for instance. With **FF** setting, a modeller can then accordingly direct his / her model simplification efforts by being aware of the non-influential inputs. The simplified model (and thus computationally less demanding one) could then be employed to perform building design optimisation for minimisation of energy loads and carbon emissions and maximisation of occupant thermal comfort. By implementing **FM** setting, the ranges or critical regions in the input space that are influencing for instance, the peak load predictions can be identified. **VC** setting would then enable to reduce the energy *performance gap* by a way of limiting the uncertainty in predictions within the threshold limits with the cutting of variances i.e. improving on the acquisition of more certain values for the input data / model parameters.

The techniques for the treatment of uncertainty in simulation studies can be classified into qualitative and quantitative methods.

3.3.1 Qualitative methods

Screening methods, which can also be termed as Qualitative methods, present a ranking of the most influential inputs indicated by the sensitivity measures qualitatively without quantifying their actual influence in numerical terms, while the quantitative ones (explained in next subsection 3.3.2) compute input variable influences as an absolute numerical value. Qualitative methods are useful in reducing the dimensions of input space by segregating the most influential input parameters from the non-influential ones, specifically in case of

complex simulation models having several uncertain input variables and significant computation time. Then, the quantitative methods / treatment (section 3.3.2) is employed for the rigorous quantification and decomposition of uncertainties on a reduced set of most influential input factors. In this subsection, we review and present the discussion on screening methods / techniques.

The qualitative sensitivity analysis techniques consist of an experiment design and the screening method. An experiment design is a sampling procedure that allows to select the small number of points from the input distributions in such a way that they faithfully represent the whole of the input space. The computational model is executed at these evenly distributed sample points. Then, the screening method is employed to identify the most influential inputs or groups containing the influential inputs. In this thesis, for the case of S-BPS, since the number of simulations runs N to be conducted are more than the number of inputs k, the screening methods, that consider a single variable as an individual entity (unlike group screening methods) termed as individual screening methods, suffice the purpose. For group screening methods (where $N << k, k$ is large), the readers are requested to refer to the literature: two-stage group screening [66], multi-stage group screening [67], Jacoby and Harrisons sequential bifurcation [68], sequential bifurcation [69], Morriss group screening [70], fractional replicate screening design [71], grouping and analysis using iterated fractional factorial design, IFFDs [72].

In screening methods, the indicators about the influence of an input x_i on the output y are determined at the local point in input space. These indicators convey the model behaviour in terms of the **total** and **interaction effects**. While the total effect accounts for the overall influence of an input on the output, an interaction effect determines the higher order influence involving interactions of x_i with other inputs due to model non-linearity and / or non-additivity. The advanced and improved screening methods also provide an indicator to detect the model monotonicity. The utility of sampling design and method for qualitative ranking of inputs is evaluated against their effectiveness in finding the most important inputs with less number of model executions. Following subsections briefly review the features of different sampling designs and screening methods described in Saltelli et al. [60].

One-at-a-time (OAT) design

OAT is a local sensitivity analysis method in which the difference in model outputs is computed by varying the value of only one input at a time between the consecutive pair of simulation runs. Typically, the model is executed between the extremes in input space and the corresponding change in model outputs is computed. The ratio of change in model outputs to the corresponding change in inputs provides the amount of impact of that given input on the output. The design requires $2k$ simulations, one each at high and low values for k inputs. The computed individual effect is a reliable indicator of the measure of sensitivity of that input variable to the output, if the underlying model exhibits linear behaviour. OAT design has shown poor effect to simulation runs efficiency when a number of the most influential input parameters is small as compared to the number of inputs investigated.

Factorial design

A typical **full factorial design** with two-level for k inputs generates 2^k grid points in k- dimensional hypercube of input space. The output mean is computed for each level of the every input by conducting total 2^k simulations. One drawback of full factorial design is that the number of simulations required increases drastically even with a small increase in number of inputs k. On the other hand, the **fractional factorial design** is a two-level truncated full factorial design in which a measure of main effect of a given input on the output is computed as half the difference of average output values for the corresponding two levels of input. The variance in the output is obtained as the square of main effect. The magnitude of variance determines the impact of that input. Also in case of two-level fractional factorial design, as the model is evaluated only at the extremes of an interval, the non-linear variation exhibited by the model at the intermediate values is not well handled.

Morris design and method

Morris design is a type of an OAT design, where the smaller changes in input values are represented by the finite levels within an interval. The local measure of incremental ratios called as an elementary effect measure is computed at

each of these levels accounting for the whole input space. The measure values are averaged out to give the single effect value that is independent of the selection of sample points. Morris [73] proposed two sensitivity measures based on the elementary effect to determine whether the effect of x_i on y is negligible, linear and additive or nonlinear and non-additive. For a given input x_i, the elementary effect is defined as:

$$EE_i = \frac{[y(x_1, x_2, .., x_{i-1}, x_{i+\Delta}, .., x_k) - y(x_1, x_2, .., x_k)]}{\Delta} \tag{3.3}$$

where k independent inputs $x_i, i = 1, 2, ...k$ define k- dimensional input space discretised into a p- level grid Ω, Δ = multiple of $\frac{1}{p-1}$ for $p \geq 2$.

The random sampling of a vector X from the grid Ω on input space is used to generate the distribution of elementary effects F_i for every i^{th} input. If p is even, Δ is assumed to be $\frac{p}{2(p-1)}$ with number of effects EE_i in F_i is equal to $p^{k-1}[p - \Delta(p-1)]$. The mean μ_i and standard deviation σ_i estimates from the elementary effects distribution F_i represent sensitivity measures. While mean μ_i measures an overall influence of input on the output, the σ_i measures the collective effects of an input due to non-linearity and/or interactions with other inputs. In the simplest design (let's say, for 2-input variable shown in Figure 3.2), a computation of a single elementary effect requires two sample points, i.e. the economy of design [3] $= \frac{1}{2}$. Therefore, for r EEs per k inputs result in $2rk$ model executions. The economy of design can be improved by building r random trajectories i.e. computing r EEs per k inputs from $(k+1)$ sample points where each trajectory provides k EEs. So with efficient Morris design, the total simulation runs required are $r(k+1)$ with the economy of design $= k/(k+1)$. Figure 3.2 illustrates a trajectory for $k = 2, p = 4$ and $\Delta = \frac{2}{3}$. Here, the trajectory $r = 1, (k+1) = 3$ sample points and model runs $= 3$ at $X^{(1)}[1/3, 1], X^{(2)}[1, 1]$ and $X^{(3)}[1, 1/3]$.

The Morris method in the current form is prone to type II[4] statistical errors [5] - failing to identify an influential input. When the underlying model exhibits a considerable non-monotonous response, a mean μ_i would measure a

[3] A number of sample points required to compute elementary effects

[4] These are purely the statistical errors and not the type of errors that lead to uncertainties introduced in section 1.1.2

[5] Let, Null hypothesis H_0 : Input is non influential. Alternative hypothesis H_a : Input is influential. Type II error occurs when one accepts a null hypothesis when it is actually not true.

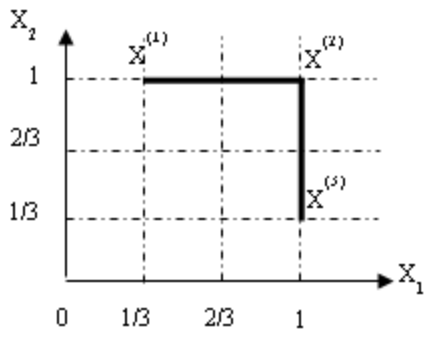


Figure 3.2: Illustration of trajectory in input space when number of inputs $k = 2$

negligible overall influence due to the cancelling effects of the opposite polarity elements. Alternatively, Morris suggested to study both the μ_i and σ_i measures simultaneously using a graphical representation in (μ, σ) plane. Although if the μ_i would measure negligible, but σ_i would certainly capture the variability in elementary effects. However, this approach has been found to be impractical in case of large models having multiple outputs.

Improved Morris measure

Campolongo et al. [74] defined an improved Morris measure μ_i^* as the mean estimate of distribution G_i of absolute values of an elementary effect $|EE_i|$. μ_i^* measure is suitable for the models with several uncertain inputs and multiple outputs. This improved model independent measure is robust against type II errors providing realistic ranking of the inputs. An efficient Morris design based generation of trajectories is susceptible to a non-uniform representation of an input space due to the small fixed number of random trajectories. This leads to biased estimates of the sensitive measures when inputs involved are highly uncertain. Campolongo et al. proposed selection of r trajectories from the set of random number of trajectories in a way that they are most widely and uniformly dispersed in an input space to represent all the regions of input space. A high number of different trajectories M are generated using an efficient Morris design discussed earlier. The optimal r trajectories are selected from M trajectories by maximising geometric distances between them. According to the experiments conducted in literature, a selection of $p = 4, \Delta = \frac{2}{3}$ and $r = 10$ has been considered as an optimum for highly dispersed represent-

ation of the input space. The measures are given as -

$$\mu_i = \frac{1}{r}\sum_{j=1}^{r} EE_i^j \tag{3.4}$$

$$\mu_i^* = \frac{1}{r}\sum_{j=1}^{r} |EE_i^j| \tag{3.5}$$

$$\sigma_i^2 = \frac{1}{r-1}\sum_{j=1}^{r} (EE_i^j - \mu_i)^2 \tag{3.6}$$

where $i = 1, 2, ..., k$.

The measures μ_i and μ_i^* represent average and absolute average of elementary effects, while σ_i represents their deviation from the mean effect. Using the settings as mentioned for Morris sampling design, the input sample values can be obtained from the characterized distributions. Campolongo et al. also compared the measure μ^* obtained using the trajectory selection based sampling design with the variance based total sensitivity index S_T and showed that EE based measure can be used as a good substitute for the total effect S_T under equivalent computational constraints.

3.3.2 Quantitative methods

Quantitative methods aid to know by what magnitude a given uncertain input is more important than the other amongst their rankings. These measures belong to a class of global sensitivity analysis techniques. Unlike qualitative measures, the effect of input x_i on the output y is quantified by exploring the entire space of uncertainty in inputs. Roy and Oberkampf [3] had described uncertainty analysis framework employed for the uncertainty estimation and decomposition due to different uncertain sources in scientific computing. The framework can be divided into two components - uncertainty propagation and decomposition. The different concepts involved and the related methods with regard to these components are briefly summarised below.

Monte Carlo (MC) approximation and sampling designs

Numerical methods for uncertainty propagation map the k-dimensional unit hypercube input space represented by N sample points to the scalar or func-

tional model output. The aim is to quantify the uncertainties in model outcomes propagated from uncertain inputs by estimating an unconditional variance of model outcomes evaluated at the sampled points in input space. Monte Carlo (MC) simulation-based method [75] for uncertainty propagation is illustrated below:

MC method involves an evaluation of a multi-dimensional integral of a square-integrable function $f(x_1, \ldots, x_k)$, representing simulation model, over a k- dimensional unit hypercube. An approximate MC estimate for the numerical integral -

$$\int \ldots \int_k f(x_1, \ldots, x_k)\, dx_1 \ldots dx_k \tag{3.7}$$

using quadrature rule[6] is given as the mathematical expectation -

$$\mathbb{E} = \frac{1}{N} \sum_{n=1}^{N} f(x_{1n}, \ldots, x_{kn}) \tag{3.8}$$

The repetitive model evaluations at the sample points forms a distribution of output quantity of interest. Considering the law of large numbers, the mean estimate converges to a true value of the integral with an estimate error $= \frac{\sigma}{\sqrt{N}}$. However, the convergence rate of $\frac{1}{\sqrt{N}}$ for an estimate error is still slow since it demands for a large number of samples and hence, the model computations. A crude form of MC uncertainty propagation method involve a simple pseudo-random sampling of input points with a replacement. Although the sampling method gives an unbiased estimate of a moment of distribution i.e. the statistical parameters (such as mean and variance), but prone to a large estimate error due to an uneven distribution of finite number of sample points drawn using random sampling. **Sampling strategies** and **designs-based MC** simulations are recommended to employ in order to improve the performance efficiency of moment estimation methods as the sampling strategies allow to achieve the convergence in estimation of moment parameter values at faster rates (i.e. with less number of sampling points) and also, reducing the moment estimate error [33] [76].

[6]Quadrature rule states that a multi-dimensional definite numerical integral can be approximated as the weighted sum of function values evaluated at the sampled points within the k-dimensional domain of input space.

Sampling strategies for the careful selection of design / training points are broadly classified into stratified sampling and quasi-random sequences. In contrast to pseudo-random sampling [34], these more sophisticated sampling techniques are designed to achieve uniformity in the design points' distribution in the input space, providing sufficient and well dispersed data / observations from which to construct a highly representative metamodel response surface. **Latin Hypercube Sampling (LHS)** is a popular stratified sampling strategy, where each input X_i is divided into say m sub-intervals, having equal marginal probability $1/m$ [33]. A sample is selected from each consecutive subinterval to form a sequence of m points. These sequences of m points from each input are randomly matched to form a k-dimensional set of m sample points [60]. In contrast, quasi-random sequences follow space filling criteria of discrepancy and geometric deviation to generate a uniformly dispersed sequence of sample points [77].

In the context of this work and for future extentions of its scope, we adopt the standard practice of using LHS, which has an estimated error convergence rate that is independent of the dimensionality of the input space [78] and is almost as fast as quasi-random sequences for large samples (N) [60] [79].

Uncertainty Propagation

MC uncertainty propagation approaches can be classified according to the type of uncertainty (epistemic or aleatory or both) being envisaged to be propagated through model evaluations [80]. Single and double loop schemes are frequently proposed in the related literature for specific situations of MC uncertainty propagation [81] [82]. Single loop scheme can be attributed to the task of propagating epistemic uncertainty (e.g. lack of knowledge about the precise value of input parameters) related to the data inputs (Figure 3.3). It is straightforward to propagate the epistemic uncertainty through model evaluations because they are conducted in the uncertainty space of controllable variables (e.g. data inputs) whose distributions are known or can be assigned a priori. While the double loop scheme involve propagation of both epistemic and aleatory (e.g. due to inherent model uncertainty) and the distribution of this inherent uncontrollable variable is unknown, both types of uncertainty are propagated in two loops: 1) an inner loop involving repetitive model execu-

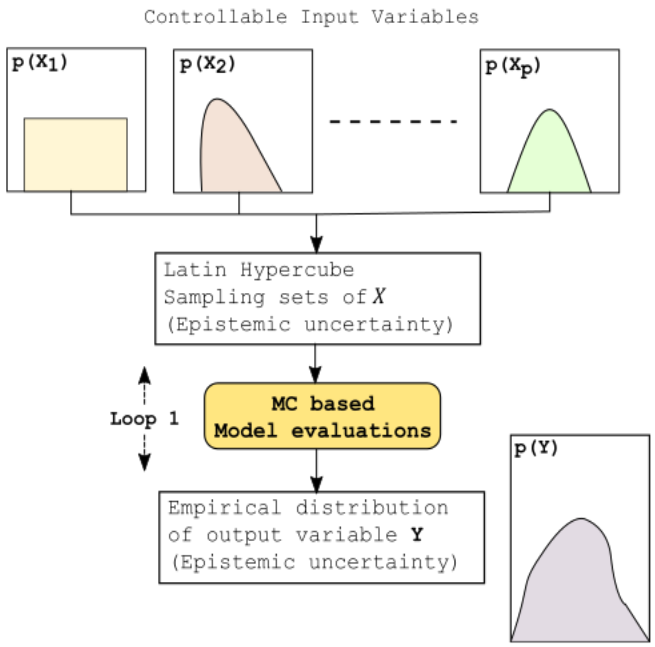


Figure 3.3: Single loop Uncertainty propagation scheme

tions for approximating the aleatory uncertainty and 2) an outer loop involving model computations at epistemic samples (Figure 3.4) [83].

Uncertainty Decomposition

Uncertainty propagation is followed by the computation of sensitivity index for the quantitative ranking of uncertain sources. Uncertainty decomposition or sensitivity analysis is defined as the process by which the total uncertainty in model output is apportioned to the respective uncertainty arising from the influential input sources. Sensitivity index defines the degree of responsiveness of model output to its uncertain sources. Quantitative methods for uncertainty quantification and decomposition can be classified into moment dependent (i.e. dependent on the statistical measure such as mean and variance) and independent[7] measures. **ANalysis Of VAriance (ANOVA)** -based moment

[7]Moment independent measures [84] [85] observe changes in entire distribution of random variables instead of utilizing statistics of the distribution as uncertainty measures. The introduction and review of these measures lie outside the scope of this thesis. We invite readers to refer to the following literature on moment independent measures: Entropy-based [86] [87] [88] and Distribution-based measures [89] [90] [91]. And for the latest review and

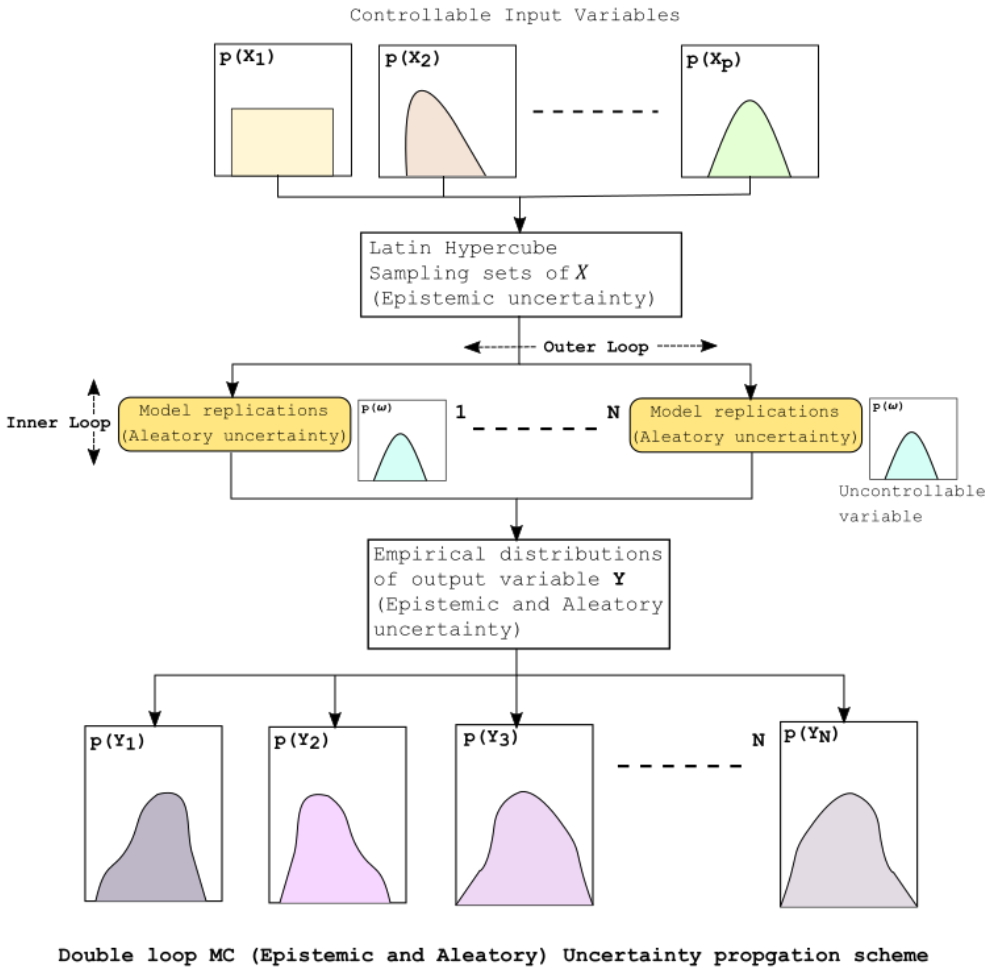


Figure 3.4: Double loop Uncertainty propagation scheme

dependent measures for the estimation of individual and total effect sensitivity indices are briefly described in the following subsections.

Fourier Amplitude Sensitivity Test (FAST) method

FAST method was proposed by Cukier et al. [93] to compute the main effect of each individual input on the total output variance. The main effect index defined is the ratio of conditional output variance due to the input under consideration to the total unconditional variance, given as the Main effect $= \frac{var_X[\mathbb{E}(Y/X)]}{var(Y)}$ where, Y and X output and input random variables respectively, $\mathbb{E}(Y/X)$ is the conditional expectation of Y for fixed X, var_X is taken over all the values of X and $var(Y)$ is the unconditional variance. Classical FAST method was limited to the computation of only the main effect terms and had drawback of computational complexity.

Sobol method

Extending FAST method in a generic way and combining Elforn and Steins work [94] on functional decomposition of the response $f(\mathbf{X})$ into a set of functions of random variables, **Sobol** [36] proposed MC-based evaluation of a multi-dimensional integral of a square-integrable [8] function $f(x_1, \ldots, x_k)$ and its decomposed functional terms over a k-dimensional unit hypercube. For the case of two input variables X_1 and X_2, according to the variance decomposition theorem

$$V = V_1 + V_2 + V_{12} \tag{3.9}$$

where unconditional variance V is computed from the distribution obtained from the evaluations of function f at sample points, $V_1 = V(\mathbb{E}[Y|X_1])$ and $V_2 = V(\mathbb{E}[Y|X_2])$. And the interaction variance from Eq. 3.9 is obtained as -

$$V_{12} = V - V_1 - V_2 = V(\mathbb{E}[Y|X_1, X_2]) - V(\mathbb{E}[Y|X_1]) - V(\mathbb{E}[Y|X_2]) \tag{3.10}$$

advances in the development of sensitivity analysis methods - Borgonovo and Plischke [92].

[8] A necessary condition for the functional decomposition to be unique.

Dividing both sides of Eq. 3.9 by the unconditional variance V gives main and interaction effect global sensitivity indices estimates i.e. $S_1 + S_2 + S_{12} = 1$ where, $S_1 = \frac{V(\mathbb{E}[Y|X_1])}{V}$, $S_2 = \frac{V(\mathbb{E}[Y|X_2])}{V}$ and $S_{12} = \frac{V_{12}}{V} = 1 - S_1 - S_2$. In general, the variance decomposition equation 3.9 can provide the unconditional variance as exhaustive sum of partial variance terms of increasing dimensionality; first, second and higher order interaction effects in sensitivity indices terms -

$$\sum_i S_i + \sum_i \sum_{j>i} S_{ij} + \sum_i \sum_{j>i} \sum_{l>j} S_{ijl} + \cdots + S_{123\ldots k} = 1 \tag{3.11}$$

Alternatively, Homma and Saltelli [95] proposed the total effect index defined as the summation of first order effect and all higher interactions involving given input, for instance X_i. As compared to original Sobol method, computing only two sets of indices (i.e. main and total effect) per input reduces the cost in terms of number of model evaluations in a numerical computation procedure. The total effect index for input X_i is then expressed as -

$$S_{T_i} = \frac{\mathbb{E}[V(Y|X_{-i})]}{V(Y)} = 1 - \frac{V[\mathbb{E}(Y|X_{-i})]}{V(Y)} \tag{3.12}$$

where, the notation X_{-i} denotes the conditioning on all other parameters except X_i as the purpose of total effect index is to account for all the first and higher order interactions variations explicitly due to variation in X_i only. For the three input model, the total effect due to X_1: $S_{T_1} = S_1 + S_{12} + S_{13} + S_{123}$.

Comparison between FAST and Sobol methods

Saltelli and Bolado [96] compared FAST and Sobol methods to explore similarities and relationships between them. The merits and demerits of the two methods described in the article are summarised in Table 3.1. The total effects computation method proposed by Homma and Saltelli [95] has computational constraints due to the calculation of computation intensive conditional expectation terms involved in the procedure. Saltelli [35] introduced a new numerical computation strategy for the estimation of sensitivity indices by making effective use of model evaluations. The results obtained using this numerical computation strategy were found in good agreement when experimented with the test functions (such as analytical Sobol g-function). The extended Sobol theorems demonstrated in this study claimed to perform more efficiently than

the FAST method. Table 3.2 compares different sensitivity analysis methods (and extensions of them) in terms of the computation costs i.e. a number of model evaluations required.

Saltelli [35] concluded that the application of quantitative methods is reasonable for the simulation models those having a single run time spanning from milliseconds to few minutes (in the context of this work, 20 and 90 seconds for proof-of-principle (5.2.4) and extended (6.2.4) case study S-BPS models respectively), while the qualitative treatment is preferred as an initial step for the models those taking hours or days for the single simulation run to complete and / or having a large number of input parameters which then requires segregation between the most influential and non-influential input parameters using qualitative screening methods. The computation expense (time and number of simulations) of estimating sensitivity indices values can be further reduced by : 1) computing only the main and total effects whilst still revealing the model structure and behaviour [95], 2) employing efficient sampling strategies in order to obtain better convergence in the values of indices estimate [97] [98] and 3) utilizing efficient numerical computation strategies / procedures (for instance, matrix based) for the sensitivity indices estimation [35] [99].

Table 3.1: Comparison between FAST and Sobol methods

Comparison indicator	**FAST** [93]	**Sobol** [36] [95]
Sensitivity index computation	Only Main effect	Main effect and interaction effects of higher order, total effects
Model behaviour detection	Non-linear, Non-monotonous and only additive	Relevant also for non-additive models
Computation costs	Cheaper and efficient, but increases with number of inputs	Intensive for large sample size but can be reduced to some extent when calculating total effects
Uncertainty in sensitivity index	Systematic bias observed from the analytical values	Converge relatively faster to the analytical values, more accurate estimates than FAST

Table 3.2: Comparison between sensitivity analysis methods. (Note: k and N are number of inputs and samples respectively.)

Effect	Classical FAST [93]	Sobol [36]	Sobol total effect [95]	Extended FAST [100]	**Extended Sobol [35]**	
					Theorem 1	**Theorem 2**
First order	N	N^2	N^2	combined	✓	✓
Total effect	-	N^2	N^2	N	✓	✓
Higher order ($[k-2]^{th}$ order, $k \geq 4$)	-	$\sum N^2$	-	-	✓	✓
Second order	-	N^2	-	-	-	✓
Total computation costs	$\mathbf{kN}$	$\mathbf{(2^k - 1)N^2 + N}$	$\mathbf{2kN^2 + N}$	$\mathbf{kN}$	$\mathbf{N(k + 2)}$	$\mathbf{N(2k + 2)}$

3.4 Uncertainty assessment approaches in BPS

Lomas and Eppel [101] considered the standard weather and operating schedules for a typical building design example and assigned a Gaussian distribution for the characterization of uncertainties in thermo-physical properties. The outcome of a Local Sensitivity Analysis (LSA)[9] measure called differential analysis revealed the linear nature of the thermal model, and the set-point temperature and window glazing U-value as the most influential inputs. De Wit [102] and his related foundational works on Uncertainty and Sensitivity Analysis of building thermal comfort performance [103] [104] applied the parameter screening method called Morris [73] and found out that the wind pressure coefficient and indoor air temperature were amongst the most influential factors. The uncertainties in all other parameters were assumed to be Gaussian distributed during their propagation and the resulted uncertainty in the thermal comfort was evaluated in terms of a coefficient of variation (= Standard deviation divided by sample mean).

The uncertainty analysis study by Macdonald [105] employed a dynamic building thermal model using ESP-r software. The study was largely focused on the identification of suitable methods for the uncertainty representation and its propagation while assessing individual and overall impacts of the parameters comprising of thermo-physical properties, casual heat gains and infiltration rates. A differential analysis was used to know the critical parameters affecting the resultant indoor temperature. The corresponding impacts of uncertainty in these parameters on the resultant indoor temperature were analysed in the context of a typical IT classroom building during early design stage in order to know the essential performance change requirements to avoid overheating risks during peak summer temperatures. The infiltration, occupant and computer gains were identified as the most influential parameters while the construction material parameters having little or no effect on the risk of overheating.

Mara and Tarantola [106] conducted ANalysis Of VAriance (ANOVA) based Global Sensitivity Analysis (GSA)[10] on the results from a test-cell build-

[9] LSA is a qualitative screening measure in which an input parameter is varied in isolation i.e. in a one-at-a-time fashion keeping other parameters at their base level or datum.

[10] GSA is a quantitative measure where all the concerned input parameters are varied simultaneously in order to capture the impact of parameter interactions on the output quantity of interest.

ing at University of Reunion Island. The objective was to better characterize and understand the thermal behaviour in the context of an existing building. The methodology employed a two-level full factorial sampling design: one level called base level assumes the actual thermal characteristics and the second level assumes perfectly adiabatic conditions with null heat transfer coefficients. The study enabled to investigate the impact of construction assemblies (wall, roof, ground and window) in order to know the influence of variation in their thermo-physical properties on the predicted indoor air temperature. A hourly computation of sensitivity indices revealed that the window and door during morning, the roof and west side wall in the afternoon and ground during night time were the most responsible for the variation in hourly indoor temperature. The results corroborated with the facts: During day time, the diffuse solar radiation transmitted through the window contributes to the warming up of indoor air and hence, having larger impact during the day and while its effect decreases during night time. In case of the ground assembly being influential during the night time, the thermal capacity of the concrete stores the absorbed irradiation during the day time and the same is released during the night.

Mechri et al. [30] analysed the variance in energy performance of a typical office building against the six design variables - thermal envelope area to volume ratio, transparent area to envelope area ratio, exterior surface absorptance, orientation, external solar shading factor and internal effective heat capacity of a building. The uncertainty analysis methodology was based on the Monte Carlo technique, which produces a set of outcomes by conducting model evaluations at the sample points in an input sample space, using a ISO standard 13790:2008 [107] based quasi-steady simplified monthly method for building energy performance assessment. The study concluded with the observations that - transparent to envelope ratio being the most significant and the surface absorptance being the least significant factor in determining the heating and cooling energy requirements of the office building.

An alternative and a more rigorous approach to the quantification of changes in building energy performance due to the changes in design parameters during early design stage was proposed by Hygh et al. [108]. A multivariate linear regression model was developed based on the dataset constructed using the sampled input parameter values and the resultant yearly energy consumption

estimates obtained from each EnergyPlus model instance within the Monte Carlo simulation framework. The impact of total 27 building design parameters, of a medium-sized rectangular office building, related to building form, orientation, fenestration, shading and thermo-physical properties was investigated under four different climatic conditions of US continent. The study claimed to fit an excellent linear regression model to the EnergyPlus simulation results and suggested it as an effective instrument for decision support during early design stages. According to the analysed sensitivity of energy loads to design parameters using standardised regression coefficients, the building area was observed to be the most influential design parameter with its orientation being the least and that too largely a similar trend was being observed across all the four climatic zones considered.

Heo [109] proposed a new Bayesian calibration framework based on normative building energy model to support risk-conscious retrofit decisions by accounting for the uncertainties associated with the energy conservation measures. The research work focused on the uncertainties related to physical (e.g. scenario, building physical and operational, model inadequacy, observation error) and financial risks (e.g. investment and utility costs) that lead to under expected energy savings predictions than the actual one. Heo used this step of uncertainty analysis to identify four dominant parameters (i.e. windows opening variable, indoor temperature, infiltration rate, air flow discharge coefficient) that have highest possible impact on the gas energy use for the Faculty of English building in the University of Cambridge, UK. This allowed to reduce the parameter space for the conduct of Bayesian calibration of normative energy model for this University building. The resulting probabilistic annual energy-savings accounting for the impact of sources of uncertainties reflected the quantified risks related to under-performing energy conservation measures (e.g. upgrading insulation, replacing windows and improving air-tightness) translated in terms of annual utility cost reduction.

A further more systematic approach by Spitz et al. [31] suggested step by step methodology to identify the most influential parameter and as well as to determine the uncertainty bands around simulation results. A simulation model of an experimental house was defined in EnergyPlus and the comparison was made between measured and simulated data. The first step

involved knowing of 10 most important parameters out of selected 139 parameters using local sensitivity analysis measure. In the next steps, the model runs were performed at the parameter sampled values of these 10 parameters to compute global sensitivity indices. The applied global sensitivity analysis measure revealed electric heating system capacity, internal gains, heat exchanger efficiency, infiltration, insulation thickness and conductivity having highest impact on the room air temperature. Sun [110] proposed methods to demonstrate parameter (e.g. ground albedo, convective heat transfer coefficients, lighting and plug load) and model form (e.g. solar diffuse irradiation model by Perez) Uncertainty quantification. The research work employed UQ repository to categorise types and sources of uncertainties at five different (operation, systems, building shell, urban and meteorology) levels of analysis for six case study buildings of the Georgia Institute of Technology campus. Sun mainly emphasised the importance of jointly considering the parameter and model form uncertainties in reducing prediction errors.

Wang [111, 112] further built upon the Sun's UQ repository to develop Georgia Tech Uncertainty and Risk Analysis Workbench (GURA-W) [32] and considered additional parameter (e.g. infiltration rate), scenario (e.g. occupancy variables) and model form (e.g. workmanship issues at a thermal bridge, HVAC system) uncertainty for the verification and reliability analysis of probabilistic energy predictions. The research work on occupancy variable uncertainty attempted to account for modeller's level of knowledge between 'total ignorance' and 'perfect knowledge' given the level of accessibility of measurement data of occupancy variables. The final outcome was the Modelling Framework to account for modeller's bias and uncertain view on typical building occupancy. The comparison of scenario profiles with mean occupancy profile revealed that the low bias and variance scenario succeeded in reproducing the actual occupancy profiles than the high bias and variance scenario. Wang's follow up research work [113] on generation of stochastic occupancy profiles, using Autoregressive Moving Average (ARMA) model, and observing their impact on aggregated HVAC energy consumption and peak demand concluded that stochastic occupant presence play less prominent role than lighting and appliance usage in the prediction of aggregated consumption due to weak occupant interaction with the building envelope and system components; and

conversely dynamic occupant stochastic interaction is important in predicting peak demand due to the inherent limitations of static profiles to represent the probabilistic nature.

Table 3.3 further summarizes some studies from the literature on the applications of uncertainty quantification approaches in BPS. The model output quantity of interest is either heating / cooling energy loads or indoor air (thermal comfort) temperature or both. The literature studies were focused largely on modelling the linear behaviour of the building simulation model by employing multi-linear regression techniques. They found to be unsuitable in the context of the requirement of approximating the complex and non-linear behaviour of the S-BPS model as set out for **Objective I** of this thesis. Secondly obvious, these studies were performed considering only the deterministic heat transfer phenomena involved in the buildings, irrespective of model / simulation program (simplified or dynamic) used. This forms the basis for thesis **Objective II** that there is a need to emulate the superposition of both deterministic and stochastic phenomena occurring inside the buildings in a plausible and computationally efficient manner while conducting Uncertainty Quantification studies based on the metamodelling techniques. With a lack of rigorousness in approximating the complex phenomena, these studies in literature are incomplete in propagation and decomposition of both the type of uncertainties (epistemic and aleatory) in energy predictions (**Objective III**). Lastly, the scope of these were limited to performing only Local Sensitivity Analysis (LSA) i.e. only one parameter at a time was varied to study its impact which is constrained in scope when the uncertain parameters exhibit combined interaction effect on the energy predictions. This points to the need (**Objective IV**) of extending the scope of Uncertainty and Sensitivity Analysis studies by considering the complex cases and as well as performing global nature of sensitivity analysis studies i.e. quantifying and decomposing the main and total effects of uncertain parameters on the energy predictions.

From the last column of most influential parameters identified in Table 3.3 and as well as from reviewed literature, it can be concluded that the global impacts (main and total effects) of uncertainty in the detailed thermophysical properties (e.g. conductivity, specific heat capacity, density, etc.) of opaque materials and that of the optical properties (e.g. solar and visible transmit-

tance, conductivity / u-factor, solar heat gain coefficient, etc.) of transparent ones and as well as that of the most uncertain parameter i.e. infiltration rate have not been thoroughly studied. Along with that the impact of aleatory uncertainties arising from the models of stochastic occupants' behaviours (i.e. stochastic internal heat gains) on the energy predictions has also not been numerically quantified and decomposed before. Therefore, this thesis prepares the foundational step of conducting Uncertainty and Sensitivity Analysis for the complex case of building performance simulation such as the stochastic BPS.

Table 3.3: Summary of applications uncertainty and sensitivity analysis approaches in BPS. (Note: ISO 13790:2008 Quasi-steady simplified monthly method, EnergyPlus and ESP-r Dynamic heat conduction and balance simulation program, Simplified MAV - Simplified Monthly Average Values)

Research article	**Building phase and type**	**Model / program used**	**Method used**	**Model behaviour**	**Most influential parameters** (in decreasing order of importance)
Heiselberg et al. [114]	Design phase and Office	Simplified MAV	Morris	-	lighting and ventilation, wall U-value
Capozzoli et al. [115]	Design and Office	ISO 13790:2008	Multi-linear regression	-	window-to-wall and surface-to-volume ratio, specific heat capacity of wall
Corrado and Mechri [116]	Operation and Residential	ISO 13790:2008	Morris	-	air change rate, number of occupants, metabolic and equipment heat gain
Booth et al. [117]	Operation and Residential	ISO 13790:2008	Morris	-	Fraction of space heated, heating set-point temperature, coefficient of performance of heating system, air leakage, glazing and wall U value, window-to-wall ratio
Sanchez et al. [118]	Operation and Residential	ESP-r	Morris and Extended	Non-linear and Non-additive	building size (width, length and height), set point temperature, ventilation rate and insulation thickness
Yildiz et al. [119]	Design and Residential	EnergyPlus	Regression	-	window area, natural ventilation, building length, window solar heat gain, air infiltration, building width
Rodriguez et al. [120]	Operation and Residential	EnergyPlus	Regression	$R^2 > 0.984$ (linear)	occupancy level and weather
Silva and Ghisi [121]	Operation and Residential	EnergyPlus	Regression	$R^2 > 0.95$ (linear)	specific heat capacity and thickness of brick wall, number of occupants, occupancy schedule
Ioannou and Itard [122]	Operation and Residential	EnergyPlus	Regression	$R^2 > 0.7$ (linear)	set-point temperature, ventilation, window U and G value, facade conductivity

3.5 Summary

For the example case of environmental models (such as HYMOD watershed model [123]), Yang [124] pointed out, in general, an absolute need (due to highly non-linear, non-monotonous and next to unknown nature of such complex and distributed models) to observe the convergence and estimate the uncertainty of the sensitivity indices by selecting and utilizing an appropriate technique amongst different sensitivity analysis techniques. Yang concluded that the Sobol method [36] is the most robust technique amongst all others (the Morris method [73], Linear Regression (LR), Regionalized Sensitivity Analysis (RSA) and non-parametric smoothing) in quantifying the uncertainties and ranking of influential parameters although at a high computational expense due to the requirement for large number of model executions.

Figure 3.5 illustrate the classification system of Uncertainty and Sensitivity Analysis techniques[11] based on the computational complexity. The techniques are classified into three categories : 1) Screening techniques are aimed at segregating the influential and non-influential parameters, 2) Ranking techniques are aimed at qualitatively ordering of the parameters according to their influence and 3) Indices ones are aimed at quantifying the relative influence of the parameters in terms of numerical importance measures.

[11] We note that the techniques such as Correlation and Regression analysis were not discussed just because of their more trivial way of obtaining output sensitivity information from the simple statistical analysis of the input/output dataset. We invite reader to refer to Kleijen and Helton [125], Helton and Davis [126], Storlie et al. [127] for the review of these approaches. Note for Figure 3.5: CC - pearson Correlation Coefficient, PCC - Partial Correlation Coefficient, SRCC - Spearman Rank Correlation Coefficient, PRCC - Partial Rank Correlation Coefficient, SRC - Standardised Regression Coefficients

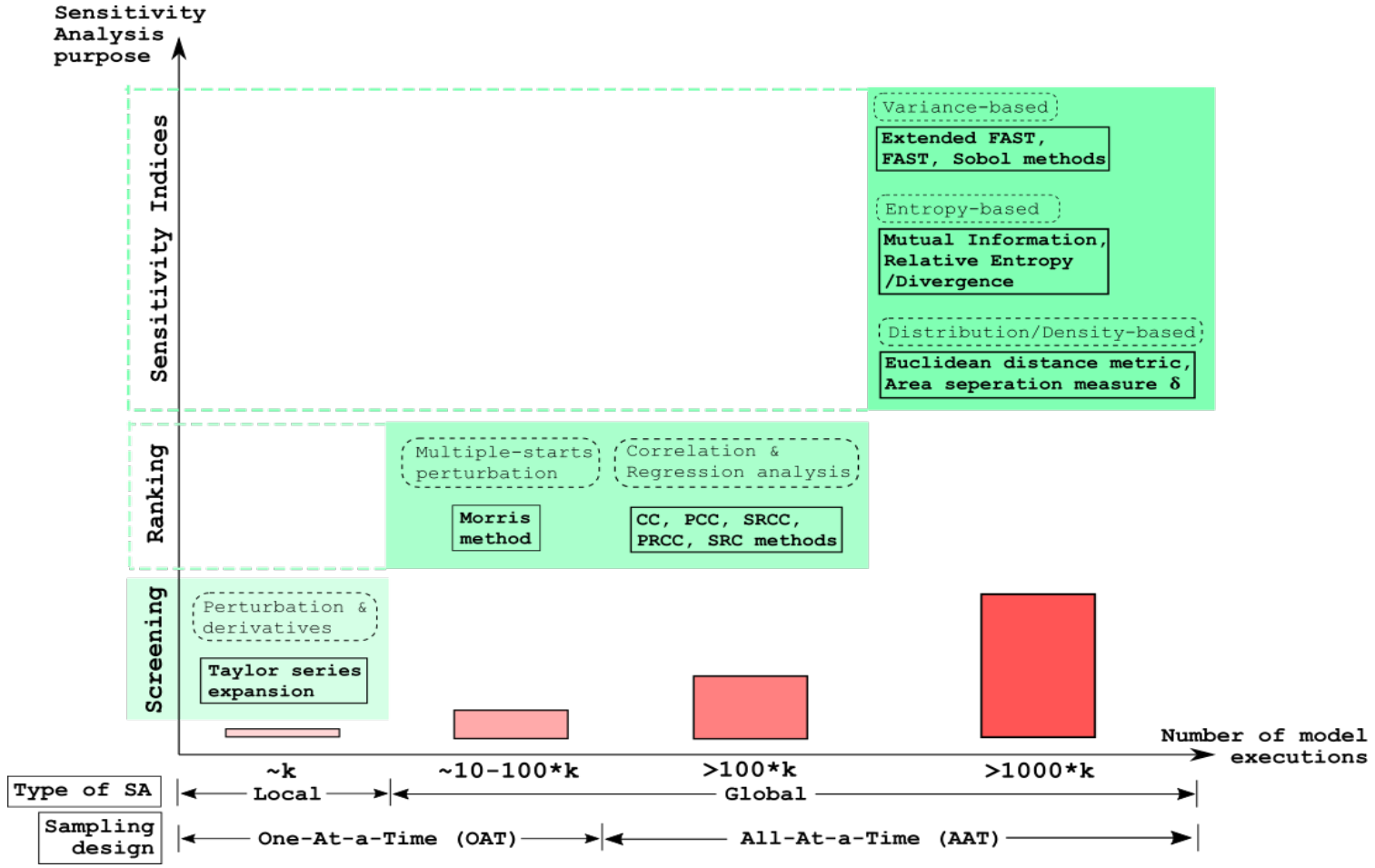


Figure 3.5: Classification system of Uncertainty and Sensitivity Analysis (USA) methods against the computational expense (Adapted from Pianosi et al. [4], Tian [5], Nguyen and Reiter [6] and Tian et al. [7] for the methods explored in earlier sections 3.3.1, 3.3.2 and applied in Chapters 5, 6) (Note: k - number of inputs)

To address the computational challenge of propagating uncertainties in building performance simulation (subsection 1.1.3), numerous statistical metamodeling techniques have been proposed in recent years for sensitivity analysis of building [122] and climatic [128] parameters, and to support more accurate prediction of building energy performance [129] [130]. The aim of metamodeling is to use outputs from simulations at carefully designed inputs to train a statistical model that provides an accurate approximation of the simulator response at a lower computation cost than that of the simulator. Once a metamodel has been constructed, the BPS tool can be substituted by the emulator and employed to conduct Monte Carlo uncertainty propagation and variance-based sensitivity analysis in providing early stage building design guidance [108] [131] and to estimate the performance of passively designed buildings [132]. It is to be noted that metamodeling is also highly relevant for computationally intensive tasks relevant to BPS such as Bayesian calibration [133, 134, 135] and optimisation [136, 137]. However, existing approaches are restricted to the emulation and uncertainty quantification of deterministic phenomena, as is the standard case in building simulation, neglecting sources of aleatory uncertainty, for example that arise from stochastic phenomena such as occupants' presence and their interactions with the building envelope, and associated heat and mass transfers. Failing to incorporate these stochastic phenomena, this thesis argues, undermines the utility of uncertainty quantification exercises, and efforts to reduce the performance gap.

The function f (Eq. 1.1) defines here a system of coupled differential algebraic equations concerning interactions between thermodynamic phenomena (such as heat conduction, convection and radiation) occurring inside the buildings [44]. A numerical solver or computer code such as the thermal building simulation program EnergyPlus employs numerical approximations to the functional form f to solve this system of equations [45]. As Wetter and Polak [138] point out, the BPS programs / solvers represent non-linear, multimodal and discontinuous function of the design parameters. As mentioned earlier, in order to reduce the computation cost of UQ studies, a metamodeling approach is highly advantageous. The chosen approach should then reliably approximate the underlying functional form as a precondition. Moreover, the approach should also consider stochastic perturbations due to occupants' behaviours

while approximating the functional response. Therefore, there is a clear need for a new metamodeling approach that: 1) estimates the functional form from training datasets whilst specifying the uncertainty in estimating it, and 2) accounts for stochastic perturbations, such as those arising from occupants' behaviours, impacting on the function response.

Chapter 4

Development of Mathematical and Computational Framework : EmUSA

In the next section 4.1, we discuss alternative candidate approaches that have been or could be applied to the emulation of building performance simulators, with a view to identifying the most promising candidate and how this can be extended for our purposes. In section 4.3, we present in detail our implementation of the emulation strategy, going on to present our related metamodeling based uncertainty quantification strategy in section 4.4.

4.1 Metamodeling Techniques

Numerous metamodeling techniques have been employed to overcome the computational costs of running thousands of expensive BP simulations for building design space exploration [139, 131], optimization, [136, 137], model calibration [135, 140] and uncertainty quantification (UQ) [108, 132]. Techniques that have been used within (deterministic) BPS workflows include ordinary least squares regression (OLS) [141], multivariate adaptive regression splines (MARS) [142], random forest (RF) [143], Gaussian process regression (GPR) [144], support vector regression (SVR) [145] and artificial neural networks (ANN) [141]. For a thorough practical evaluation of these techniques, we refer the interested reader to : the works of Cheng and Cao [129] for the application

of MARS, Yidil *et. al.* [146] and Amasyali and El-Gohary [147] for reviews of regression models, Østergård *et. al.* [148] and Wei *et. al.* [149] for comparative studies, and Lim and Zhai [150] for an analysis of the impact of metamodels on bayesian calibration.

While the literature above reveals that there exists no single emulator-based approach that is suitable for all purposes, we review these metamodeling techniques based on the following criteria defined by Jin et al. [151] in our search for most suitable candidate metamodel that could satisfy our building simulation requirement of performing UQ under occupants' behavioural variability.

- **Accuracy** - the ability of an emulator in making predictions over an input space, usually reported by unit less measure of error called coefficient of determination or predictivity (R^2) (defined in subsection 4.2.1 Eq. 4.9).

- **Robustness** - the measure of an emulator ability to predict accurately for different nature of problems, indicated by mean and standard deviation of R^2 when analysed for different types of model function and output quantity of interest related to application cases.

- **Efficiency** - the computational effort required for training an emulator and making new predictions, assessed in terms of amount of computation time required against the size of training set and number of inputs.

- **Transparency** - the nature of an emulator structure in revealing the simulator behaviour against the input variables and interactions between them, interpreted in terms of additivity, linearity and monotonicity.

- **Simplicity** - easy to use and the parameter settings can be readily modified for the problem with minimum user input.

In remainder of this section, we proceed with a briefing of above mentioned techniques, their applications in BPS and performance assessment against these criteria.

Ordinary Least Squares regression (OLS)

OLS regression is expressed as the dependent variable Y being the linear combination of the independent variables $(X_1, X_2, ..., X_p)$ weighted by the regression coefficients plus an error term ϵ_i. OLS regression are quite commonly applied techniques - to quantify building energy performance [108] and to investigate the importance of building design variables in early design stage by constructing a multi-linear regression over a range of variable values [115], to analyse uncertainties in building simulation [121] and perform sensitivity analysis of macroparameters [120] [122] using standardised regression coefficients. Although the multi-linear regression techniques are fast, easy to use and interpret, the linear modelling assumption to fit a curve or surface to the training data appears restrictive when the underlying nature of simulator f is expected to be nonlinear. This affects the accuracy of the metamodel in making predictions at new test points and as well as its robustness to emulate different nature of functions. Rastogi et al. [152] compared the accuracy of linear and nonlinear regression models in approximating BPS model and concluded that nonlinear regression modes are $3-4$ times more accurate than the linear counterparts.

Multivariate Adaptive Regression Splines (MARS)

MARS adaptively builds a model of weighted nonlinear functions, called basis functions by a forward/backward iterative approach. In forward phase, the training data is recursively split by assigning a pair of basis functions that maximizes the reduction of the residual sum of squares. In backward phase, the least effective terms of the fitted model are successively removed using cross validation to avoid overfitting. Cheng and Cao [129] proposed a technique based on MARS to predict heating and cooling loads of residential buildings. The study also compared proposed MARS approach with other techniques including neural networks and support vector regression using cross validation approach and showed to obtain maximum reduction in Root Mean Square errors (RMSEs) compared to these other methods. MARS performed well in terms of prediction accuracy of BPS, depicting robustness and transparency in its ability to represent model nonlinearity and additivity using basis functions but may become computationally inefficient with large input size due to

laborious recursive cross validation in backward phase to remove least influential inputs. Moreover, given the complexity of S-BPS model considered in this thesis, the recursive validation procedure may again hinder the practicability of MARS.

Random Forest (RF)

RF is an extension of decision trees - classification and regression trees, grown using many such trees which are combined and averaged to make predictions. A tree is constructed by binary and recursive splitting of training set into branches at the threshold values of response variables that leads to largest reduction in residual sum of squares. The process continues until the number of specified splits is reached or all leafs contain a single observation. Tsanas and Xifara [130] proposed an application framework based on RF to study the impact of building design variables on heating and cooling demands. RF has been shown to outperform parametric least square regression techniques in terms of accuracy and robustness in representing a nonlinear structure. But, for large number of inputs, an internal structure of the model may become difficult to interpret due to aggregation of multiple trees [141] [148].

Support Vector Regression (SVR)

SVR is a nonlinear generalization of the linear regression model in a high dimensional feature space, represented by the kernel function and its parameters. In addition, SVR also estimates the regression function with an uncertainty band specified by an additive term ϵ. The parameters and uncertainty are estimated by the optimization algorithm, based on the structural risk minimization approach, by minimizing the sum of residuals. Amasyali and El-Gohary [147] review the applications of SVR using radial basis function for the prediction of building energy performance. SVR technique offers higher levels of prediction accuracies and are computationally efficient upto a few hundred training points but are less simple requiring additional settings for the diverse nature of problems and as well as less robust when the number of input variables are large [148]. The number of input variables can be of the order of magnitude of two for some practical uncertainty and sensitivity analysis studies of building performance [31] and as well as for the potential future practical

applications of this work.

Artificial Neural Networks (ANN)

A multi-layered **ANN** typically consists of an input layer, hidden layers and output neuron layer. Each layer has number of neurons that connects and propagates information to next layers using weights and transfer functions. After specifying network parameters, the network is trained using learning data and the output is compared with desirable response values using mean square errors, followed by optimization algorithm to update neuron weights so that the mean square error is minimum. A numerous BPS applications of ANN have been presented, for a comprehensive review [153] [154] - Kalogirou works [155] on the prediction of energy consumption and indoor temperatures and their long term forecasting [156], analysis of HVAC system behaviour and their optimization [157]. ANN has been relatively faster than the other methods particularly during training and making new predictions with large datasets. ANN is a black-box approach where the relationship between the inputs is derived from the data which is flexible but also can be least interpretable due to the structure of the function that is being derived from the data remains hidden between the layers [148]. Moreover, they also require homogeneous training datasets - complete and exhaustive in terms of their representation [158].

Gaussian Process Regression (GPR)

GPR is a non-parametric approach to fitting a curve or a surface to the training data where the data generating process is viewed as an unknown function. The function values at design points is a collection of random variables having multivariate Gaussian distribution. The assumed prior probability over functions is updated upon the evidence of training data under Bayesian framework resulting into a posterior distribution which is used to make predictions at new design points. The relationship between the input space design points is expressed by the covariance function which along with the mean function define Gaussian process. The hyperparameters can be estimated in a non-parametric way by maximizing marginal likelihood conditioned on the training data. As GPR is a probabilistic approach, it presents emulator uncertainty at new pre-

dictions in the form of distribution. A number of studies demonstrated the application of GPR in BPS applications − (1) studies to optimise the design of glazing systems for office buildings [159]; real time optimization of cooling energy systems for efficient energy management in high-rise office buildings [160]; building design optimization [161] [162], (2) calibration and diagnosis of building energy simulation models for the formulation of energy conservation and model predictive control strategies [163], and (3) sensitivity analysis of energy performance predictions to weather and building envelope characteristics [128].

4.1.1 Selection of a candidate metamodelling technique

Numerous evaluations (see for example [150] and [148]), studying the effects of training set size, the nature of the simulator functions to be emulated and the number of inputs points to be represented, have shown that GPR is consistently able to emulate the underlying simulator response more accurately than competing techniques. Østergård et al. [148] compare the six mentioned metamodeling techniques applied to building performance simulations, judging these against performance indicators such as accuracy, efficiency, robustness, simplicity and transparency. They conclude that GPR metamodels are the most accurate (with a coefficient of determination or predictivity $R^2 > 0.99$) in approximating the nonlinear nature of BPS problems.

GPR shown to be also robust and easy to use towards different nature of problems having varying complexity and dimensionality with readily modifiable hyperparameter settings requiring less user input for their estimation during optimization. When compared with ANN technique, GPR has been proven to be computationally less efficient with the large training datasets, typically for thousands of observations and GPR required more time for training and in making new predictions depending upon the chosen hyperparameter optimization procedure [150].

Although the complex covariance functions are required for sufficient flexibility while GP fitting which may become at times difficult to comprehend, they would definitely be more interpretable than the completely black-box approach like ANN where the emulator structure is almost invisible i.e. hidden between the layers. However, in case of GPR, after learning from the data,

Table 4.1: Comparison between metamodeling techniques against the performance criteria (Note: fully - 5 points, partially - 3 points, not-at-all - zero. The numbers in the brackets are the points scored by the technique based on the subjective judgement of fully / partially / not-at-all.)

Technique	**Performance metrics**				
	Accuracy	Robustness	Efficiency	Transparency	Simplicity
OLS(15)	not-at-all	not-at-all	fully	fully	fully
MARS(18)	fully	fully	not-at-all	fully	partially
RF(15)	partially	partially	partially	partially	partially
SVR(17)	fully	partially	partially	partially	partially
ANN(18)	fully	fully	fully	not-at-all	partially
GPR(21)	fully	fully	partially	partially	fully

the estimated hyperparameter parameters such as length scale, actually reveal the simulator behaviour against input variables in terms of their influence on the response i.e. how dynamic is the output response toward variation in input variables. GPR models were also termed to be intuitive inter- and extrapolators much more than just surface smoothers by O'Hagan [164].

Table 4.1 summarizes the extent to which each technique meets the criteria using the subjective judgements of : fully / partially / not-at-all in the context of their applications for building performance simulations. GPR metamodelling technique scores highest (21/25 points) in terms of the presented performance metrics despite being slightly less computationally efficient than the ANN but that only with large datasets and also depending upon the comparable model optimization settings for weights / hyperparameters estimation. GPR still promises a competitive and efficient performance during the training and predictions with medium-sized datasets which is found to be advantageous and sufficient to address the problem in hand for this thesis. In addition, in the current context of emulating a stochastic response from BPS, the features of GPR allows enough flexibility to modify the core configuration in order to account for the non-constant stochastic occupants' variability as a noise variance by still conforming to its mathematical framework. This also makes feasible to perform UQ for such increasingly complex and highly computation intensive case of stochastic building performance simulator.

Addressing of practical issues with the Gaussian Process emulators

However, there are few practical situations as pointed by Bastos and O'Hagan [165] and Al-Taweel [166] that are required to be taken into account while constructing a valid Gaussian process emulator of a simulator: 1) A covariance function inherently assumes a stationarity where the simulator output may respond with similar degrees of smoothness at all points in the input space, 2) The hyperparameters are inferred from the data and it may happen that the inferring procedure estimate them poorly and 3) The Gaussian process emulators may result in poor predictions of the simulator outputs if inappropriate forms of the mean and covariance function are assumed while training them. These have following consequences which aid in identifying such situations and potentially avoiding them when detected during the emulator construction procedure: 1) In case of non-stationarity in simulator response (behaving much more quickly in some parts of the input space than the others), the confidence intervals of the emulator predictions can be too wide or too narrow in the regions when the corresponding response is very less or highly dynamic. 2) In case of poor estimation of hyperparameters, the confidence intervals of the emulator predictions can be systematically too wide or too narrow in the regions of training data points.

In addition, Bastos and O'Hagan [165] and Al-Taweel [166] proposed numerical and graphical diagnostic methods to validate the assumptions that are used in constructing the Gaussian process emulator of a simulator. The methods statistically compare simulator and emulator predictions at the validation data points. For instance, the numerical methods include diagnostic based on individual prediction errors (between the observed simulator outputs and the predictive mean output at the same inputs), a summary diagnostic called mahalanobis distance (comparing emulator predictions with the simulator outputs under the same inputs) and decomposition indices (variance, eigen, cholesky and pivoted cholesky). Graphical methods are actually the plots of individual errors against the emulator's predictions, the validation data index, the inputs and quantile-quantile plots for the emulator and simulator outputs. Other diagnostics include the credible interval diagnostic (proportion of validation outputs that lie in confidence intervals) and predictive density diagnostic (a

decreasing function of the Mahalanobis distance). The diagnostics essentially check whether the emulator and its uncertainty can represent the simulator. And when that is not the case, then the diagnostics can aid in signposting about where the problematic situation might be.

In the context of this thesis, the credible interval diagnostic (Eq. 4.10) has been used to evaluate the uncertainty in the predictions of emulators trained for the proof-of-principle (Table 5.3) and extended studies (subsection 6.2.2). The simulator responses have been checked for the non-stationarity condition during emulator fitting procedure by experimenting with the covariance function structure and then the highly dynamic and non-representative response surfaces were discarded by also verifying their prediction accuracies (subsection 5.2.2). The stationarity or smoothness in emulator fitting has been observed by plotting the response surfaces of the fitted mean and variance GPs (Figure 5.3). The proportion of stochastic simulator predictions contained within the confidence intervals conform that the intervals are not too wide or too narrow in the regions of data points (Figure 5.4). This is to observe for the quality of hyperparameters estimated from the data. These are few practical issues with the Gaussian process emulators that need due consideration when deciding on their usefulness for building performance simulation applications.

4.1.2 The proposal

In this chapter, we propose a new metamodeling approach to simulate the stochastic response of a BPS that simulates occupants' behaviour. In this, we follow the generic stochastic Kriging metamodeling approach of Ankenman et al. [167] which uses repetitions of the stochastic simulator, at a given design point $\mathbf{X}$, to characterise/approximate the input-depdendent mean response and variance of the stochastic building performance simulator, via Gaussian process regression (GPR). The proposed approach aims to provide GPR emulators to the following quantities

$$f_d(\mathbf{X}) = \mathbb{E}_\omega(Y|\mathbf{X}), \qquad V(\mathbf{X}) = \mathbb{E}_\omega(Y|\mathbf{X} - f_d(\mathbf{X}))^2, \tag{4.1}$$

where $\mathbb{E}_\omega$ denotes the expectation with respect to ω. By means of off-the-shelff GPR algorithms [168], we fit a GP to the variance $V(\mathbf{X})$ given the sample

variance from repeated simulator runs computed at design / training points. We subsequently use predictions from the corresponding metamodel of $V(\mathbf{X})$ to fit a non-heteroscedastic GP to f_d.

To the best of our knowledge, metamodeling approaches for BPS have only been applied where the simulator response is deterministic (for a particular choice of inputs). In this thesis we address the case when BPS is coupled with a stochastic model for occupants' behaviour. As described in the introductory chapter 1 (section 1.2), standard metamodeling techniques (including GPR) cannot be directly applied to emulate the response of a stochastic BPS. The main challenge arises from the fact that the variance of the simulator (with respect to an uncontrollable variable), $V(\mathbf{X})$ is a function of input parameter $\mathbf{X}$ (see equation(4.1)) which is, in general, unknown.

From the statistical literature we identify two possible routes to address this challenge. With the approach proposed in [169, 170] the unknown values of the variance $V(\mathbf{X})$, at the training points, can be treated as hidden variables within the GPR framework with non-heterosedastic variance. Alternatively, when repetitions of the computer code are available for each input, the Stochastic Kriging (SK) method can be applied to emulate the variance (given the sample variance from repetitions) [171, 167, 172]. Once $V(\mathbf{X})$ has been approximated via an emulator, GPR can be applied to emulate the mean response surface using information from the GP emulator for the variance. There is good evidence to suggest that SK lends itself more naturally to variance based sensitivity analysis which, in turn, requires an explicit emulation of both the mean and variance [41]. A SK approach is proposed in Section 4.3 for the emulation for the stochastic response of BPS, for which we demonstrate its efficiency for the computation of sensitivity indices in Section 5.2.4. Since standard GPR is at the core of the SK method, we introduce in the following section the basic elements of standard GPR (for deterministic computer simuators) and refer the interested reader to [144] for further details.

4.2 Standard Gaussian Process (GP) Metamodeling for deterministic simulators

Let us assume that we have a deterministic computer simulator that depends on a vector of input parameters denoted by $\mathbf{X} = (X_1, \ldots, X_p)$, that we assume belong to a space of admissible inputs $\mathcal{X} \subset \mathbb{R}^p$. We further assume that the outputs of the simulator can be described via the evaluation of a deterministic function $F : \mathcal{X} \to \mathbb{R}$, which maps inputs $\mathbf{X} \in \mathcal{X}$ into the corresponding simulator response (output) $F(\mathbf{X})$. For simplicity, in what follows we consider the case of a univariate simulator response, but the results will be extended to the multivariate case.

As discussed earlier, when running the simulator (i.e. evaluating $F(\mathbf{X})$ for a given input $\mathbf{X}$) is a computationally costly exercise, propagating inputs uncertainty through the simulator can become unfeasible. The aim of metamodeling is to replace F with a function that emulates the simulator response at much lower computational cost. GPR provides for this, by inferring the underlying functional form of F, given observations of the evaluation of F at a carefully selected number of *design* or *training points* $\{\mathbf{X}_n\}_{n=1}^N$ in the input space $\mathcal{X}$. More specifically, we consider the training set defined by $\mathcal{D} = \{(\mathbf{X}_n, Y_n)\}_{n=1}^N$ where Y_n are obsevations of $F(\mathbf{X}_n)$. Following standard assumptions we consider an observational model of the form [144]

$$Y_n = F(\mathbf{X}_n) + \epsilon, \tag{4.2}$$

where $\epsilon \in \mathbb{R}^d$ is Gaussian noise with zero mean and covariance $\sigma^2 I$.

Before observing the data $\mathcal{D}$, the underlying Bayesian framework for GPR metamodeling requires the specification of a prior distribution (over a space of functions) for the function F that we wish to infer. In the standard framework, the prior on F is a Gaussian distribution of functions with mean (function) $m(\mathbf{X})$ and covariance $k(\mathbf{X}, \mathbf{X}')$ (we denote this by $F \sim \mathcal{GP}(m(\mathbf{X}), k(\mathbf{X}, \mathbf{X}'))$). For the sake of clarity, we consider here the case $m(\mathbf{X}) = 0$ for all $\mathbf{X} \in \mathcal{X}$ which corresponds to the case in which we have no prior knowledge of the simulator.

The covariance function $k(\mathbf{X}, \mathbf{X}')$, on other hand, characterizes the vari-

ability of the family of functions defined by the GP prior. A covariance that enables a wide class of functions to be characterised is the Mattern covariance, defined as [144] :

$$k(\mathbf{X}, \mathbf{X}') = \sigma_Y^2 \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{||\mathbf{X} - \mathbf{X}'||}{l} \right)^{\nu} K_\nu \left(\frac{||\mathbf{X} - \mathbf{X}'||}{l} \right), \tag{4.3}$$

where Γ is the gamma function, l is the characteristic length scale, σ_Y^2 is the signal variance, K_ν is the modified Bessel function of the second kind of order ν, and $||\cdot||$ is the Euclidean norm (in $\mathbb{R}^p$). The parameter ν controls the regularity/smoothness of the samples. For the present work we consider ν to be fixed and so the Gaussian is fully characterised via the specification of the prior hyperparameters l, σ_Y that are contained in the single vector $\theta = (l, \sigma_Y^2)$.

In order to make predictions of F at a test location $\mathbf{X}^*$ (i.e. $F(\mathbf{X}^*)$), we consider the predictive distribution, $p(F(\mathbf{X}^*)|\mathbf{X}^*, \mathcal{D}, \theta)$, which from the standard GPR framework, is Gaussian with mean $\hat{F}(\mathbf{X}^*)$ and variance $\hat{K}(\mathbf{X}^*)$ defined by:

$$\hat{F}(\mathbf{X}^*) = K_\theta^T(\mathbf{X}^*)(K_\theta(\mathbf{X}) + \sigma^2 I)^{-1}\mathbf{Y} \tag{4.4}$$

$$\hat{K}(\mathbf{X}^*) = k_\theta(\mathbf{X}^*, \mathbf{X}^*) - K_\theta^T(\mathbf{X}^*)(K_\theta(\mathbf{X}) + \sigma^2 I)^{-1}K_\theta(\mathbf{X}^*), \tag{4.5}$$

where K_θ satisfies: $[K_\theta]_{i,j} = k_\theta(\mathbf{X}_i, \mathbf{X}_j)$, $K_\theta(\mathbf{X}^*) = (k_\theta(\mathbf{X}^*, \mathbf{X}_1), \ldots, k_\theta(\mathbf{X}^*, \mathbf{X}_N))^T$, and $\mathbf{Y} = (Y_1, \ldots, Y_N)^T$. We note that if θ and σ are known, expression (4.4) can be used to make predictions of the simulator F, at testing point $\mathbf{X}^*$. A measure of the uncertainty in those predictions is provided by the variance (4.5). In practice, however, the (hyper)parameters θ and σ are unknown and hence these must be inferred from the training set $\mathcal{D}$. The standard practice for estimating these parameters is to maximise the posterior density $p(\theta, \sigma^2|\mathcal{D})$, which from Bayes' rule, satisfies the condition:

$$p(\theta, \sigma^2|\mathcal{D}) \propto p(\mathcal{D}|\theta, \sigma^2)p(\theta, \sigma^2) \tag{4.6}$$

where $p(\theta, \sigma^2)$ is the (hyper-)prior distribution on (θ, σ^2) and $p(\mathcal{D}|\theta, \sigma^2)$ is the marginal likelihood given by the following expression [144]:

$$p(\mathcal{D}|\theta, \sigma^2) = \frac{1}{2}\mathbf{Y}^T(K_\theta + \sigma^2\mathbf{I})^{-1}\mathbf{Y} + \frac{1}{2}\log|K_\theta + \sigma^2\mathbf{I}| + \frac{N}{2}\log 2\pi \tag{4.7}$$

The maximum a posteriori (MAP) estimate of the prior hyperparameters is:

$$(\hat{\theta}, \hat{\sigma}^2) = \arg\max_{\theta,\sigma^2} p(\theta, \sigma^2|\mathcal{D}) = \arg\min_{\theta,\sigma^2} \Big[-\log p(\mathcal{D}|\theta, \sigma^2) - \log p(\theta, \sigma^2) \Big] \quad (4.8)$$

Finally, predictions of $F(\mathbf{X}^*)$ together with a measure of their uncertainty can be used by employing $(\hat{\theta}, \hat{\sigma}^2)$ in expressions (4.4)-(4.5).

As stated earlier, expressions (4.4) -(4.5) are derived under the assumption of zero prior mean $m(\mathbf{X}) = 0$. It is worth mentioning, however, that it is often beneficial to incorporate a non-zero mean function. A common choice is to consider $m(\mathbf{X})$ a polynomial of some degree, with coefficients that can be inferred by including them in the setting of (4.6). We refer the reader to [144] for further details and the corresponding equations that generalise those in (4.4) -(4.5) for the case of non-zero prior mean.

Finally, we emphasize that the Gaussian assumptions in the GPR framework relate to the family of functions that approximate the functional form of the simulator (i.e. F). The GPR framework impose neither restrictions on the distribution of the input space $\mathcal{X}$ nor on the distribution of the simulated outputs $F(\mathcal{X})$. The GPR framework can be used to propagate uncertainties for an arbitrary distribution of inputs which may, in turn, yield a distribution of non-Gaussian outputs.

4.2.1 Validation measures

Numerous validation measures can be employed to determine the goodness of fit of an emulator [137]. A common choice is a unit-less measure called the predictivity coefficient (or coefficient of determination R^2) Q defined by:

$$Q(Y, \hat{Y}) = 1 - \frac{\sum_{i=1}^{N_t} (Y_i - \hat{F}(X_i^*))^2}{\sum_{i=1}^{N_t} (\bar{Y} - Y_i)^2} \quad (4.9)$$

where $Y = \{Y_i\}_{i=1}^{N_t}$ denotes the observed simulator outputs over a set of N_t testing points $\{X_i^*\}_{i=1}^{N_t}$ (distinct from training data), $\bar{Y}$ denotes their empirical mean, and $\hat{Y} = \{\hat{F}(X_i^*)\}_{i=1}^{N_t}$ are the GP metamodel predictions computed by (4.4).

Van Gelder et al [137] propose that a metamodel be constructed with as few training samples as possible, through a new strategy that is independ-

ent of the type of metamodel, by systematically adding samples in steps until the constructed metamodel meets defined validation criteria. If the evaluation measures indicate a poor fit (low Q and high error) then the number of sampling points, within a fixed computation budget, are increased and the model fit is re-analysed; the process being repeated until the goodness of fit criteria (high Q and low error) are satisfactory i.e. addition of more sampling points to the training datasets does not lead to any significant increase or change (say by more than 5%) in the predictivity coefficient. Since the desired accuracy is problem-dependent, the goodness of fit of the metamodel is evaluated based on the converged value of our chosen indicator of accuracy i.e. predictivity coefficient. We apply and demonstrate this new metamodel-type independent strategy in Section 5.2.1, to obtain mean and stochastic variance GP response surfaces by analysing their goodness of fit at test data points, and by observing the convergence in predictivity coefficient values over sampling sets of increasing size.

In addition, since the GP framework produces a statistical distribution of predicted outputs, it is important to quantify the uncertainty in the GP emulator response at validation/testing data points. To this end, we make use of a credible interval technique, as proposed in Bastos and O'Hagan [165]. At any given test point $\mathbf{X}^*$, the GP framework that yields expressions (4.4) and (4.5) enables us to compute credible intervals defined by

$$C_\alpha(\mathbf{X}^*) = [\hat{F}(\mathbf{X}^*) - \tau\sqrt{\hat{K}(\mathbf{X}^*)}, \hat{F}(\mathbf{X}^*) + \tau\sqrt{\hat{K}(\mathbf{X}^*)}]) \tag{4.10}$$

where τ_α is the $(1-\frac{\alpha}{2})$ quantile of the standard Normal distribution and $\hat{F}(\mathbf{X}^*)$ and $\hat{K}(\mathbf{X}^*)$ are the values that we obtain from expressions (4.4) and (4.5), respectively. According to Bastos and O'Hagan [165], evaluating the statistical accuracy of the emulator response can be performed by computing the percentage of observations (obtained from a set of testing points) that fall within this interval. An accurate emulator response should yield credible intervals that contain a proportion of approximately $(1 - \frac{\alpha}{2})$ of those observations. In Section 5.2.1, we evaluate the uncertainty of the emulators trained in this study.

4.3 Proposed methodology for the emulation of S-BPS

In this section we propose a methodology to emulate the stochastic response of S-BPS described by (Eq. 4.1). Again we reiterate that while $\mathbf{X}$ is a vector of random variables representing design input parameters (e.g. wall thickness, window transmittance), these are assumed to have a well-characterised probability density function. In contrast, ω is a variable associated with the aleatoric nature of the S-BPS associated with occupants presence and their interactions with the envelope (e.g. use of windows and shading devices). We assume that we only have access to the distribution of ω via repetitions of the simulator at a training point (as per formulation Eq. 1.1 and discussed in sections 1.2, 2.3). In this case, our aim is to construct metamodels for the mean response $f_d(\mathbf{X})$ and the variance, $V(\mathbf{X})$, of the stochastic component of the simulator. More precisely, we consider the log-transformed [1] variance $\log V(\mathbf{X})$ to ensure positivity of the estimates produced by the metamodel.

Figure 4.1 illustrates the deployment of the S-BPS (Figure 2.4) within our Monte Carlo uncertainty propagation scheme for the generation of training datasets with which to construct metamodels.

4.3.1 Gaussian process fitted to the log-variance $\log V(\mathbf{X})$.

We consider a set of training points $\{\mathbf{X}_n\}_{n=1}^N \in \mathcal{X}$ in the input space and denote by $Y_n = f(\mathbf{X}_n, \omega)$ [2] the random variable that arises from evaluating the stochastic simulator at each training point $\mathbf{X}_n$. For each $\mathbf{X}_n$ we generate a set of J_n realisations (of order 10^2 to 10^3 for the case studies in this thesis) of the stochastic simulator. We denote these realisations by $\{Y_n^{(j)}\}_{j=1}^{J_n}$, and we assume that these are independent and identically distributed (iid) according to the law of Y_n. We compute the sample mean, $\overline{Y}_n$, and sample variance, s_n^2, of these realisations via the following expressions:

[1] Since variance is always positive, the log transformation of the sample variance avoids prediction of negative variance values.

[2] For the purposes of simplicity in notation, in this section we also use Y_n to denote the simulator response, so that Y_n is now a random variable since the function f is stochastic .

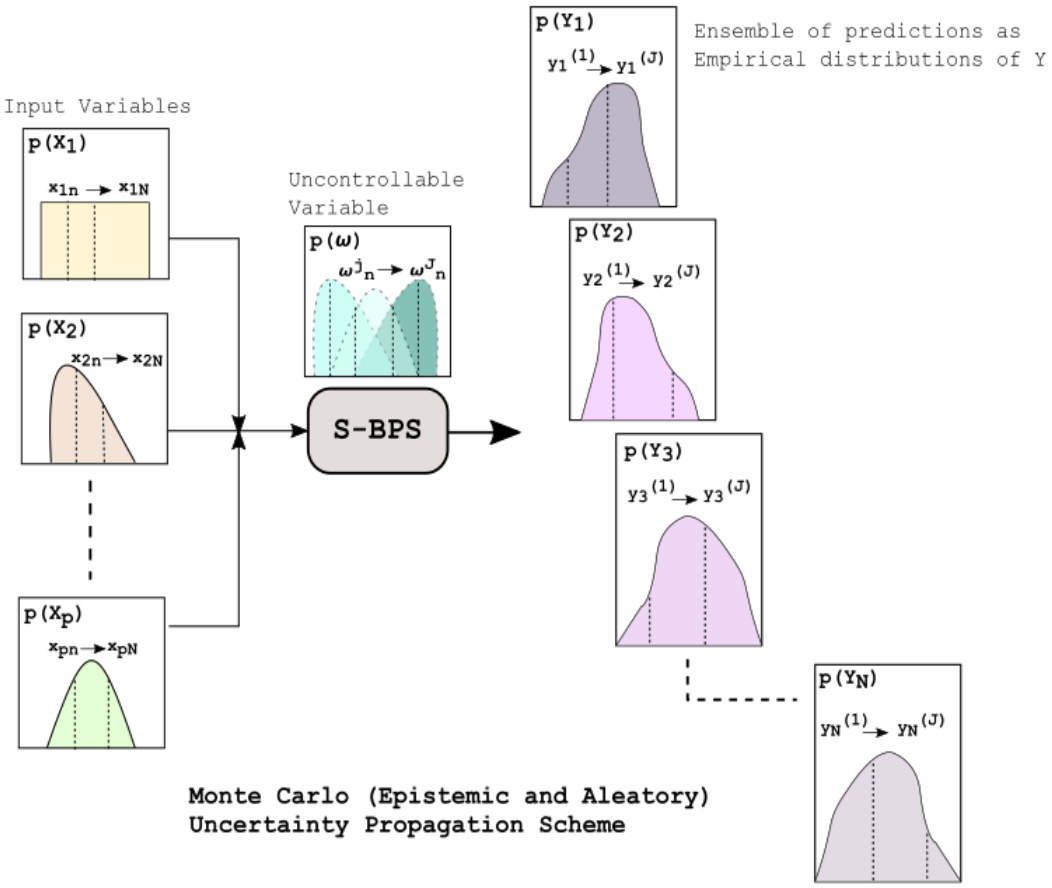


Figure 4.1: Monte Carlo Uncertainty Propagation Scheme for S-BPS

$$\overline{Y}_n = \frac{1}{J_n}\sum_{j=1}^{J_n} Y_n^{(j)} \quad \text{and} \quad s_n^2 = \frac{1}{(J_n-1)}\sum_{j=1}^{J_n}(Y_n^{(j)} - \overline{Y}_n)^2. \tag{4.11}$$

Following the work of [172, 171], we use the log of the sample variance $\{\log(s_n^2)\}_{n=1}^N$ computed at the training points to construct a GP metamodel for the unknown log-variance $\log(V(\mathbf{X}))$. In order to introduce an observational model of the form (4.2) suitable for GPR, we note that under Gaussian assumptions on the distribution of the observations $\{Y_n^{(j)}\}_{j=1}^{J_n}$, the log-transform introduces a bias β_n. More specifically, according to [172, 171], we know that:

$$\log V(\mathbf{X}_n) = \mathbb{E}(\log(s_n^2)) + \beta_n \tag{4.12}$$

where

$$\beta_n = -\Psi\left(\frac{J_n-1}{2}\right) + \log\left(\frac{J_n-1}{2}\right).$$

In the previous expression Ψ denotes the digamma function. Motivated by (4.12) we thus consider the following observational model for the true log-variance at the training points:

$$\log(s_n^2) + \beta_n = \log V(\mathbf{X}_n) + \eta_n \tag{4.13}$$

where $\eta_n \sim \mathcal{N}(0, \sigma^2)$. The training set for the metamodel $\log V(\mathbf{X})$ is thus defined by $\mathcal{D}_v = \{(\mathbf{X}_n, \log(s_n^2) + \beta_n)\}_{n=1}^N$. We use the GPR described in subsection 4.2 to fit a stationary Gaussian process $\log V(\mathbf{X})$, i.e. we assume a centred (i.e. zero mean) GP prior on $\log V(\mathbf{X})$:

$$\log V(\mathbf{X}) \sim \mathcal{GP}(0, k_{\theta_v}(\mathbf{X}, \mathbf{X}^{'})) \tag{4.14}$$

with a covariance function defined by (4.3) and hyperparameters θ_v. From the GP framework we obtain mean predictions for $\widehat{\log V}(\mathbf{X}^*)$ and variance $\hat{K}_v(\mathbf{X}^*)$ defined via the following expressions:

$$\widehat{\log V}(\mathbf{X}^*) = K_{\theta_v}^T(\mathbf{X}^*)(K_{\theta_v} + \sigma^2\mathbf{I})^{-1}\boldsymbol{\xi} \tag{4.15}$$

$$\hat{K}_v(\mathbf{X}^*) = k_{\theta_v}(\mathbf{X}^*, \mathbf{X}^*) - K_{\theta_v}^T(\mathbf{X}^*)(K_{\theta_v} + \sigma^2\mathbf{I})^{-1}K_{\theta_v}(\mathbf{X}^*) \tag{4.16}$$

where K_{θ_v}, $K_{\theta_v}(\mathbf{X}^*)$ are defined analogous to those in (4.4)-(4.5), and where $\boldsymbol{\xi} = (\log(s_1^2)+\beta_1, \ldots, \log(s_N^2)+\beta_N)$. Hyperparameters θ_v and σ^2 are estimated by maximising the posterior density as described earlier. In this case the marginal likelihood is given by:

$$p(\mathcal{D}_v|\theta_v, \sigma^2) = \frac{1}{2}\boldsymbol{\xi}^T(K_{\theta_v} + \sigma^2\mathbf{I})^{-1}\boldsymbol{\xi} + \frac{1}{2}\log|K_{\theta_v} + \sigma^2\mathbf{I}| + \frac{N}{2}\log 2\pi \quad (4.17)$$

4.3.2 Gaussian process fitted to the mean response $f_d(\mathbf{X})$.

Let us note that if the realisations $\{Y_n^{(j)}\}_{j=1}^{J_n}$ are iid samples from a Gaussian / non-Gaussian distribution, then from the Central Limit Theorem [173] we know that $\overline{Y}_n$ follows a Gaussian distribution with mean $f_d(\mathbf{X}_n)$ and variance $V(\mathbf{X}_n)/J_n$; we write this as follows:

$$\overline{Y}_n \sim \mathcal{N}(f_d(\mathbf{X}_n), V(\mathbf{X}_n)/J_n), \qquad n = 1, \ldots N \quad (4.18)$$

If the assumption of normality on $\{Y_n^{(j)}\}_{j=1}^{J_n}$ is not satisfied, expression (4.18) provides a good approximation for sufficiently large sample size J_n. Therefore, from (4.18) we consider the following observational model for $f_d(\mathbf{X})$:

$$\overline{Y}_n = f_d(\mathbf{X}_n) + \epsilon_n, \qquad n = 1, \ldots, N \quad (4.19)$$

where $\epsilon_n \sim \mathcal{N}(0, V(\mathbf{X}_n)/J_n)$. We now follow the approach proposed in [167] where the true (unknown variance) $V(\mathbf{X}_n)$ is replaced by the predictions of the metamodel constructed in (4.15). More specifically, we assume that:

$$\overline{Y}_n = f_d(\mathbf{X}_n) + \epsilon_n, \qquad \epsilon_n \sim \mathcal{N}\Big(0, \frac{1}{J_n}e^{\widehat{\log V}(\mathbf{X}_n)}\Big). \quad (4.20)$$

We note that (4.20) defines an observational model for $f_d(\mathbf{X})$ that can be used within a GPR framework with heteroscedastic (i.e. non-constant and input-dependent) variance [172, 171] to construct an emulator for $f_d(\mathbf{X})$. To this end, we define the training set $\mathcal{D}_Y = \{(\mathbf{X}_n, \overline{Y}_n)\}_{n=1}^N$. We further assume a centred GP prior for f_d. That is:

$$f_d(\mathbf{X}) \sim \mathcal{GP}(0, k_{\theta_y}(\mathbf{X}, \mathbf{X}')) \quad (4.21)$$

Under the GPR framework, predictions at testing points are then given via the mean $\hat{f}_d(\mathbf{X}^*)$ and variance $\hat{K}_y(\mathbf{X}^*)$ defined by:

$$\hat{f}_d(\mathbf{X}^*) = K_{\theta_y}^T(\mathbf{X}^*)(K_{\theta_y} + \Sigma_{\hat{\mathbf{V}}})^{-1}\overline{\mathbf{Y}} \tag{4.22}$$

$$\hat{K}_y(\mathbf{X}^*) = k_{\theta_y}(\mathbf{X}^*, \mathbf{X}^*) - K_{\theta_y}^T(\mathbf{X}^*)(K_{\theta_y} + \Sigma_{\hat{\mathbf{V}}})^{-1}K_{\theta_y}(\mathbf{X}^*), \tag{4.23}$$

where K_{θ_y} and $K_{\theta_y}(\mathbf{X}^*)$ are defined in a similar fashion to K_{θ_v} and $K_{\theta_v}(\mathbf{X}^*)$ above (i.e. with k_{θ_v} replaced by k_{θ_y}) and where $\Sigma_{\hat{\mathbf{V}}} = \text{diag}(\frac{1}{J_1}e^{\widehat{\log V}(\mathbf{X}_1)}, \ldots, \frac{1}{J_N}e^{\widehat{\log V}(\mathbf{X}_N)})$ and $\overline{\mathbf{Y}} = (\overline{Y}_1, \ldots, \overline{Y}_N)^T$. Finally, we estimate the parameter θ_y from the pre-averaged marginal likelihood given by [172]:

$$p(\mathcal{D}_y|\theta_y) = \frac{1}{2}\overline{\mathbf{Y}}^T(K_{\theta_y} + \Sigma_{\hat{\mathbf{V}}}\mathbf{I})^{-1}\overline{\mathbf{Y}} + \frac{1}{2}\log|K_{\theta_y} + \Sigma_{\hat{\mathbf{V}}}| + \frac{N}{2}\log 2\pi \tag{4.24}$$

Similar to the discussion of the preceding subsection, predictions of $f_d(\mathbf{X}^*)$ can be computed via (4.22)-(4.23) with θ_y replaced by $\hat{\theta}_y$. The methodology is now summarised in Algorithm 1. Figure 4.2 illustrates the methodology workflow, according to algorithm 1, for obtaining a validated pair of mean and variance GPs.

Algorithm 1 *Algorithm to construct metamodels for $f_d(\mathbf{X})$ and $\log V(\mathbf{X})$.*

(1) Select/sample N training points from the input space $\{\mathbf{X}_n\}_{n=1}^N \subset \mathcal{X}$

***for** $n = 1, \ldots, N$ **do** (e.g. via LHS)*

(1.1) Compute realisation j of the stochastic simulator $Y_n^{(j)} = F(\mathbf{X}_n, \omega^{(j)})$, $(j = 1, \ldots, J_n)$

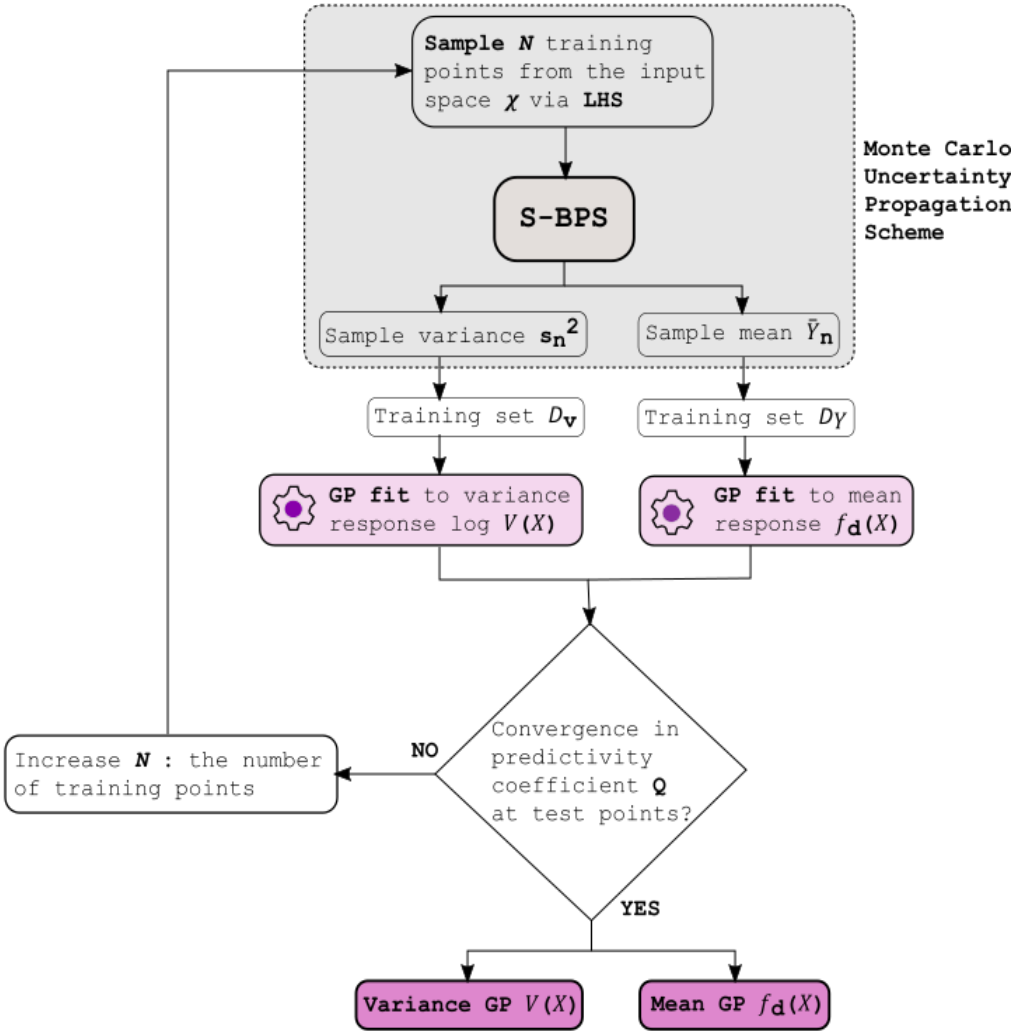
(1.2) Compute the sample mean and variance $\overline{Y}_n$ and s_n^2

end for

(2) Use the training set $\mathcal{D}_v = \{(\mathbf{X}_n, \log(s_n^2) + \beta_n)\}_{n=1}^N$ to fit a GP to $\log V(\mathbf{X})$ (given by (4.15) with θ_v and σ^2 computed for example via MAP estimation).

(3) Use the estimate from (2) and the training set $\mathcal{D}_Y = \{(\mathbf{X}_n, \overline{Y}_n)\}_{n=1}^N$ to fit a GP with input-dependent variance to the mean response $f_d(\mathbf{X})$ (given by (4.22)-(4.23) and θ_y computed via MAP estimation).

end



Algorithm 1 : Construction of Gaussian Process (GP) Emulator to S-BPS

Figure 4.2: Flow diagram for Algorithm 1

4.3.3 Emulation of the stochastic response

Constructing the metamodels in the previous sections enable us to capture uncertainty due to the stochasticity of the simulator at a given input point $\mathbf{X}$. This is relevant for the propagation of uncertainty through complex models [here through an inexpensive metamodel - a plausible approximation of our simulator] and its subsequent quantification under stochasticity. Note that a particular case of the stochastic simulator from (1.1) arises when the function f can be written in the following form:

$$Y = f(\mathbf{X}, \omega) = f_d(\mathbf{X}) + f_s(\mathbf{X}, \omega), \qquad f_s(\mathbf{X}, \omega) \sim \mathcal{N}(0, V(\mathbf{X})). \tag{4.25}$$

where, as before, $f_d(\mathbf{X})$ is the mean response. In this case $f_s(\mathbf{X}, \omega)$ is a Gaussian random variable with input-dependent variance. Under the modelling assumption given by (4.25), following formulation in [41], the mean response $f_d(\mathbf{X})$ and the variance $V(\mathbf{X})$ fully characterise the stochastic response at a given point $\mathbf{X}$ in the input space. Constructing metamodels according to the methodology described in subsections 4.3.1-4.3.2 allows us to simulate the response from the stochastic simulator. Indeed, simulations of the variable Y at at testing point $\mathbf{X}$ can be obtained by:

$$Y = f_d(\mathbf{X}) + \sqrt{V(\mathbf{X})}\xi, \qquad \xi \sim \mathcal{N}(0, 1). \tag{4.26}$$

Similarly, two-sided α-tailed credible intervals can be defined by:

$$\mathcal{CI}_\alpha(\mathbf{X}) = \left[f_d(\mathbf{X}) - \tau_\alpha \sqrt{V(\mathbf{X})}, f_d(\mathbf{X}) + \tau_\alpha \sqrt{V(\mathbf{X})} \right] \tag{4.27}$$

where, as before, τ_α is the $(1 - \frac{\alpha}{2})$ quantile of the standard Normal distribution. The predictors from the metamodels for $f_d(\mathbf{X})$ and $V(\mathbf{X})$ can be used in expressions (4.26)-(4.27) for the efficient emulation of our stochastic simulator and to approximate credible intervals. In Section 5.2.3, we assess using (4.27) the capabilities of the emulator to capture the stochastic observations obtained from the stochastic building performance simulator.

4.4 Global sensitivity analysis through metamodeling

We now discuss how the proposed approach to quantify uncertainties in S-BPS can be combined with a Global Sensitivity Analysis (GSA) technique to decompose the total uncertainty in the model output into the respective sources of uncertainty in the model inputs [65]. We consider a variance based method, based on sensitivity indices that enable us to quantify the importance of a given uncertain input relative to others. These indices are the quantitative measures of knowing an impact of uncertain sources on the model predictions (so that those uncertainties can be subsequently addressed if possible).

4.4.1 The deterministic case

Let us consider again the case of a deterministic simulator $Y = F(\mathbf{X})$ discussed in subsection 4.2. We denote by I a set of indices from $\{1, \ldots, p\}$ with cardinality denoted by $|I|$. We use standard index notation for which X_I denotes a vector that contains the components of $\mathbf{X} = (X_1, \ldots, X_p)$ with indices in I. For example, if $I = \{1, 4\}$, then $X_{\{1,4\}} = (X_1, X_4)$. Furthermore, we denote by $\mathbf{X}_{\sim I}$ the vector with all components of $\mathbf{X}$ except those that belong to the set I. Using the total variance decomposition theorem, the unconditional variance of the model output Y can be expressed as [95]:

$$\mathbb{V}_{\mathbf{X}}[Y] = \sum_{i=1}^{p} \sum_{|I|} \mathcal{V}_I(Y), \tag{4.28}$$

where

$$\mathcal{V}_I(Y) = \mathbb{V}_{X_I}[\mathbb{E}_{\mathbf{X}_{\sim I}}(Y|X_I)]. \tag{4.29}$$

In (4.29) we use $\mathbb{E}_\kappa(Z)$ and $\mathbb{V}_\kappa(Z)$ to denote expectation and variance of Z with respect to κ. Let us note that when $I = \{i\}$ and $I = \{i, j\}$, for example, we have

$$\mathcal{V}_i(Y) = \mathbb{V}_{X_i}[\mathbb{E}_{\mathbf{X}_{\sim i}}(Y|X_i)] \tag{4.30}$$

and

$$\mathcal{V}_{i,j}(Y) = \mathbb{V}_{X_i,X_j}[\mathbb{E}_{\mathbf{X}_{\sim i,j}}(Y|X_i, X_j)] - \mathcal{V}_i(Y) - \mathcal{V}_j(Y), \tag{4.31}$$

respectively.

Sobol indices of $Y = F(\mathbf{X})$ with respect to controllable parameters is defined by [95]:

$$S_I(Y) = \frac{\mathcal{V}_I(Y)}{\mathbb{V}_{\mathbf{X}}(Y)}. \tag{4.32}$$

For $I = \{i\}$, S_i represents the main (first order) effect giving the measure of X_i on the model predictions Y. The second order effect S_{ij} expresses sensitivity of the model outcomes to the interaction between the variables X_i and X_j without the individual effects of X_i and X_j; higher order indices represent third, fourth, ... order interaction effects. From (4.28) it follows that the sum of these indices (all positive, since the variance is always positive) is equal to one (fractions of unconditional variance sum to unity). The larger and close to one an index value is, the higher will be the importance of that input or the group of inputs that is linked to this index. Note that for a model with p inputs, the number of Sobol indices would amount to $2^p - 1$; leading to an intractable number of indices as p increases.

The total sensitivity of the model output variance to an input X_i is given by [95]:

$$S_{T_i}(Y) = \sum_{i \in \subset I} S_I \tag{4.33}$$

For instance, for a model with three input parameters, $S_{T_1} = S_1 + S_{12} + S_{13} + S_{123}$. In practice, when the model function F has multiple input parameters, only the Sobol indices for the main and total effect are computed.

4.4.2 The stochastic case

We follow the approach of [41] for GSA for our S-BPS of the form $Y = f(\mathbf{X}, \omega)$, in which variable ω is treated as an additional uncontrollable parameter. Using the total variance decomposition theorem, the unconditional variance of the

model output Y can be expressed as:

$$\mathbb{V}_{\mathbf{X},\omega}[Y] = \mathcal{V}_\omega(Y) + \sum_{i=1}^{p}\sum_{|I|}[\mathcal{V}_I(Y) + \mathcal{V}_{I,\omega}(Y)], \tag{4.34}$$

where:

$$\mathcal{V}_\omega(Y) = \mathbb{V}_\omega[\mathbb{E}_{\mathbf{X}}(Y|\omega)], \qquad \mathcal{V}_I(Y) = \mathbb{V}_{X_I}[\mathbb{E}_{\mathbf{X}_{\sim I},\omega}(Y|X_I)] - \sum_{I'\subsetneq I}\mathcal{V}_{I'}(Y)), \tag{4.35}$$

$$\mathcal{V}_{I,\omega}(Y) = \mathbb{V}_{X_I,\omega}[\mathbb{E}_{\mathbf{X}_{\sim I}}(Y|X_I,\omega)] - \sum_{I'\subsetneq I}\mathcal{V}_{I',\omega}(Y) - \sum_{I'\subsetneq I}\mathcal{V}_{I'}(Y) - \mathcal{V}_\omega(Y) \tag{4.36}$$

Similar to the deterministic case, Sobol indices of $Y = f(\mathbf{X},\omega)$ with respect to controllable parameters $\mathbf{X}$ can be defined by:

$$S_I(Y) = \frac{\mathcal{V}_I(Y)}{\mathbb{V}(Y)}, \qquad S_{T_i}(Y) = \sum_{i\in\subset I} S_I \tag{4.37}$$

Sensitivity indices that involve the uncontrollable parameter ω are given by:

$$S_{I,\omega}(Y) = \frac{\mathcal{V}_{I,\omega}(Y)}{\mathbb{V}(Y)}, \qquad S_\omega(Y) = \frac{\mathcal{V}_\omega(Y)}{\mathbb{V}(Y)}. \tag{4.38}$$

and the total sensitivity with respect to ω is:

$$S_{T_\omega}(Y) = \sum_{i=1}^{p}\sum_{|I|} S_{I,\omega} + S_\omega. \tag{4.39}$$

which from (4.28) can be written as:

$$\begin{aligned} S_{T_\omega}(Y) = \sum_{i=1}^{p}\sum_{|I|=i} S_{I,\omega} + S_\omega &= \frac{1}{\mathbb{V}(Y)}\Big[\sum_{i=1}^{p}\sum_{|I|=i}\mathcal{V}_{I,\omega}(Y) + \mathcal{V}_\omega(Y)\Big] \\ &= 1 - \frac{1}{\mathbb{V}(Y)}\sum_{i=1}^{p}\sum_{|I|=i}\mathcal{V}_I(Y). \end{aligned} \tag{4.40}$$

The method proposed in [41] then allows us to write (4.37) and (4.41) in terms of $f_d(\mathbf{X})$ and $V(\mathbf{X})$, for which we have constructed GP-based metamodels. Indeed, from (4.30), (4.1) and the assumption of independence between ω and

X_i we have that:

$$\begin{aligned}\mathcal{V}_i(Y) = \mathbb{V}_{X_i}\mathbb{E}_{\mathbf{X}_{\sim i},\omega}(Y|X_i) &= \mathbb{V}_{X_i}[\mathbb{E}_{\mathbf{X}_{\sim i}}[\mathbb{E}_{\omega}(Y|\mathbf{X})|X_i]] \\ &= \mathbb{V}_{X_i}[\mathbb{E}_{\mathbf{X}_{\sim i}}(f_d|X_i)] = \mathcal{V}_i(f_d(\mathbf{X})).\end{aligned} \tag{4.41}$$

More generally, it can be shown that $\mathcal{V}_I(Y) = \mathcal{V}_I(f_d(\mathbf{X}))$. Thus, (4.37) becomes:

$$S_I(Y) = \frac{\mathcal{V}_I(Y)}{\mathbb{V}(Y)} = \frac{\mathcal{V}_I(f_d(\mathbf{X}))}{\mathbb{V}(Y)} = S_I(f_d(\mathbf{X}))\frac{\mathbb{V}_{\mathbf{X}}(f_d(\mathbf{X}))}{\mathbb{V}(Y)} \tag{4.42}$$

where $S_I(f_d(\mathbf{X})) = \frac{\mathcal{V}_I(f_d(\mathbf{X}))}{\mathbb{V}_{\mathbf{X}}(f_d(\mathbf{X}))}$ is the Sobol index of $f_d(\mathbf{X})$. In addition, from the total variance theorem and the definitions (4.1) we have:

$$\mathbb{V}[Y] = \mathbb{V}_{\mathbf{X}}[\mathbb{E}_{\omega}(Y|\mathbf{X})] + \mathbb{E}_{\mathbf{X}}[\mathbb{V}_{\omega}(Y|\mathbf{X})] = \mathbb{V}_{\mathbf{X}}[f_d(\mathbf{X})] + \mathbb{E}_{\mathbf{X}}[V(\mathbf{X})] \tag{4.43}$$

Substitution of (4.43) in (4.42) yields:

$$S_I(Y) = S_I(f_d(\mathbf{X}))\frac{\mathbb{V}_{\mathbf{X}}(f_d(\mathbf{X}))}{\mathbb{V}_{\mathbf{X}}[f_d(\mathbf{X})] + \mathbb{E}_{\mathbf{X}}[V(\mathbf{X})]} \tag{4.44}$$

Note that that the variance of $f_s(\mathbf{X})$ can now be decomposed into the respective contributions of its inputs $\mathbf{X}$:

$$\mathbb{V}_{\mathbf{X}}[f_d(\mathbf{X})] = \sum_{i=1}^{p}\sum_{|I|=i}\mathcal{V}_I(f_d(\mathbf{X})) = \sum_{i=1}^{p}\sum_{|I|=i}\mathcal{V}_I(Y) \tag{4.45}$$

Therefore, combining (4.45) with (4.41) and (4.43) we obtain:

$$\begin{aligned}S_{T_\omega}(Y) = 1 - \frac{1}{\mathbb{V}(Y)}\sum_{i=1}^{p}\sum_{|I|=i}[\mathcal{V}_I(Y)] &= 1 - \frac{\mathbb{V}_{\mathbf{X}}[f_d(\mathbf{X})]}{\mathbb{V}(Y)} \\ &= \frac{\mathbb{E}_{\mathbf{X}}[V(\mathbf{X})]}{\mathbb{V}_{\mathbf{X}}[f_d(\mathbf{X})] + \mathbb{E}_{\mathbf{X}}[V(\mathbf{X})]},\end{aligned} \tag{4.46}$$

Expression (4.44) and (4.46) enable us to compute Sobol indices by using the approximation of $f_d(\mathbf{X})$ and $\log V(\mathbf{X})$ that we obtained from the GP metamodel constructed in subsections 4.3.1-4.3.2. In other words, we use the

following approximations:

$$S_I(Y) \approx S_I(\hat{f}_d(\mathbf{X})) \frac{\mathbb{V}_{\mathbf{X}}(\hat{f}_d(\mathbf{X}))}{\mathbb{V}_{\mathbf{X}}[\hat{f}_d(\mathbf{X})] + \mathbb{E}_{\mathbf{X}}[\exp(\widehat{\log V}(\mathbf{X}))]}, \tag{4.47}$$

$$S_{T_\omega}(Y) \approx \frac{\mathbb{E}_{\mathbf{X}}[\exp(\widehat{\log V}(x))]}{\mathbb{V}_{\mathbf{X}}[\hat{f}_d(\mathbf{X})] + \mathbb{E}_{\mathbf{X}}[\exp(\widehat{\log V}(\mathbf{X}))]} \quad . \tag{4.48}$$

For the computation of sensitivity indices, accounting for both individual and total effects of our parameters, we use the *Sensitivity Analysis For Everybody (SAFE)* package developed in [174] and available at `safetoolbox.info`. After obtaining well trained and validated GP emulators for $f_d(\mathbf{X})$ and $\log V(\mathbf{X})$, we study the convergence of our sensitivity indices (4.47) and (4.48) by running this inexpensive substitute to the stochastic BPS for different sample sizes, and correspondingly perform uncertainty propagation and decomposition. Since the sensitivity indices (4.47) and (4.48) involve statistical quantities computed from samples, a convergence study is required to ensure that the enough samples are used to accurately approximate those sensitivity indices.

For the convergence of sensitivity indices we adopt the criterion proposed in [39], as the index value remains stable (or changes to a limited degree within the credible interval) when computed at multiple independent sample sets of increasing size. This convergence criterion uses the bootstrap technique to provide a measure of uncertainty in the value of the sensitivity index [175]. Since the sensitivity index value is an estimate obtained from the drawn samples, so that its true value cannot be known, it is advisable to account for uncertainty in estimates of sensitivity index due to the sampling variability. The application of a bootstrap technique, along with the chosen sampling scheme, allows for the estimation of an approximate probability distribution around the sensitivity index value, giving confidence intervals as lower and upper bounds for the measure of uncertainty in its estimate. This is particularly helpful when assessing the error in the index estimate at sampling sets of increasing size. In bootstrapping, the different resamples are randomly drawn with replacement from the original sample of the model input / output. To assess the convergence of the sensitivity indices, we observe the width of the 95% credible intervals of the index distribution obtained via bootstrapping.

In particular, for each input and across different sample sizes, we observe:

$$CI_{indices} = (S_{ub} - S_{lb}) \tag{4.49}$$

where, S_{ub} and S_{lb} are the upper and lower bounds of the probability distribution of the sensitivity index of the input factor at the given sample size. When the value of the width of the confidence interval is close to or stabilizes near to zero or a small threshold value (say, 0.05), the sensitivity index is deemed to have converged. In Section 5.2.4, we present the convergence of main and total effect sensitivity indices of the inputs involved in this study, across multiple sample sizes. Algorithm 2 summarises the algorithm to estimate sensitivity indices for the stochastic simulator with respect to controllable variables.

Algorithm 2 *Algorithm to compute Sobol indices S_I and S_{T_ω}. Consider approximations $f_d(\mathbf{X}) \approx \hat{f}_d(\mathbf{X})$ and $V(\mathbf{X}) \approx \exp(\widehat{\log V}(\mathbf{X}))$ from the metamodels constructed from Algorithm 1.*

(1) Compute Sobol sensitivity indices for $\hat{f}_d(\mathbf{X})$ (i.e. $S_I(\hat{f}_d(\mathbf{X}))$) using standard techniques.

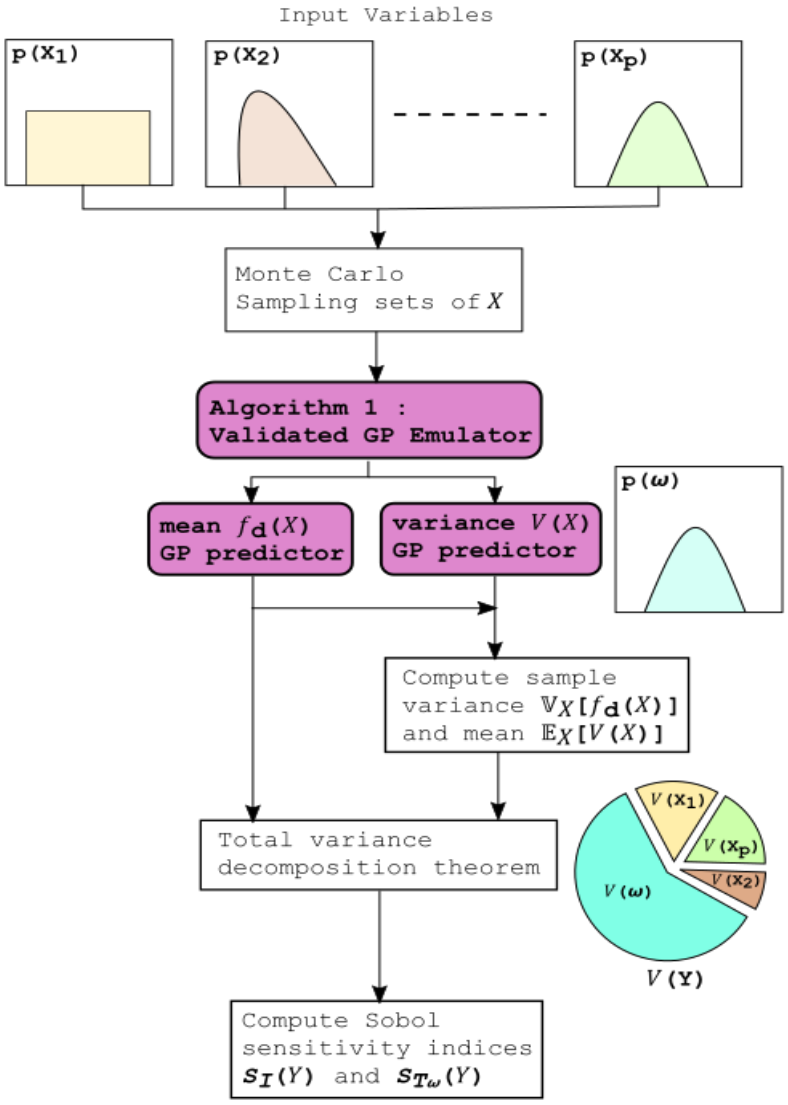
(2) Compute $\mathbb{V}_{\mathbf{X}}[\hat{f}_d(\mathbf{X})]$ via Monte Carlo sampling (i.e. Sample from the distribution of $\mathbf{X}$, evaluate $\hat{f}_d(\mathbf{X})$ via the GP predictor (4.22) and compute sample variance).

(3) Compute $\mathbb{E}_{\mathbf{X}}[\exp(\widehat{\log V}(\mathbf{X}))]$ via Monte Carlo sampling (i.e. Sample from the distribution of $\mathbf{X}$, evaluate $\exp(\widehat{\log V}(\mathbf{X}))$ via the GP predictor (4.15) and compute sample mean).

(4) Use (1)-(3) to compute $S_I(Y)$ and $S_{T_\omega}(Y)$ via expressions (4.47)-(4.48)

end

Figure 4.3 illustrates algorithm 2, employing validated metamodels following algorithm 1 for the efficient computation of sensitivity indices i.e. the quantification and decomposition of the total uncertainty into its respective sources.



Algorithm 2 : Computation of Sobol sensitivity indices through metamodeling

Figure 4.3: Flow diagram for Algorithm 2

4.5 Computational aspects

For the selection of training points to construct the GPs involved in Algorithm 1 we use the LHS sampling strategy. To generate the GP metamodels for the log-variance and the mean response we use the *GPstuff* package for Gaussian process models developed in [168] and available at `github.com/gpstuff-dev/gpstuff`. Note that the GP fitted to the mean response f_d requires a modification of GPstuff to incorporate non-constant variance given by (4.15), as well as the corresponding routines that compute the MAP estimator of hyperparameters.

For a given computational budget, in terms of the maximum number of simulation runs allowed, Algorithm 1 can be implemented with different combinations of N_d training/design vs N_r number of repetitions/samples. In a similar vein to a sequential search methodology, as proposed in [176], to optimise the specification of the input space and the number of stochastic simulation replicates, we specify a fixed total computational budget $C = N_d N_r$. Then, we conduct a number of numerical experiments corresponding to different combinations of the number of design points and associated replicates by systematically increasing the number of design points in steps, and by reducing the replicates to respect our set computational budget. For each combination, we evaluate the goodness of fit of GP response surfaces using the predictivity coefficient and observe the set of these values for the considered computational budget. Depending upon the quality of GP fits (higher the predictivity coefficient better is the GP fit), the initial computational budget is incrementally increased and the procedure for conducting numerical experiments for design point and replicate combinations is repeated for this new budget. In Section 5.2.1 we compare the goodness of GP fits evaluated at test data for different combinations of design points and repetitions of them.

4.6 EmUSA methodology workflow : A summary

Figure 4.4 summarizes the automated workflow for the generation of parameterized EnergyPlus input data files (.idf) combining the base definition of a

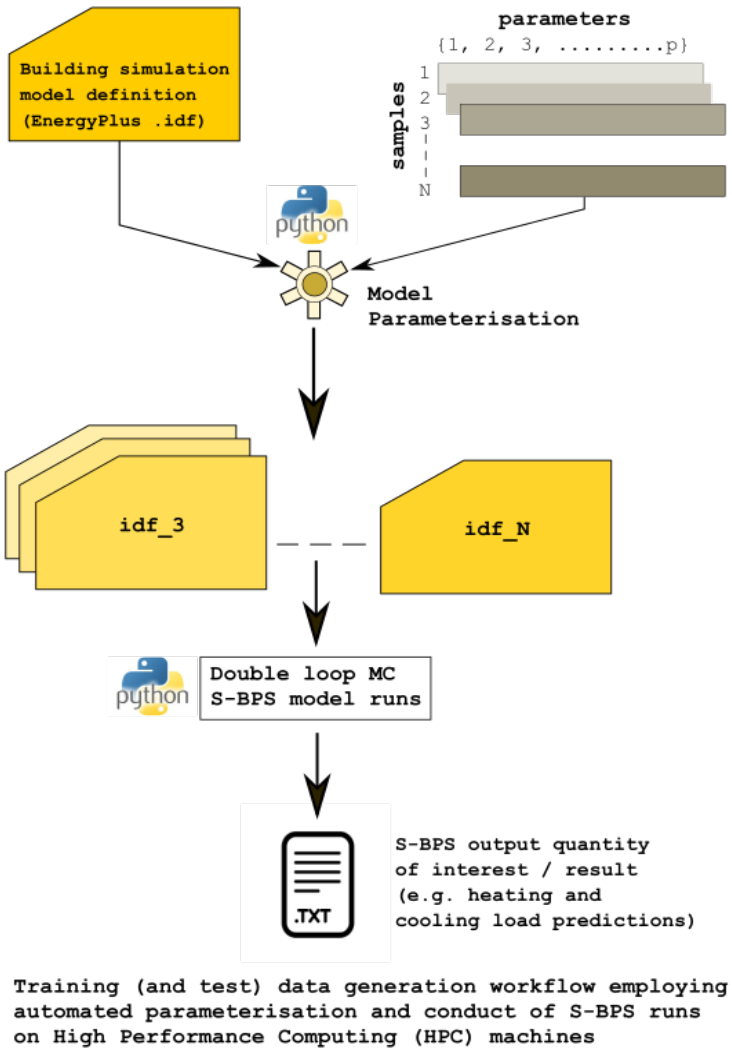


Figure 4.4: Model Parameterisation Workflow

simulation model with the sampling set and writing new input data files for each sample. The automated workflow is flexible in terms of parameters that can be varied (depending upon the uncertain quantity of interests) and as well as the underlying simulation to be used. Figure 4.5 abstracts the computational workflow behind the mathematical methodology of EmUSA framework. The computational workflow can be readily extended to include additional workflow steps such as qualitative parameter screening as shown in Chapter 6.

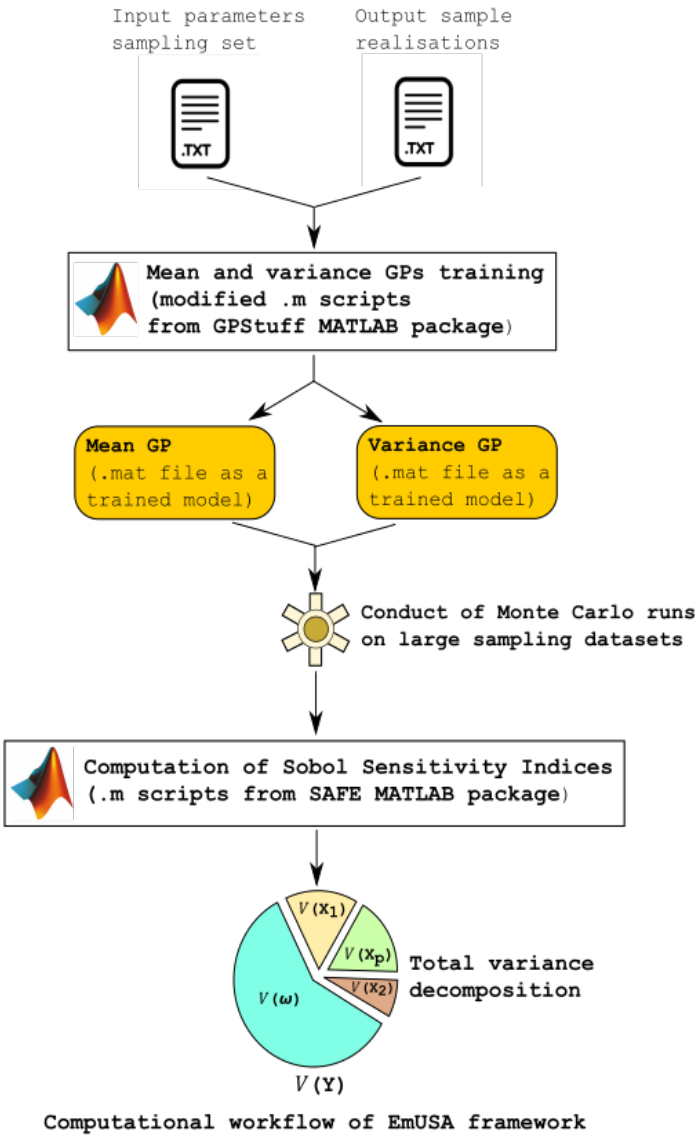


Figure 4.5: Computational Workflow

Chapter 5

Application I - A proof of principle study

In this chapter, we apply the proposed metamodeling approach to simulate the stochastic response from a hypothetical office building model. In this we show, for our particular choice of BPS settings in conjunction with a straightforward model of a monozone office building, that our GPR of $V(\mathbf{X})$ and $f_d(\mathbf{X})$ is both computationally efficient and provides an accurate emulation of the simulator response, in the sense that the stochastic predictions of the simulator fall within confidence intervals provided by the proposed framework. In this, we will also show that the pair of predicted GP emulator response surfaces plausibly capture the underlying behaviour of the detailed BPS program whilst representing the uncertainty bounds due to the emulators themselves; emulators that successfully propagate both the input parameter uncertainty and the stochastic uncertainty due to occupants' interactions, by reproducing the prediction probability distribution at test points, thereby quantifying the total uncertainty in energy predictions. Finally, we combine the proposed metamodeling approach for $V(\mathbf{X})$ and $f_d(\mathbf{X})$ with a variance-based decomposition of total uncertainty in predictions into its respective sources, computed in terms of sensitivity indices.

5.1 Case Study

We analyse the individual and collective impacts of two deterministic phenomena: (d1) heat conduction through a wall construction, and (d2) incoming shortwave irradiation[1]; and through three stochastic phenomena: (s1) occupants' radiative and convective metabolic heat gains due to their presence, (s2) advective heat losses due to ventilation through the opening and closing of windows, and (s3) radiative heat gains due to the raising and lowering of shading devices. The impacts of variation in the deterministic phenomena are studied by assessing the impacts of variations in wall insulation thickness and window transmittance, in plausible ranges of their design uncertainty [2]. We have selected insulation thickness for convenience, as this has a similar effect on heat transfer to variations in conductivity [3]. We realise that we could also have varied density and specific heat capacity for opaque materials, as well as emissivity and conductivity for transparent materials. However, as our primary aim here is merely to present the proof of principle of our framework and to compare the impact of deterministic phenomena against their stochastic counterparts, we believe that our pragmatic choices of parameters are sufficient for our purposes; that this peculiarity does not negatively impact on the usefulness of our evaluation exercise. For stochastic phenomena, we simply repeat simulations with the relevant models enabled, randomly seeding a random number generator for each run.

5.1.1 Experiment set up

As noted earlier (2), we generate our training data through a co-simulation approach, using the Functional Mockup Interface co-simulation standard, in which the dynamic BPS tool EnergyPlus is the master algorithm, controlling

[1]Here, the incoming shortwave irradiation through a window is assumed to be a deterministic component because for the purpose of our study we are considering deterministic input weather excitations.

[2]The basis for this variation is the source of uncertainty in these parameters which can be potentially due to various reasons ranging from manufacturing defects, poor workmanship, early design decisions, etc. In order to accommodate for all the sources of uncertainty that are relevant, the less conservative ranges (which are highly likely during early design) have been considered as plausible for the proof-of-principle case study.

[3]However, we do acknowledge that conductivity is more likely to vary as a result of manufacturing practices, and in response to variations in moisture content, than is thickness; but this latter remains useful for the purpose of demonstration.

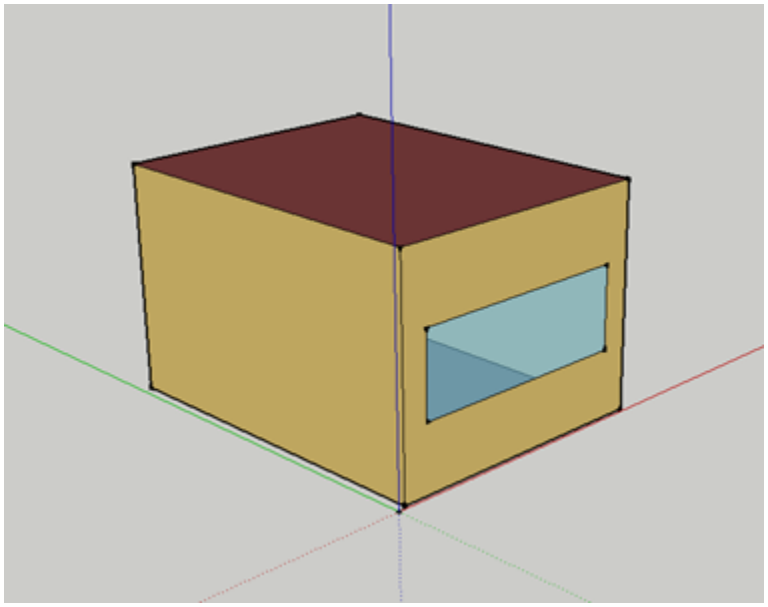


Figure 5.1: Building geometry of hypothetical office

the execution of the multi-agent stochastic simulation tool, No-MASS [22]. In common with Haldi and Robinson's 2011 study of the impact of stochastic models on simulated building performance [28], we employ a model of a simple hypothetical monozone (shoe-box) office building, also located in Geneva, Switzerland ($46.25°N, 6.13°E$, elevation 416m), for which the corresponding Energy Plus weather data file was used. This hypothetical office building has dimensions 3.5m×4.5m, and a floor to ceiling height of 2.8m (Figure 5.1).

The south facade of the building has a double glazed window occupying 40% of its surface area, located above the working plane height of 0.8m above the ground. The wall and window construction compositions used in this study are given in Table 5.1. The sensible internal heat gains are assumed to be a maximum of 400W, collectively accounting for metabolic heat gains due to occupants' presence and casual heat gains due to lighting and any electrical equipment in use whilst occupants are present. The heating (and cooling) and ventilation temperature set-point (SP) schedules are set to 21°C (and 25°C) and 24°C during working hours, according to ANSI/ASHRAE/IES Standard 90.1 for a small office prototype building model, representing an ideal loads HVAC system operation. The graphical user interface to Energy Plus, DesignBuilder 4.3, was used to generate the input data file (.idf) according to the above specifications.

Table 5.1: Thermophysical properties of construction layers : thickness, thermal conductivity, specific heat capacity, density and U-value

Surface	Material	t(m)	λ(W/(mK))	C(J/(kgK))	ρ(kg/m^3)	U(W/(m^{2}K))
Wall	Brick	0.1	0.84	800	1700	0.28
	Insulation	0.111	0.034	1400	35	
	Concrete	0.1	0.51	1000	1400	
	Plaster	0.013	0.4	1000	1000	
Ground	Foam	0.1327	0.04	1400	10	0.26
	Concrete	0.1	1.13	1000	2000	
	Screed	0.07	0.41	840	1200	
	Flooring	0.03	0.14	1200	650	
Roof	Asphalt	0.01	0.7	1000	2100	0.26
	Glass Wool	0.1445	0.04	840	12	
	No mass	Thermal	Resistance	= 0.18		
	Plasterboard	0.013	0.25	896	2800	
Window	Glass	0.003	0.9	750	2500	2
	Air	0.013	0.0262	1005	1.17	
	Glass	0.003	0.9	750	2500	

5.1.2 Parameters studied

We consider two controllable parameters in the S-BPS model described earlier: insulation thickness X_1 and window transmittance X_2. The prediction outputs are annual heating Y_h and Y_c cooling demand per unit floor area. Our objective is to quantify the uncertainty in Y_h and Y_c given uncertainty in $\mathbf{X} = (X_1, X_2)$ as well as the (uncontrollable) uncertainty that arises from occupants' behaviour, encoded in the model described earlier. We assume that X_1 and X_2 are independent and characterised by uniform distributions, and that the values for these design variables are equally likely within the possible uncertain ranges specified [30] [5]. The assignment of a uniform distribution is desirable for the case of poorly defined design variables (e.g. during early design stage) where only an estimate of the lower and upper bounds may be available [105].

For the distribution of insulation thickness we use values between 0.06m and 0.16m to account for plausible variations in insulation thickness (e.g. from low to moderate thickness) within the office building's facades, entailing a corresponding variation in heat conduction through walls. For glazing solar transmittance, the distribution takes values between 0 and 1, effectively modulating the quantity and quality of glazing in the south facade, entailing a corresponding variation in transmitted shortwave radiation [4]. The solar transmittance is applied to the facade glazing ratio, where zero transmittance represents a combination of glazing utilisation (the proportion of the glazed surfaces that is taken up by frames) and transmission (more or less transparent materials) that effectively renders the glazing opaque, and a value of unity represents completely unobstructed glass of perfect transmission properties; with values tending to lie between these (unphysical) extremes. The formulation of the occupant behavioural models and their related parameter settings are adopted as described in Chapter 2 (following Chapman's work on No-MASS [2]).

To apply the proposed framework described in Sections 4.3- 4.4 we treat the annual heating and cooling demand Y_h and Y_c distinctly, and thus construct GP metamodels for each of these output variables. More specifically, we apply Algorithm 1 for each of the model outputs Y_h and Y_c, and subsequently use those GPs within Algorithm 2 to conduct sensitivity analysis for each output

[4]Note that we have not considered in this work uncertainty in visible transmittance, and its possible intercorrelation with solar transmittance, as we are not modelling daylight responsive controls of artificial lights and associated heat gains.

variable. Technical details relevant to the GP prior settings used within the proposed approach can be found in 7.3.

5.2 Results and Discussion

5.2.1 Design points vs repetitions

We note from Algorithm 1 that the proposed metamodeling framework applied to the case study introduced above involves constructing GP emulators of both the mean response $f_d(X_1, X_2)$ and the log variance $V(X_1, X_2)$ for each of the S-BPS response outputs Y_h and Y_c. We also observe that the methodology requires us to select a number of design points (via LHS) within the input space for (X_1, X_2) that we introduced in subsection 5.1.2, and to conduct a number of repetitions of the model outputs at each of those design points. The accuracy of the proposed methodology relies on having a large enough sample of design points (so that input uncertainty $\mathbf{X}$ is well characterised) and a sufficient number of repetitions that can accurately inform, at those design points, the mean and variance response (with respect to ω) of each model output. To arrive at the reasonable computational budget of 3000 simulator runs, we follow the sequential search approach outlined in 4.5 and observe the accuracy of the predictions of each GP involved in the proposed methodology using different combinations of design points and replicates.

Once all GPs have been constructed for each combination, we compute the predictivity coefficient Q and the CI validation measure discussed in subsection 4.2.1. The test data for this case study consists of simulation runs at 60 new sampling points in the input design space, each with 600 repetitions of the stochastic simulator. For the design point and replicate combinations having a budget of 3000 simulator runs to train the emulator, Table 5.2 reports the accuracy in terms of Q obtained at the test data points by the GP emulators of the mean and log-variance. The coefficient of predictivity saturates to a value of more than 0.92 when around 25 design points with 120 repetitions are available for training (Figure 5.2). Any further increase in the number of design points does not lead to significant increase in predictive accuracy. Table 5.3 reports the proportion of test observations that lie inside the 95% CI (according to Eq. 4.10)) predicted by the GP emulators. Also, for the

training combination (25×120), the proportion of test observations within the emulator uncertainty interval is more than 80% without a loss in predictive accuracy. We note that the cooling demand log-variance GP displays less accuracy in the CI measure, arguably due to the large variation in non-constant variance exhibited by occupants' behaviours (observed in the regions of high insulation thickness and window transmittance, in Figure 5.3 right bottom).

Table 5.2: Comparison of quality of prediction Q of fitted GP response surfaces at test data given a fixed computation budget (of 3000 simulation runs) for training. Here the item entry 5×600 represents 600 simulation repetitions at 5 design points and so on.

Output target variable	Emulated quantity	Predictivity coefficient Q					
		5×600	10×300	15×200	20×150	25×120	30×100
Heating demand Y_h	$f_d(\mathbf{X})$	0.9367	0.9949	0.9997	0.9999	0.9999	0.9999
	$\log V(\mathbf{X})$	0.3179	0.9186	0.9159	0.9235	0.962	0.9483
Cooling demand Y_c	$f_d(\mathbf{X})$	0.9704	0.9858	0.9831	0.9782	0.9841	0.9719
	$\log V(\mathbf{X})$	0.0841	0.7765	0.932	0.8993	0.9184	0.9092

Table 5.3: Comparison of emulator uncertainty (Eq. 4.10) of fitted GPs in terms of proportion of test points lying inside 95% uncertainty interval

Output target variable	Emulated quantity	Emulator uncertainty interval CI					
		5×600	10×300	15×200	20×150	25×120	30×100
Heating demand Y_h	$f_d(\mathbf{X})$	0.9999	0.95	0.9999	0.9999	0.9833	0.9999
	$\log V(\mathbf{X})$	0.9999	0.8167	0.6333	0.7	0.8667	0.7167
Cooling demand Y_c	$f_d(\mathbf{X})$	0.9999	0.9833	0.8667	0.7333	0.8167	0.75
	$\log V(\mathbf{X})$	0.6667	0.3833	0.6333	0.35	0.4333	0.4333

5.2.2 Gaussian Process emulation of the mean $f_d(\mathbf{X})$ and log-variance $\log V(\mathbf{X})$ of the S-BPS outputs

Based on the fixed computation budget study of prediction accuracy of GPs (Table 5.2 and 5.3), the combination 25 design / sample points with 120 repetitions of stochastic building performance simulator has been selected to be optimum in order to obtain a plausible pair of GPs predicting mean and log-variance of energy demands Y_h and Y_c. The GP regression surfaces for the computed energy demands' mean and log-variance are shown in Figure 5.3 together with the training set used for the construction of each GP. Each GP response surface is an interpolation surface that predicts mean energy demand value and variance around it due to stochastic occupants' behaviours for any given test point within the uncertainty ranges of the input parameters. The GP regression surfaces allow us to understand the effect of input variables on predictions.

These results show that an increase in insulation thickness lowers heating demand, due to the expected reduction in heat loss through building envelope, analogously to a lowering of thermal conductivity (see Figure 5.3 top-left). As variations in only solar (not visible) transmittance is considered, its direct impact on heating demand is somewhat dampened. Furthermore, with a relatively moderate availability of irradiation during the heating season, variations in transmittance have a relatively lower impact than those due to insulation thickness. Conversely, during the cooling season, our results show the inverse trend; with variations in transmittance have a greater impact, owing to the increased magnitude of available solar irradiation. Transmittance and cooling demand now show positive correlation, as an increase in transmitted shortwave radiation entails a corresponding increase in cooling loads (Figure 5.3 bottom-left). Furthermore, the internal storage of heat can be exacerbated by envelope thermal insulation, which inhibits heat transfer to the outside whilst the indoor temperature is higher, so increasing cooling loads (as shown in Figure 5.3 bottom-left). These results are consistent with the underlying physics abstracted by the BPS.

Furthermore, Figure 5.3 (top and bottom-right) illustrate the heteroscedastic (i.e. non-constant) nature of the variance due to occupants' behaviours and it exhibits a positive correlation with energy demand i.e. the non constant

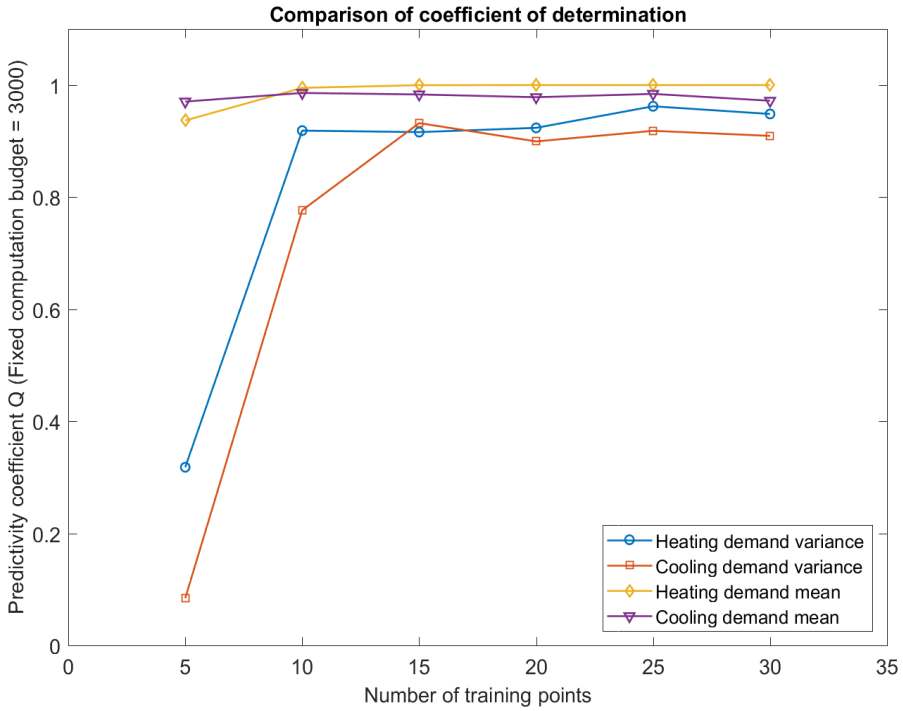


Figure 5.2: Comparison of prediction accuracies of mean and log-variance GPs

variance is high at higher energy demands and vice-versa when compared with their respective mean GP counterparts (top and bottom-left).

5.2.3 Uncertainty quantification using the GP emulator

The validation measures discussed in subsection 5.2.1 enable us to assess the accuracy of the predictive capabilities of each GP (for mean and log-variance). We now assess the accuracy of the proposed methodology to statistically reproduce uncertainty in the outputs of the S-BPS within the same input-space. To this end, we consider the credible intervals introduced in (4.27) for $\alpha = 0.05$ (i.e. 95% CI), where the mean and variance are replaced with those obtained via the GP emulators obtained in the previous subsection. We generate test data with 100 replicates of the S-BPS at a collection of 60 randomly selected input points. Some 95.2% and 96.08% of the simulator's random response values [at each test point] for heating and cooling demands lie within the 95% CI predicted by the pair of (mean and variance) GPs (Figure 5.4), suggesting that this new GP emulator can be used with confidence.

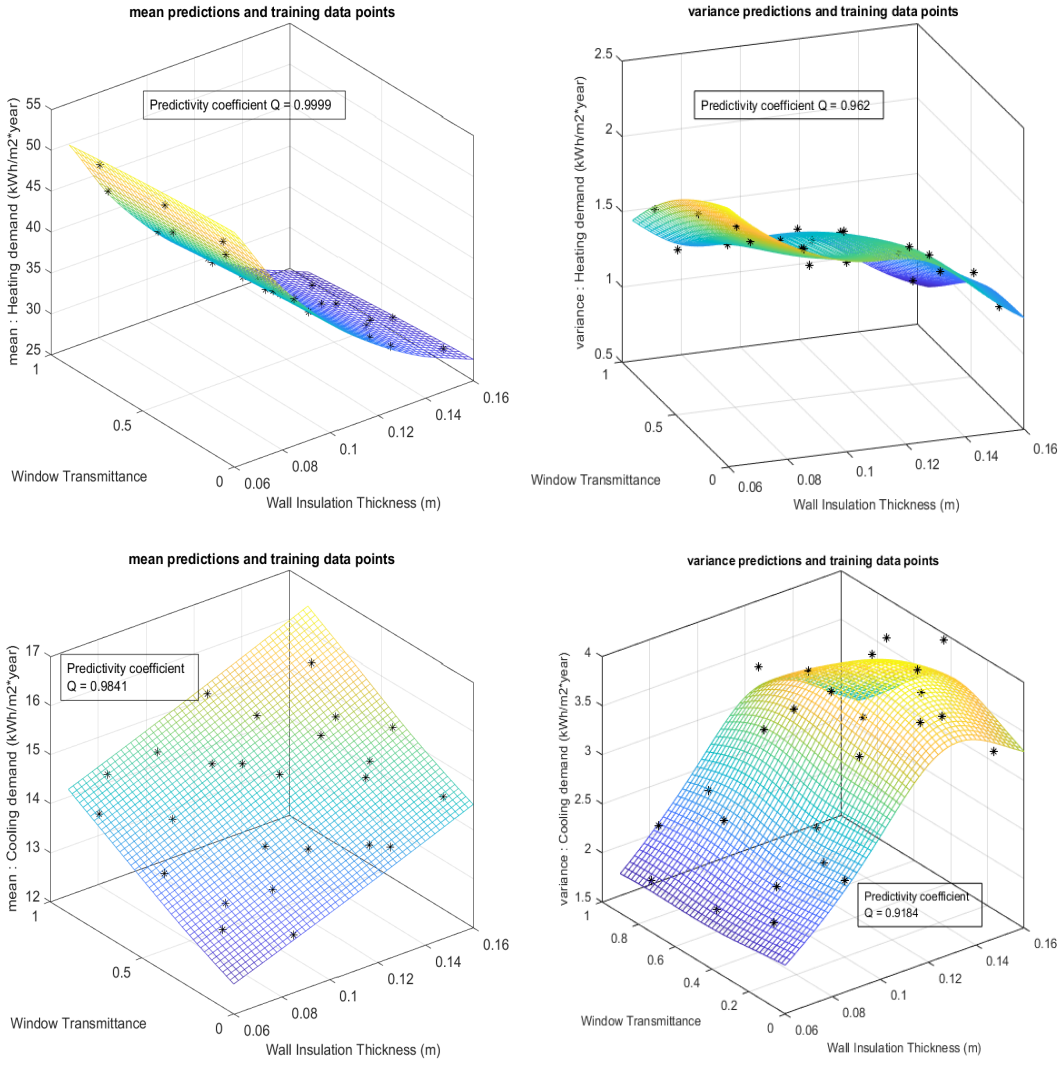


Figure 5.3: GP regression surface fitting to sampled mean (left) and log-variance (right) of heating demand (top) and cooling demand (bottom). The black asterisks are the training points and every point on the mesh is a test data point [not used for training]. The predictivity coefficients values are computed based on the evaluation of fitted GP surfaces at 60 randomly selected test data points, in Figure 5.4 subsection 5.2.3.

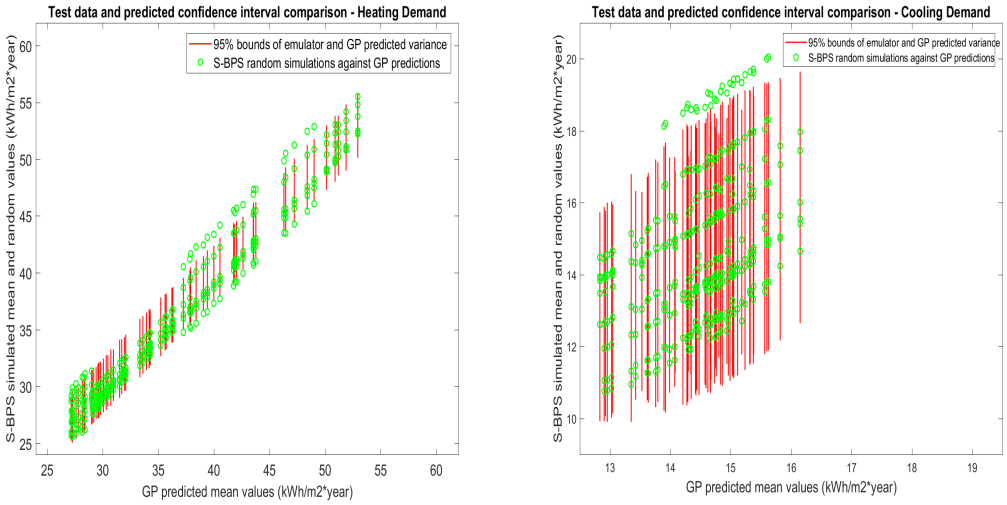


Figure 5.4: Comparison of S-BPS simulated random values within GP predicted credible intervals at test data for heating (a) and cooling (b) demand

To further evaluate the effectiveness of the proposed GP-based metamodeling approach, we compare in Figures 5.5-5.6 the probability density functions (PDFs), at some specific test points, computed by Monte Carlo Simulation, the S-BPS outputs (in red) and outputs from the GP emulator (in blue); this latter by simulating stochastic observations with (4.26) using GP emulators for $f_d(\mathbf{X})$ and $V(\mathbf{X})$).

In Figure 5.7 we also compare cumulative distribution functions (CDFs) of predicted energy demands from the GP emulator with those of the S-BPS simulations, again obtained using Monte Carlo Simulation. A single CDF in these plots represent variation in prediction due to occupants' stochastic behaviour [aleatory uncertainty] and their spread represents uncertainty in predictions due to input parameters [epistemic uncertainty]. Figures 5.5-5.7 offer numerical evidence that the proposed framework enables us to accurately and robustly reproduce the underlying PDFs and CDFs obtained with the S-BPS, at a small fraction of the computational cost, since once the GP emulators have been constructed, S-BPS runs are no longer required for subsequent uncertainty analyses.

In the specific case of heating demand (the top row in Figure 5.7), the collective spread of CDFs (25 kWh/m^2 difference in magnitude) is more than the spread of an individual one (5 kWh/m^2 magnitude difference) - illustrating that epistemic uncertainty due to inputs plays a more dominant role than the

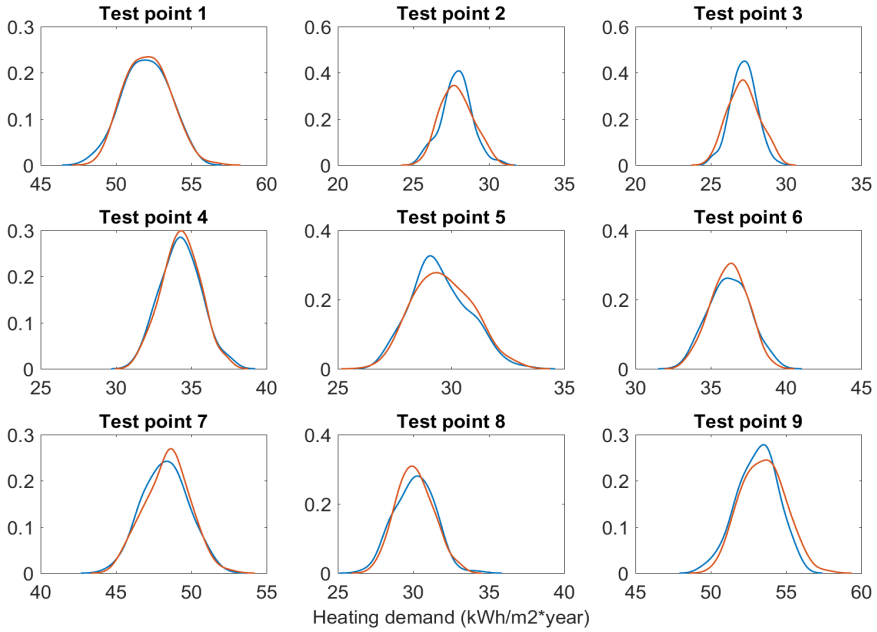


Figure 5.5: Uncertainty propagation comparison between GP predicted (blue) and S-BPS simulated PDFs (red) for heating demand

aleatory uncertainty due to occupants' behaviours. And in the case of cooling demand (the bottom row in Figure 5.7), we observe the opposite trend, the spread of an individual CDF (10 kWh/m^2 magnitude difference) exceeds the collective one (4 kWh/m^2 magnitude difference), showing that the aleatory uncertainty is more dominant than its epistemic counterpart. This is expected, because the stochastic models used in this work (windows and blinds) are essentially biased towards more frequent user interactions during warmer weather (higher ambient temperatures and incident irradiances, these stimulating the opening of windows and the lowering of shading devices) when cooling loads are at their largest.

Finally, the quantitative comparison between the means of heating and cooling demand replicates presented in Figure 5.8 underline how accurate the GP emulator's predictions are. Having demonstrated that the GP emulator is an acceptably accurate means for accelerating stochastic building performance simulations, we proceed now to decompose the total uncertainty in predictions into its component parts by computing sensitivity indices for our case study.

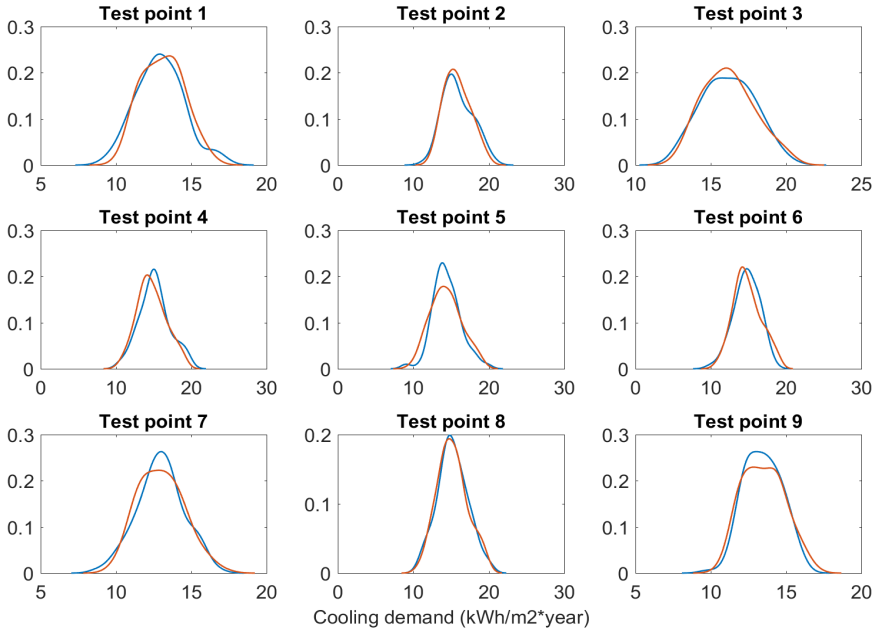


Figure 5.6: Uncertainty propagation comparison between GP predicted (blue) and S-BPS simulated PDFs (red) for cooling demand

5.2.4 Uncertainty decomposition using the GP emulator : Sensitivity analysis

We now employ our GP-metamodeling framework in conjunction with the GSA approach discussed in section 4.4. More specifically, we decompose the total uncertainty in energy demand prediction using the individual and total effect indices of (4.47) and (4.48). The individual effect indices account for the impact of design variables 'insulation thickness' and 'window transmittance' and the total effect index includes that of occupants' behaviour. The index scale is between 0 and 1, with highly influential factors having values close to unity. Table 5.4 reports these indices, computed for both small (500) and large (20000) input sample sizes. In the case of heating demand, and as discussed earlier, epistemic uncertainty due to insulation thickness dominates the uncertainty in predictions, at an individual index value of 0.97. During heating season envelope heat losses (and thus impact of uncertainty in insulation thickness) dominates the energy balance, much more so than occupants' behaviour, since occupant are relatively less likely to open windows (to avoid excessive heat losses) and to lower shades (thus reducing useful transmitted solar heat

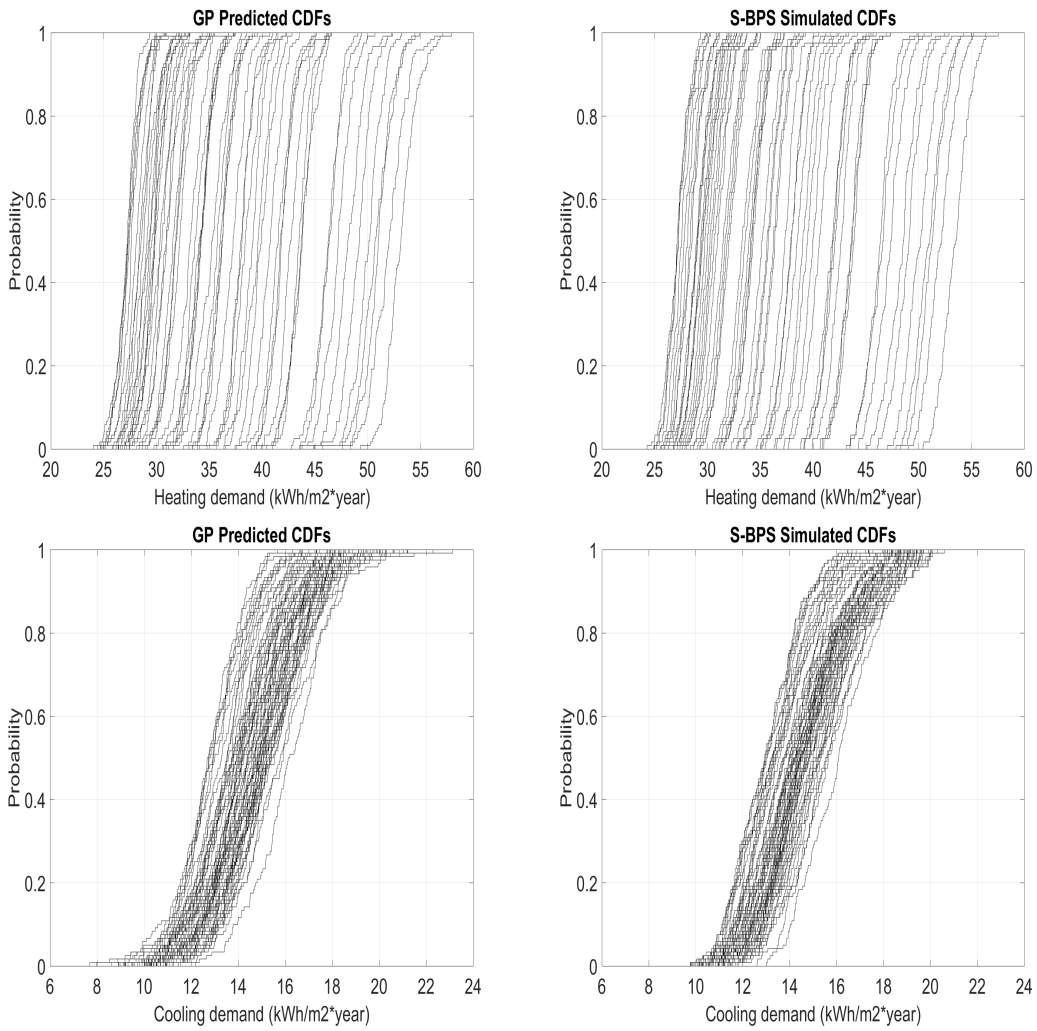


Figure 5.7: Comparison of GP predicted (left) and S-BPS simulated (right) CDFs for heating (top) and cooling (bottom) demand

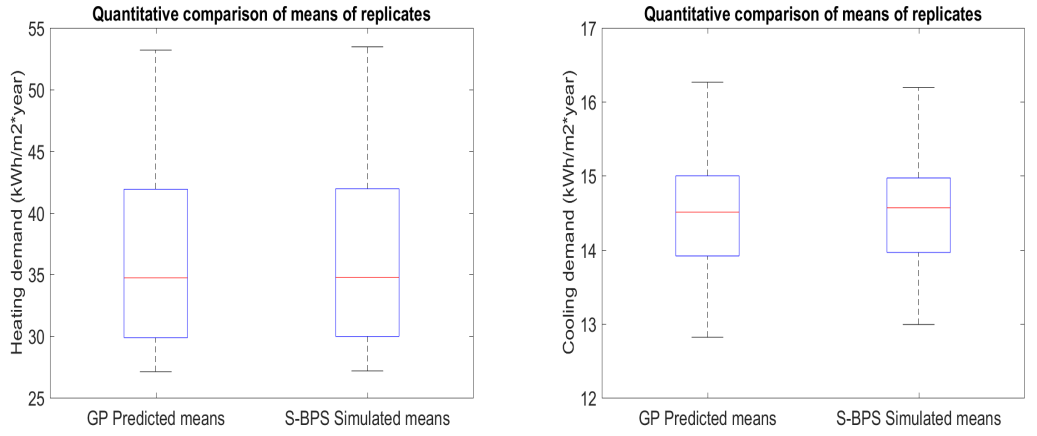


Figure 5.8: Comparison between S-BPS Simulated and GP predicted means of heating (left) and cooling (right) demand replicates in Figures 5.7

Table 5.4: Sensitivity indices quantifying the impacts of thickness, transmittance and occupants' behaviour on the heating and cooling energy demand predictions. S_1 and S_2 are individual effects due to thickness and transmittance and S_{T_w} total effect due to occupants' behaviours.

Output variable	Sample size	Sensitivity indices		
		S_1	S_2	S_{T_w}
Heating demand	Small (500)	**0.9709**	-0.0654	0.0266
	Large (20000)	**0.9715**	-0.0729	0.0266
Cooling demand	Small (500)	0.0752	0.1279	**0.829**
	Large (20000)	0.0527	0.1277	**0.8275**

gains) during the heating season. In contrast, for cooling demands, aleatory uncertainty due to occupants' behaviour is the most influential factor, at total index value of 0.83, as occupants are correspondingly more likely to interact with windows and shading devices, as explained in the previous section.

Figure 5.9 illustrates the combined sensitivity analysis of energy demands as a tornado plot. Sensitivity analysis tornado plot is a plot of percentiles of the distribution of output quantity of interest with median (50^{th} percentile) values as a base value. The tornado plot depicts the amount of variation that resulted into the output variable due to the uncertainties in input variables and it is plotted in decreasing order of importance starting with the positively correlated variables at the top and negatively correlated at the bottom while those in the middle are of negligible importance. In other words, the tornado plot collectively ranks the sensitivity of variables between the two ends of spectrum: positive and negative correlation. For both the energy loads, the amount of variation in values corroborate with the sensitivity indices (Table 5.4). In case of heating demand (plot (a)), 'insulation thickness' is being the most influential variable and also it is negatively correlated (Figure 5.3 top left) with occupancy stochastics are being less influential. On the other hand, in case of cooling demand, the occupancy stochastics are being the most important than 'insulation thickness' and 'window transmittance' again corroborating with the indices in Table 5.4.

Figures 5.10 and 5.11 (top rows) show the convergence of indices values for the emulator evaluations at different sample sizes, ranging from 500 to 20,000. We consider 0.05 as the convergence threshold for the width of the 95% credible interval $CI_{indices}$ (Eq. 4.49), below or close to which the index

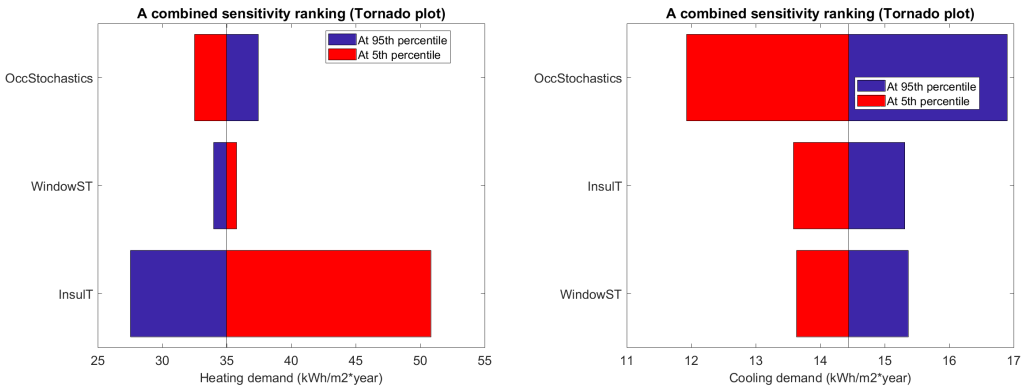


Figure 5.9: Sensitivity Tornado plots for heating (a) and cooling (b) demand

value is deemed to be stable. In the case of heating demand, for the main effect index of 'insulation thickness', the index value converges rapidly (at a small sample size of 1,000) below the threshold, while the other indices values approach the threshold (at values of around 0.2) at large sample sizes of 20,000. The bottom rows of the figures show the convergence of the actual credible intervals around indices values obtained via bootstrapping, demonstrating the trend for convergence at large sample sizes.

Now, the computational cost of a classical GSA framework is given by $N(k+2)$ [35] [60], where N and k are number of samples and inputs, respectively. For the present example with 2 input parameters and 20,000 samples (large sample sizes at which indices converge as shown), the number of simulator runs that we need in our proposed Emulation based Uncertainty and Sensitivity Analysis (EmUSA) framework is a small fraction (3,000 / 80,000 = 0.0375) of that required by the standard approach (see Table 5.5), as the emulation framework only requires simulator runs to train GPs, which can then be employed to conduct sensitivity analyses at negligible computational cost. This represents a significant saving not only in terms of the number of simulator runs, but also of the computational time available to conduct an Uncertainty and Sensitivity Analysis (USA) study. Consequently, the computation time reduces to a manageable c.17 hours, in contrast to the infeasibly large number of days that are required by a classical approach (see Table 5.6). Moreover, these training runs can be readily distributed over a High Performance Computing cluster, thus benefiting from additional acceleration from computational hardware. In this way, it is possible to complete an entire

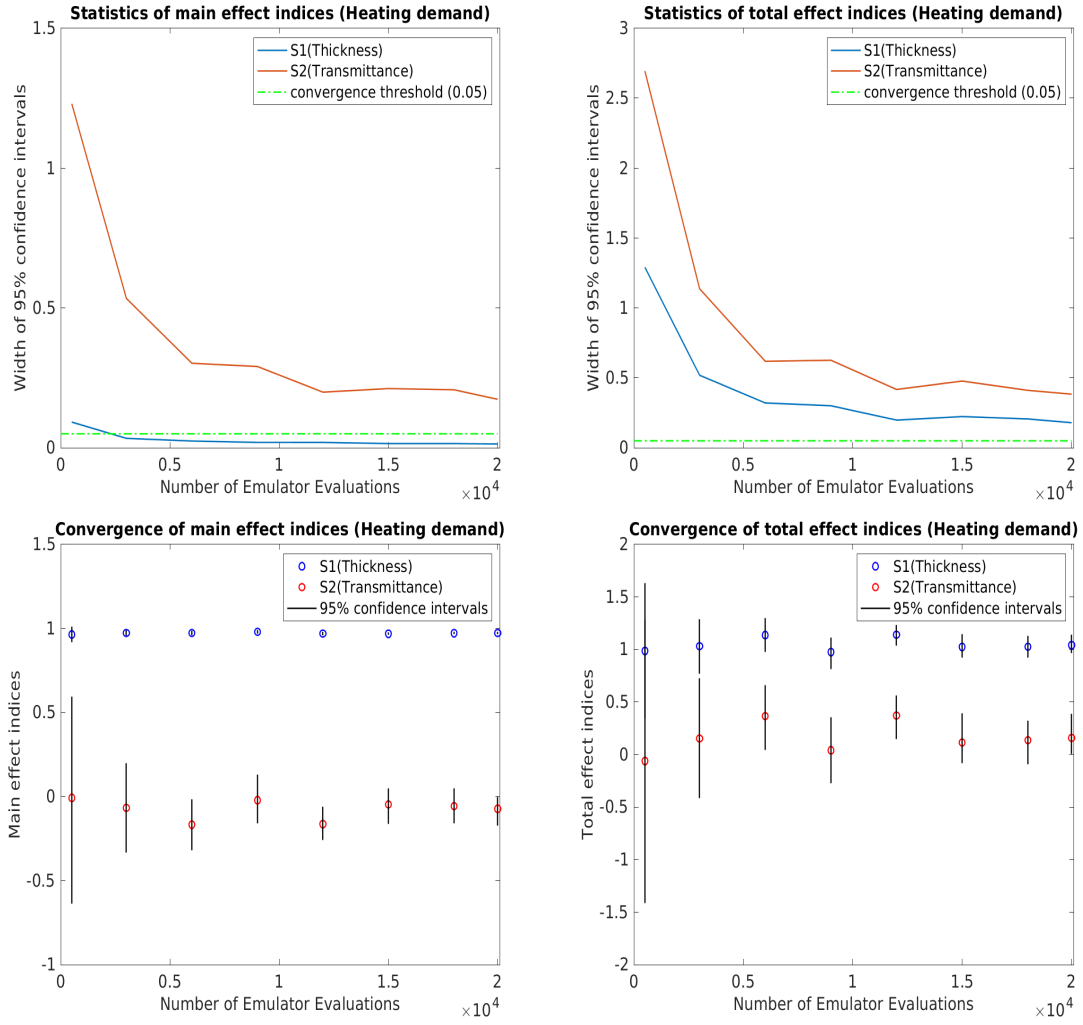


Figure 5.10: Convergence of main (left) and total (right) effect sensitivity indices of design variables for heating demand at distinct sample sets of increasing size. The top row shows the convergence of the width of index credible interval and bottom row shows the actual credible interval (Eq. 4.49).

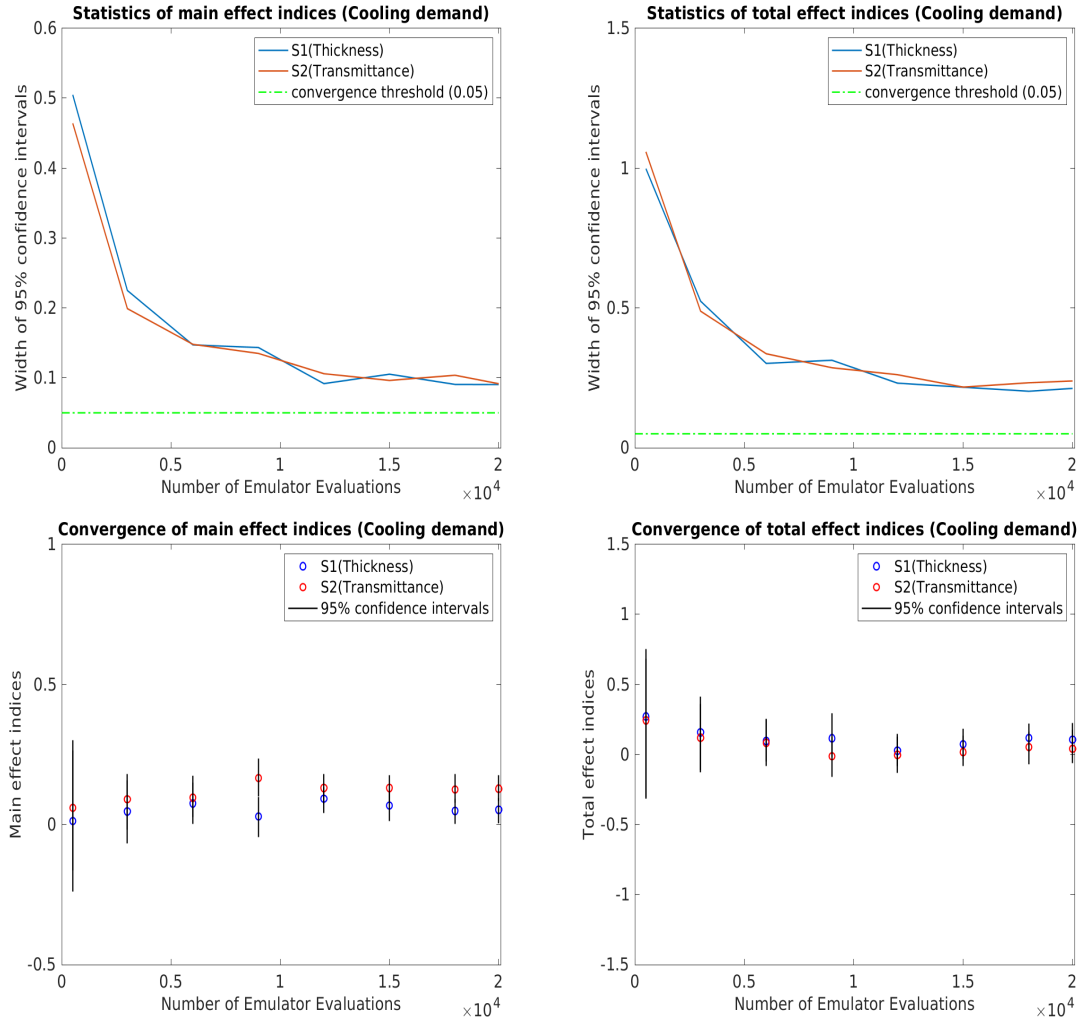


Figure 5.11: Convergence of main (left) and total (right) effect sensitivity indices of design variables for cooling demand at distinct sample sets of increasing size. The top row shows the convergence of the width of index credible interval and bottom row shows the actual credible interval (Eq. 4.49).

Table 5.5: Comparison of computational runs between Classical and EmUSA framework

	Classical	Emulator based	
		Training	Sensitivity Analysis
Simulator runs	80×10^3	3×10^3	-
Emulator runs	-	-	0.5 to 20 $\times 10^3$

Table 5.6: Comparison of computational expense between Classical and EmUSA framework (One simulation run benchmark = 20 seconds)

	Classical	Emulator based	
		Training	Sensitivity Analysis
Simulator runs	18.5 days	16.67 hours	-
Emulator runs	-	-	15 to 20 minutes

USA study using our proposed EmUSA framework in just a couple of hours for this proof-of-principle.

5.3 Conclusions

The purpose of this study was to develop a computationally efficient emulator of a Stochastic Building Performance Simulator (EnergyPlus co-simulated with NoMASS), to support the quantification and decomposition of both epistemic and aleatory uncertainties. Our approach has been to estimate Gaussian Process Regression emulators for both the mean response and the variance of the stochastic component of the simulator, and to deploy these emulators to decompose the total uncertainty in predictions into its respective sources, computed in terms of sensitivity indices.

We have demonstrated the application of this new framework using a relatively simple use case: a small monozone (shoebox) office building, computing both individual and total effect indices. The former accounts for the effect of design variables (e.g. insulation thickness or glazing transmittance), while the latter also accounts for stochastic variance arising from occupants' behaviours (e.g. opening windows and lowering shading devices). From these results we conclude that:

- The pair of fitted GP emulators successfully emulate the behaviour of the stochastic building performance simulator (S-BPS) at the chosen test

design points, with predictivity coefficient $Q_2 > 0.95$, (save for a small number of moderate exceptions).

- The proposed GP Emulator-based UQ framework, requiring simulator runs only to fit the emulators, is highly computationally efficient; reducing the computational cost to a small fraction (0.0375) of its classical counterpart, as simulator runs are required only for training the emulator which subsequently computes at negligible cost.

- The variance-based uncertainty quantification and decomposition framework is an effective means for studying the combination of epistemic and aleatory uncertainties that arise from a S-BPS tool.

- The effects of uncertainties in insulation thickness on heating demand are more significant than those arising from stochasticity in occupants' behaviours; as these behaviours are relatively constrained during the heating season.

- Conversely, the effects of uncertainties arising from stochasticity in occupants' behaviours on cooling demand exceed those arising from uncertainties in effective glazing system transmittance; as these behaviours are less constrained - representing occupants' interactions with windows and shading devices to restore their thermal comfort, with corresponding impacts on cooling demand.

This work demonstrates, for the first time [5] in building simulation studies, the viability of an Emulation based Uncertainty and Sensitivity Analysis (EmUSA) framework for the computationally efficient quantification and decomposition of both epistemic (design variable) and aleatory (occupants'

[5]However, an analogy can be drawn with the applications in other fields of research. For instance, a couple of Uncertainty Quantification studies: 1) on oil reservoir modelling [40] (demonstrated in [41]) and 2) on groundwater flow modelling through rocks [42] (for radioactive waste disposal) employed Gaussian process joint emulation / metamodelling of deterministic and stochastic processes involved therein in order to obtain computationally efficient estimates of mean and variance of output quantity of interests for uncertainty and sensitivity analysis. These studies are important to place this thesis contribution in the wider context of emulating stochastic models to achieve uncertainty analysis estimates in shorter time. In relation to that, the work in this thesis attempts to address the UQ computational challenge for an instance of stochastic modelling (corresponding to occupants' behaviours) in building performance simulation studies. Since, so far until now, this was not addressed in the field of building performance modelling and simulation research and industry.

stochastic behaviours) uncertainties in building energy performance predictions. So far this demonstration has been constrained in scope, in terms of the number of design variables and the range of occupants' behaviours investigated as well as the complexity of the adopted use case. The scope of applicability of this new framework has been extended and demonstrated, in terms of the design variables and occupants' behaviours that are addressed, and applied to a more complex use case. The next chapter on Application II report on the outcomes of this extension.

Chapter 6

Application II - An extended study

In this chapter, the proposed framework has been extended and applied to a more complex use case represented by a multizone residential building model.

6.1 Case Study

The aim of this case study is to analyse the impact of thermophysical properties (e.g. conductivity, density, specific heat capacity) of materials used in building envelope components (wall, roof, ground floor and window) on the heating and cooling load predictions. The infiltration rates for the zones of multizone house are also considered to be uncertain with a goal to understand its impact along with the thermophysical properties of envelope materials. In addition to that, the stochastic phenomena relating to occupants' interactions with windows and shades are also enabled.

6.1.1 Experiment set up

A model of a complex multizone residential building located in Nottingham, UK ($51.15°N, 0.18°W$, elevation 62m) has been employed for which the corresponding EnergyPlus weather data file was used. This building consists of two occupied floors and an unoccupied attic having floor dimensions 7.4m×4.4m, and ceiling heights of 3.5m each and an attic floor with height of 1.45m (Figures 6.1, 6.2, 6.3). The wall, roof, ground and window construction compositions

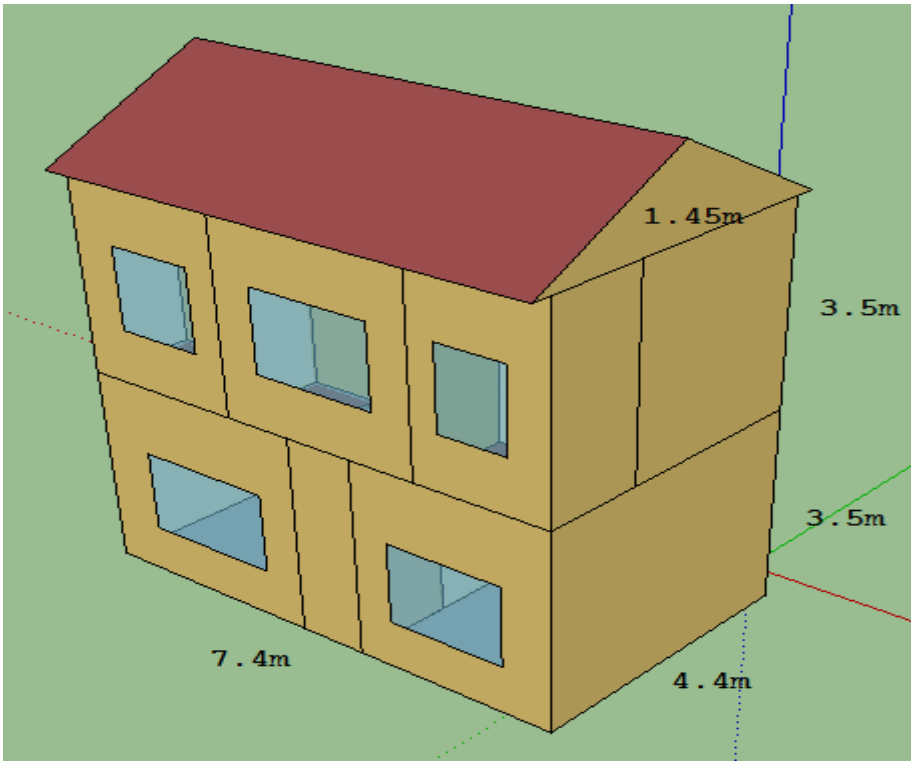


Figure 6.1: Building geometry of multizone house

used in this study are given in Table 6.1. The sensible internal heat gains for a couple residing in this house are taken from ISO 7730 [177] depending upon their activity (e.g. sleeping, passive, audio / visual, IT, etc.) and location (Bedroom, Living Room, Office, Kitchen, etc.) as specified in Chapman [2]. The heating (and cooling) and ventilation temperature set-point (SP) schedules are set to 21°C (and 25°C) and 24°C during occupation hours of respective building zones, according to ANSI/ASHRAE/IES Standard 90.1 for a residential prototype building model, representing an ideal loads HVAC system operation. A fraction 0.5 air changes per hour is assigned as a zone infiltration design flow rate for all the building zones.

6.1.2 Parameters studied

Fifty-nine design and physical properties input parameters (can be termed as controllable parameters from an analyst view point) have been considered to be uncertain in the S-BPS model with their assigned probability distributions given in Table 6.2. The prediction outputs are annual heating Y_h and cooling

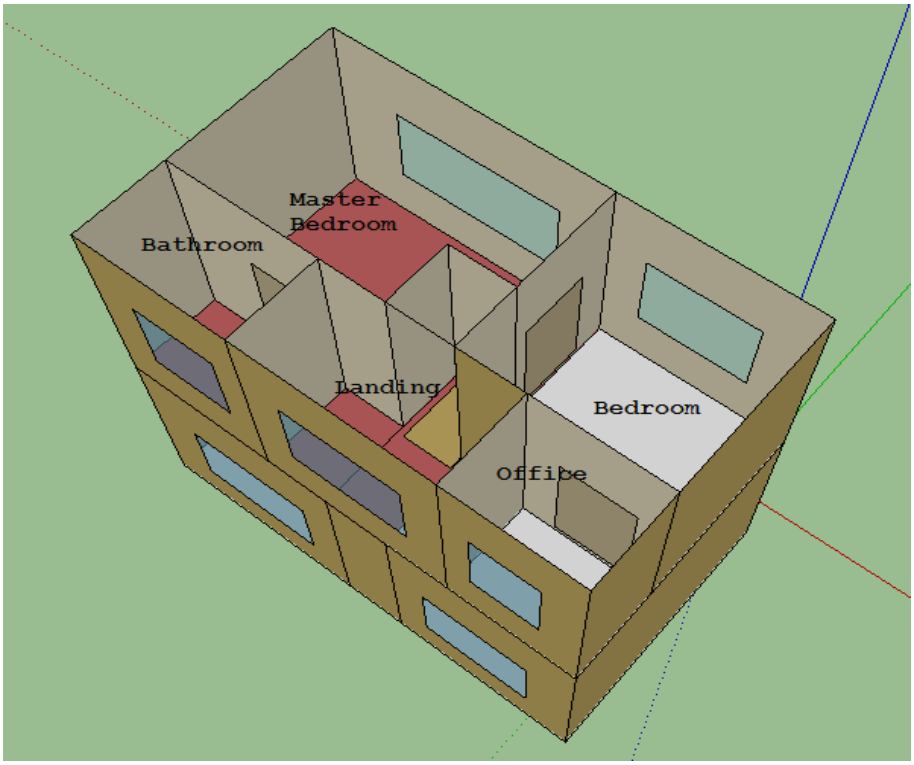


Figure 6.2: Building geometry of multizone house : First floor

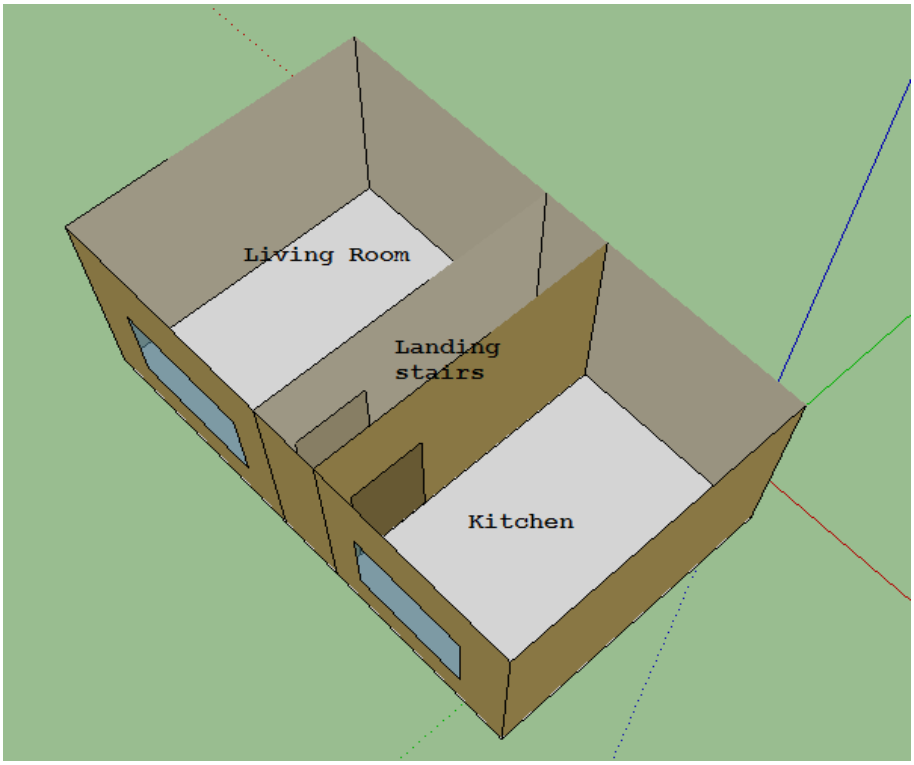


Figure 6.3: Building geometry of multizone house : Ground floor

Table 6.1: Thermophysical properties of construction layers : thickness, thermal conductivity, specific heat capacity and density

Surface	Material	t(m)	λ(W/(mK))	C(J/(kgK))	ρ(kg/m^3)
Wall	Brick	0.1	0.84	800	1700
	Insulation	0.0795	0.034	1400	35
	Concrete	0.1	0.51	1000	1400
	Plaster	0.013	0.4	1000	1000
Ground	Foam	0.1327	0.04	1400	10
	Concrete	0.1	1.13	1000	2000
	Screed	0.07	0.41	840	1200
	Flooring	0.03	0.14	1200	650
Roof	Clay tiling	0.025	1	800	2000
	No mass	Thermal	Resistance	= 0.15	
	Felt	0.005	0.19	837	960
Window	Glass	0.003	0.9	750	2500
	Air	0.013	0.0262	1005	1.17
	Glass	0.003	0.9	750	2500

Y_c demand per unit floor area. The objective here is to quantify the uncertainty in Y_h and Y_c given uncertainty in controllable input parameters as well as the uncertainty that arises from occupants' stochastic behaviour, due to uncontrollable parameter ω (from an analyst view point), encoded and inherent in the model described earlier. The parameters are assumed to be independent and characterized by normal and uniform distributions (Table 6.2). The assignment of particularly these type of distributions and the respective values for their parameters (e.g. mean, standard deviations, lower and upper bounds) is based on the measured material properties documented in Clarke [178] and processed for the uncertainty characterisation in Macdonald [105] [179]. As per these studies, a variability of 30% in thermal conductivity values and of $10-12\%$ in specific heat measurements and the other properties around their base case values (as noted in Table 6.1) have been assumed. The proposed EmUSA framework (sections 4.3-4.4) has been extended and a pre-step of Morris sampling and screening (subsection 3.3.1) is included, before the conduct of mean and variance GPs training, for the purpose of qualitatively singling out the most influential parameters out of initial fifty-nine parameters. Then, the proposed framework has been applied and the GP metamodels for Y_h and Y_c are constructed using Algorithm 1 and subsequently used within Algorithm 2 to conduct sensitivity analysis for each output variable.

Table 6.2: Thermophysical properties of construction layers : thickness t(m), thermal conductivity λ(W/(mK)), specific heat capacity C(J/(kgK)), density ρ(kg/m^3), thermal emittance ϵ and solar absorptance α. In brackets of Surface and Material columns, a number of parameters (in total 59) considered to be uncertain. "N" denotes Normal distribution with mean and standard deviation and "U" denotes Uniform distribution with lower and upper bounds.

Surface	Material	t	λ	C	ρ	ϵ	α
Wall(22)	Brick(6)	N(0.1,0.01)	N(0.84,0.261)	N(800,90)	N(1700,301)	U(0.85,0.95)	U(0.25,0.89)
	Insulation(4)	N(0.0795,0.008)	N(0.034,0.011)	N(1400,298)	N(35,5)	-	-
	Concrete(6)	N(0.1,0.01)	N(0.51,0.16)	N(1000,89)	N(1400,258)	U(0.85,0.95)	U(0.56,0.69)
	Plaster(6)	N(0.013,0.0013)	N(0.4,0.12)	N(1000,120)	N(1000,120)	U(0.85,0.95)	U(0.3,0.5)
Ground(22)	Foam(4)	N(0.1327,0.013)	N(0.04,0.012)	N(1400,168)	N(10,1.2)	-	-
	Concrete(6)	N(0.1,0.01)	N(1.13,0.339)	N(1000,120)	N(2000,240)	U(0.85,0.95)	U(0.6,0.8)
	Screed (6)	N(0.07,0.007)	N(0.41,0.123)	N(840,101)	N(1200,144)	U(0.9,0.95)	U(0.7,0.75)
	Flooring(6)	N(0.03,0.003)	N(0.14,0.042)	N(1200,144)	N(650,148)	U(0.88,0.92)	U(0.76,0.8)
Roof(12)	Clay tiling(6)	N(0.025,0.0025)	N(1,0.3)	N(800,96)	N(2000,600)	U(0.85,0.95)	U(0.6,0.82)
	No mass	Thermal	Resistance	= 0.15			
	Felt(6)	N(0.005,0.0005)	N(0.19,0.057)	N(837,100)	N(960,288)	U(0.88,0.92)	U(0.8,0.89)
Window(2)	Glass(2)	0.003	N(0.92,0.01)	750	2500	solar transmittance	U(0.74,0.84)
	Air	0.013	0.0262	1005	1.17		
	Glass	0.003	N(0.92,0.01)	750	2500	solar transmittance	U(0.74,0.84)
Infiltration	rate(1)	(air changes	per hour)	N(0.4,0.15)			

6.2 Results and Discussions

6.2.1 Parameter screening

Following Morris screening method described in subsection 3.3.1, a qualitative analysis has been conducted to identify the most influential parameters out of fifty-nine considered uncertain parameters. With a number of elementary effects (= 10) to be computed for each parameter and for the total of fifty-nine parameters, a total 600 runs (samples obtained according to Morris sampling) of BPS model of multizone house have been conducted. Morris screening method is applied on this input-output dataset and the μ-σ (mu-sigma) plots for the heating and cooling demands are obtained (Figures 6.4 and 6.5). A mu-sigma plot qualitatively compares the parameters on account of their means and standard deviations obtained from elementary effects distribution. A number of elementary effects (=10 assumed for this study) are actually a number of samples drawn from the elementary effect distribution for that parameter and these realisations when conducted determine the elementary effect distribution. Mean μ parameter accounts for the main or individual effect and sigma σ parameter for the interaction effect due to the impact of a combined variation in two or more interacting parameters on the model output. A high value for these parameters indicates higher influence on the output variable of interest. Table 6.3 describes the short annotations that are used for the parameters in these plots.

Morris measures (μ and σ) aid in understanding the model behaviour as well as to compute the qualitative ranking of influential parameters. Sanchez et al. [118] introduced an indicator $\frac{\sigma}{\mu}$ to detect the underlying model behaviour with respect to the parameters that causing the variations in model output. The three lines on the mu-sigma plots are the slope lines representing slope values (sigma/mu) less than 0.1, between 0.1 and 1 and greater than 1. A solid line (slope ≤ 0.1) represent a linear response of model output to its inputs i.e. the parameters lying in the region below this line exhibit linear impact. This is because the deviations in elementary effect are so small that all the effects are very close to their mean or are constant depicting linear response with respect to the parameter. The parameters in the the region between the solid line and dot-dash line (i.e. slope between 0.1 and 1) have monotonic to almost

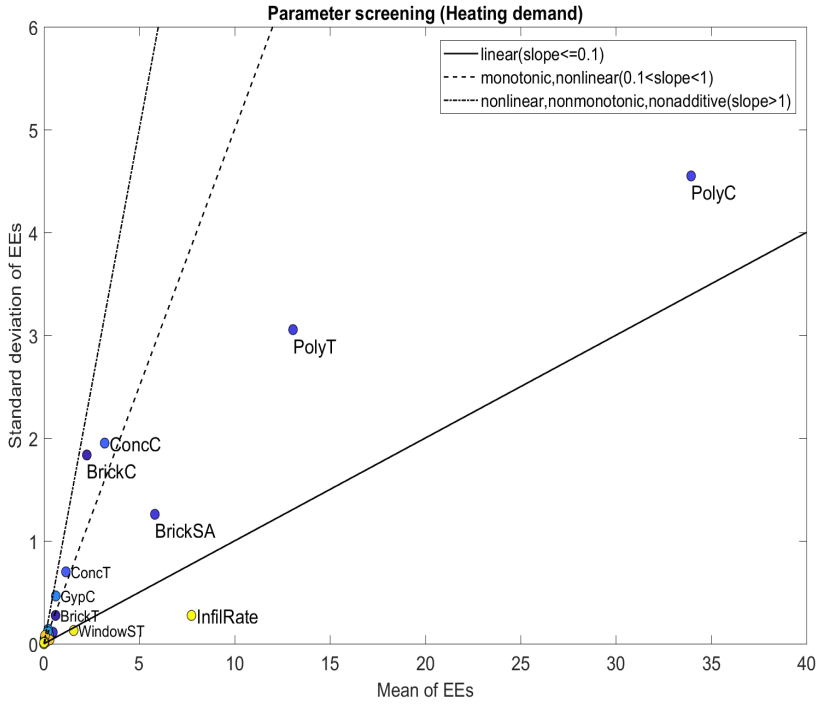


Figure 6.4: mu-sigma plot : Heating demand

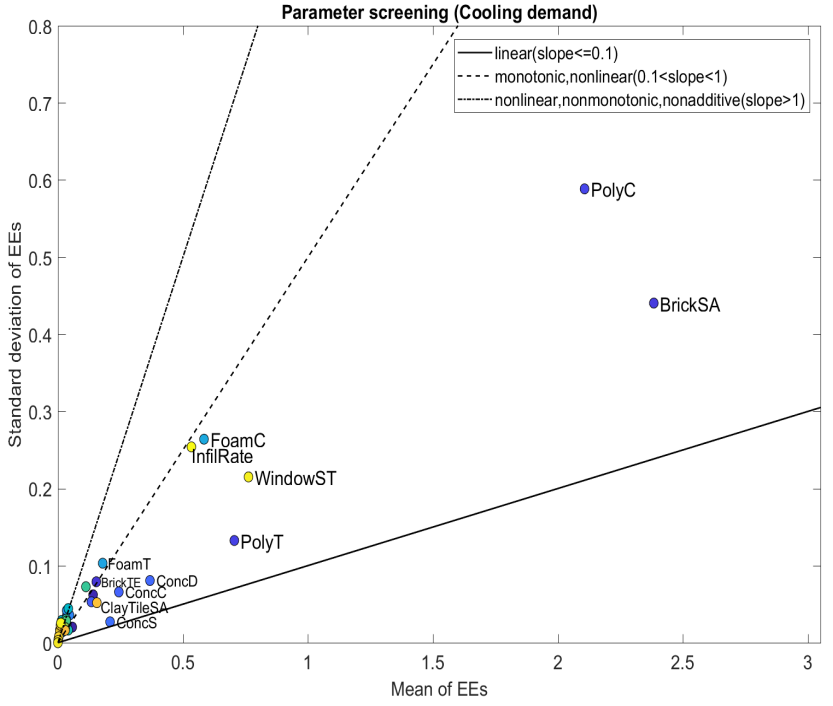


Figure 6.5: mu-sigma plot : Cooling demand

Table 6.3: The uncertain parameters' annotations and their descriptions

Annotation	Description
PolyC	Wall insulation (polystyrene) Conductivity
PolyT	Wall insulation Thickness
ConcC	Wall Concrete block Conductivity
BrickC	Brickwork Conductivity
BrickSA	Brickwork Solar Absorptance
InfilRate	Infiltration Rate
ConcT	Wall Concrete block Thickness
GypC	Gypsum plastering Conductivity
BrickT	Brickwork Thickness
WindowST	Window Solar Transmittance
FoamC	Ground insulation (urea formaldehyde foam) Conductivity
FoamT	Ground insulation (urea formaldehyde foam) Thickness
ConcD	Wall Concrete block Density
ConcS	Wall Concrete block Specific heat capacity
ClayTileSA	Clay Tile roofing Solar Absorptance
BrickTE	Brickwork Thermal Emittance

monotonic and nonlinear response. Because there are low to high deviations in effects (+ve/-ve) resulting into monotonically increasing or decreasing effect of parameter on the model output. The parameters in the region represented by the line with slope more than one exhibit most diverse behaviour of non-linearity, non-monotonicity and non-additivity on the model output, as the elementary effects deviate highly from their mean.

In case of heating demand, the parameters wall insulation conductivity PolyC and thickness PolyT (having high values for both mean and standard deviations) have highest impact on the model outcomes. Qualitatively, they tend to have monotonic and nonlinear impact on the heating demand. Conversely, the window transmittance WindowST and inflitration rate InfilRate having lower and linear response to the heating load predictions than PolyC and PolyT. On the other hand, brickwork and concrete conductivity having more nonlinear impact on the output. In case of cooling demand, again the wall insulation conductivity PolyC and a new parameter brickwork solar absorptance BrickSA dominate the cooling load predictions. The latter is due to the increased solar gains in summer season as uncertainty in solar absorptance is crucial and affects predictions when cooling demands are high due to resulting variations in longwave radiations. The other parameters ground insulation

conductivity FoamC, InfilRate, WindowST and PolyT indicating monotonic and nonlinear trend due to their respective uncertainties. From this qualitative analysis and plots, a total of eight most influential and common (to both heating and cooling demand output variables) input variables have been selected for further extensive quantification and decomposition of uncertainties in output variables. The impact of parameters BrickC, BrickSA, PolyT, PolyC, ConC, FoamC, WindowST and InfilRate is further investigated to quantitatively identify their dominating effects on the load predictions (subsection 6.2.3).

6.2.2 Gaussian Process Emulation of the S-BPS outputs

For this extended case study, following the operational procedure (as described in section 5.2.1) for training of GPs, the values of 0.9876 and 0.7027 were obtained as predictivity coefficient for mean and variance GP of heating demand at 50×1000 training combination of design points vs S-BPS repetitions. For cooling demand, these values stand at 0.9909 and 0.9573 respectively. According to Eq. 4.10, for this pair of GPs trained on multizone house dataset, the proportion of test observations that lie inside the 95% CI predicted by the GP emulators are 0.96 (mean), 0.99 (variance) and 0.88 (mean), 0.99 (variance) for heating and cooling demands respectively. This plausible pair of GPs is further employed for predicting mean and variance of energy demands Y_h and Y_c.

The GP regression curves for the computed energy demands' mean with respect to each of eight parameters are shown in Figures 6.6 and 6.7 together with the testing dataset, plotted for the convenience by varying only one parameter at a time but employing the GP trained on eight dimensional input space. Each GP curve is an interpolation curve that predicts mean energy demand value (black circles) and variance (red bars in Figures 6.6 and 6.7) around due to the uncertainty in emulation itself.

The regression curves allow us to understand the individual or main effect of input variables on the load predictions. In decreasing order of their individual effects: 1) the uncertainty in wall insulation thickness (PolyT) and conductivity (PolyC), complying with the underlying heat conduction physics (with heating load decreasing and increasing as insulation thickness and conductiv-

ity increases respectively), results in a variation of approximately 50 to 35 and 25 to 55 kWh/m^2 respectively, 2) the infiltration rate InfilRate and wall brickwork solar absorptance BrickSA show relatively moderate effect, both leading to a variation between 37 and 45 kWh/m^2, although in reverse order. This is in accordance with the underlying building physics resulting into the heat losses due to convection for InfilRate and heat gains due to increased short and longwave radiations for BrickSA inside the building envelope. 3) the wall brickwork BrickC and concrete block ConcC conductivity show partly nonlinear effect and positive trend in heating load also with a moderate variation of 38 to 42 kWh/m^2, 4) the window solar transmittance WindowST being relatively less influential to the load predictions and 5) the ground insulation conductivity FoamC having negligible impact. These findings also verify the qualitative analysis reported in previous section 6.2.1 with PolyC and PolyT having highest impact relative to InfilRate and BrickSA; BrickC and ConcC showing non-linear behaviour; WindowST being the lowest and FoamC with near to zero value for the main effect (mean of respective Elementary effects).

In cooling season, PolyC and BrickSA showing the highest impact on cooling load predictions but interacting against each other resulting into the variation between 2 to 8 kWh/m^2. This is again complying the underlying physics with BrickSA raising the cooling demand due to increased heat gains and PolyC tending to lower it down by loosing heat through the walls simply because of increased in conductivity. The next four uncertain parameters PolyT, WindowST, FoamC and InfilRate driving moderate variations between 4 to 5.5 kWh/m^2 in cooling loads with PolyT and WindowST driving the positive change while FoamC and InfilRate on the negative side. The brickwork BrickC and concrete block ConcC conductivity having lowest but partly nonlinear impact on the cooling demand. These again confirm the qualitative analysis results obtained in previous section 6.2.1.

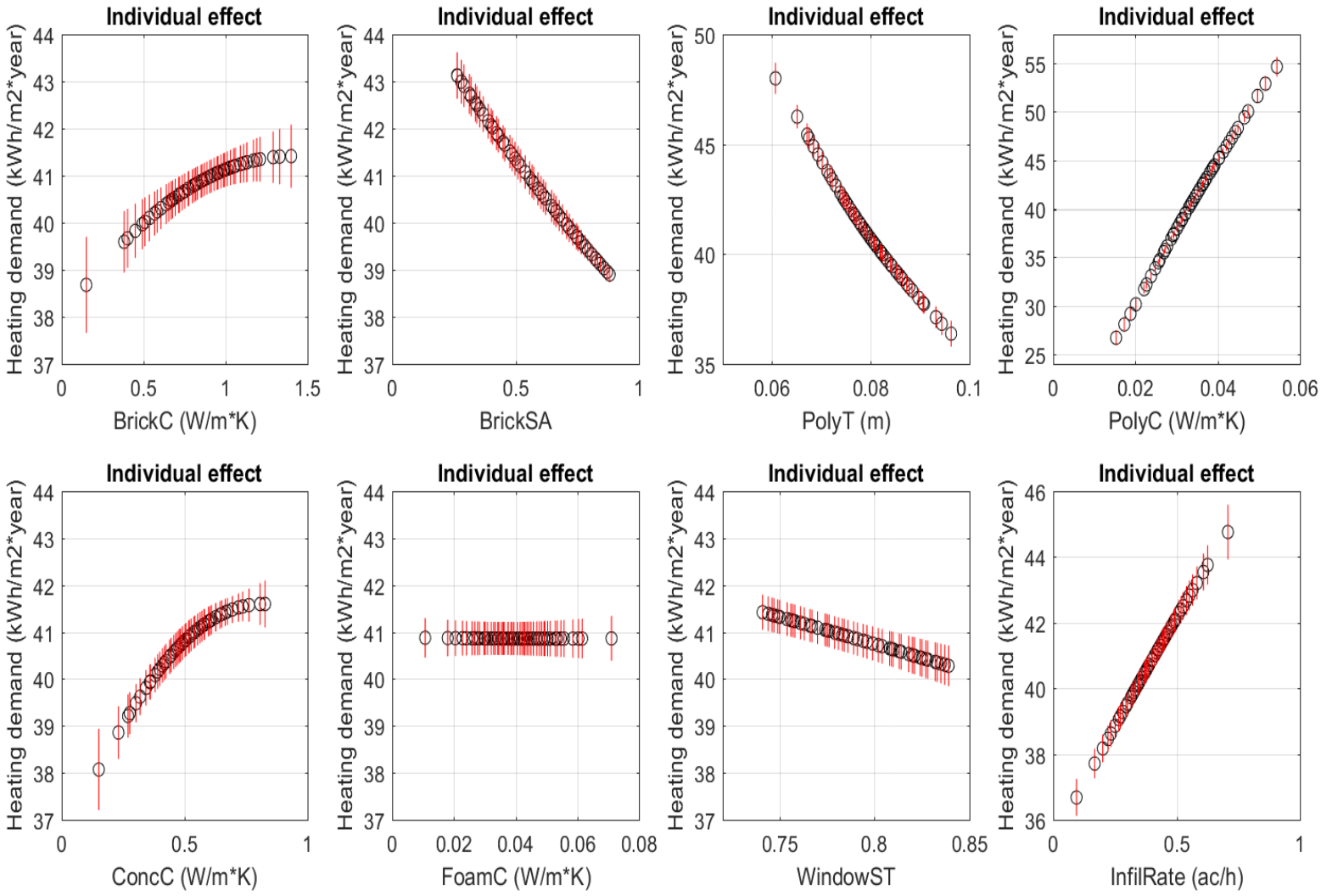


Figure 6.6: Main / Individual effect plots : Heating demand

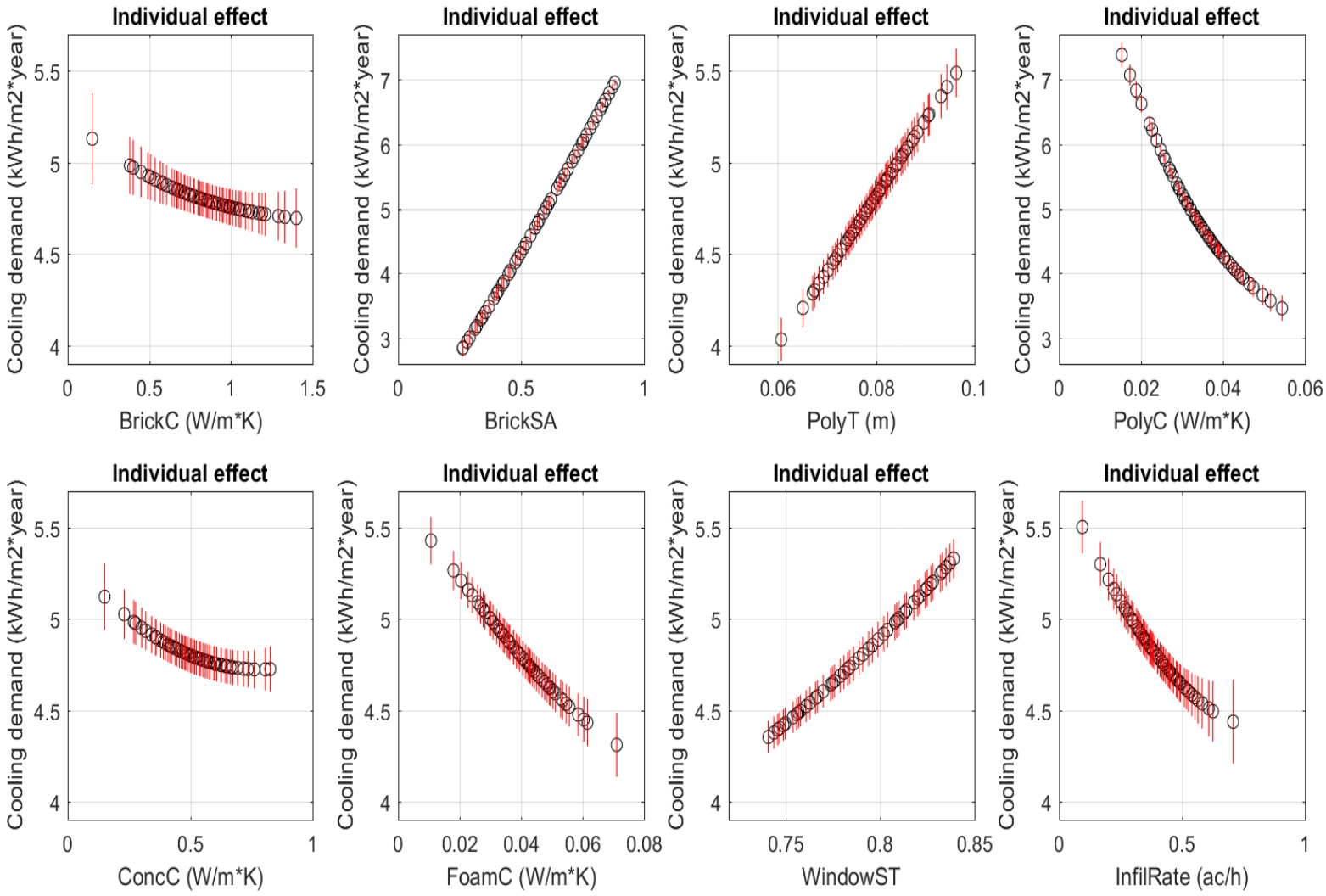


Figure 6.7: Main / Individual effect plots : Cooling demand

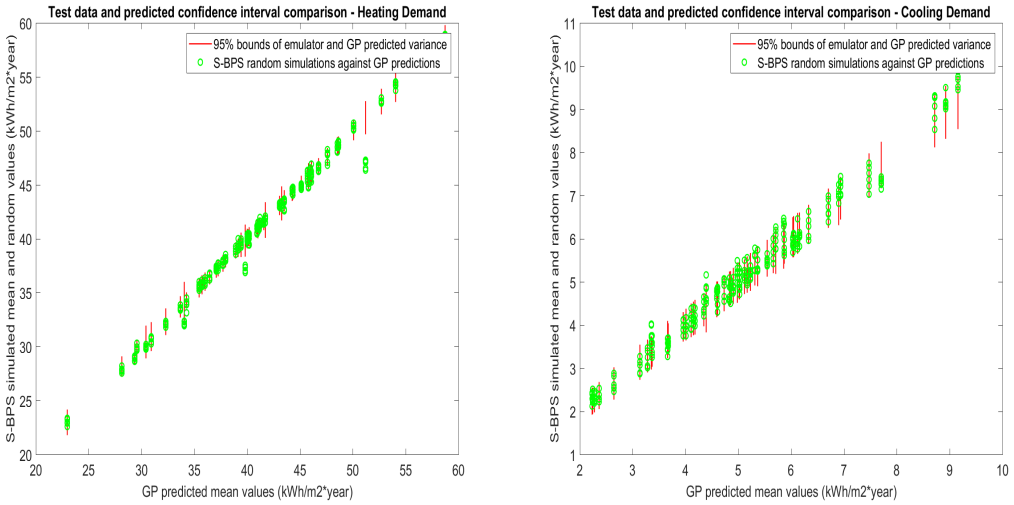


Figure 6.8: Comparison of S-BPS simulated random values within GP predicted credible intervals at test data for heating (a) and cooling (b) demand

6.2.3 Uncertainty Quantification using the GP emulator

The capability of the trained GP emulator pair to statistically reproduce uncertainty in the outputs of the S-BPS within the same input-space is assessed by considering the credible intervals introduced in (4.27) for $\alpha = 0.05$ (i.e. 95% CI). A test dataset with 1000 replicates of the S-BPS at a collection of 50 randomly selected input points is generated. Some 94.6% and 93.1% of the simulator's random response values [at each test point] for heating and cooling demands lie within the 95% CI predicted by the pair of (mean and variance) GPs (Figure 6.8), suggesting that this new GP emulator trained on extended study dataset can be used with confidence.

In Figures 6.9-6.10, the probability density functions (PDFs) are compared, at some random test points, computed by Monte Carlo Simulation, the S-BPS outputs (in red) and outputs from the GP emulator (in blue); this latter by simulating stochastic observations with (4.26) using GP emulators for $f_d(\mathbf{X})$ and $V(\mathbf{X})$ with regard to this extended case study. In Figure 6.11 the cumulative distribution functions (CDFs) of predicted energy demands from the GP emulator are compared with those of the S-BPS simulations, again obtained using Monte Carlo Simulation. A single CDF in these plots represent variation in prediction due to occupants' stochastic behaviour [aleatory uncertainty] and their spread represents uncertainty in predictions due to input

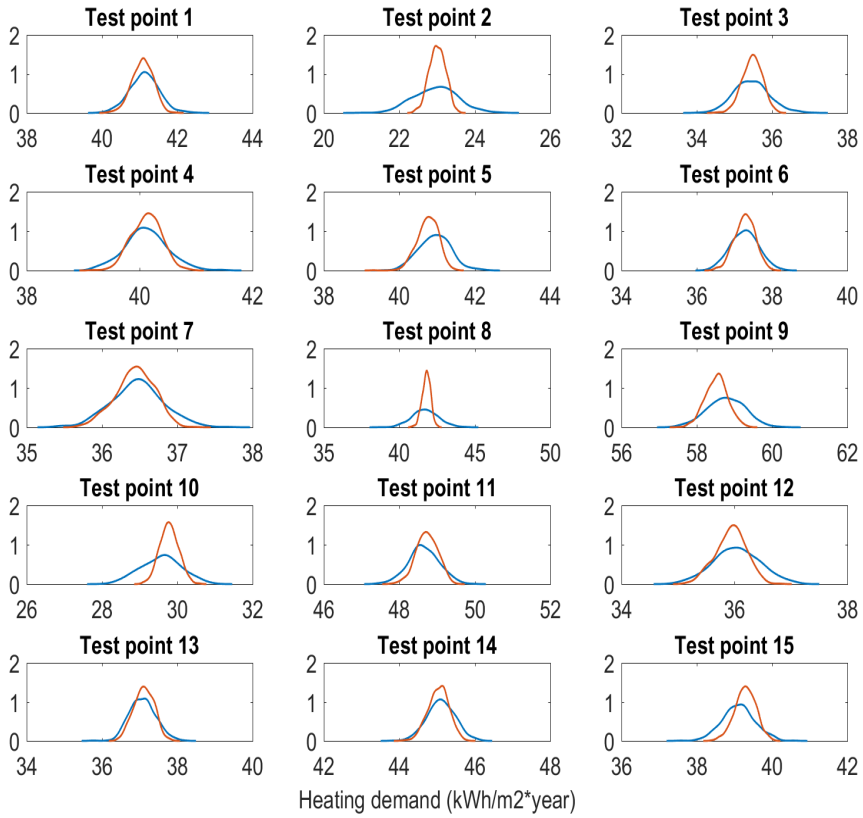


Figure 6.9: Uncertainty propagation comparison between GP predicted (blue) and S-BPS simulated PDFs (red) for heating demand

parameters [epistemic uncertainty]. Figures 6.9-6.11 offer numerical evidence that the proposed framework enables us to confidently reproduce the underlying PDFs and CDFs obtained with the S-BPS, at a small fraction of the computational cost, since once the GP emulators have been constructed, S-BPS runs are no longer required for subsequent uncertainty analyses.

In both the cases of heating and cooling demand (Figure 6.11), the collective spread of CDFs (35 and 7 kWh/m^2 difference in magnitude for heating and cooling respectively) is more than the spread of an individual one (2 and 1.5 kWh/m^2 difference in magnitude for heating and cooling respectively) - illustrating that epistemic uncertainty due to design and physical properties inputs plays a more dominant role than the aleatory uncertainty due to occupants' behaviours. This is expected in both the cases, given the nature of multizone house case study enclosing the large heated volume where the total

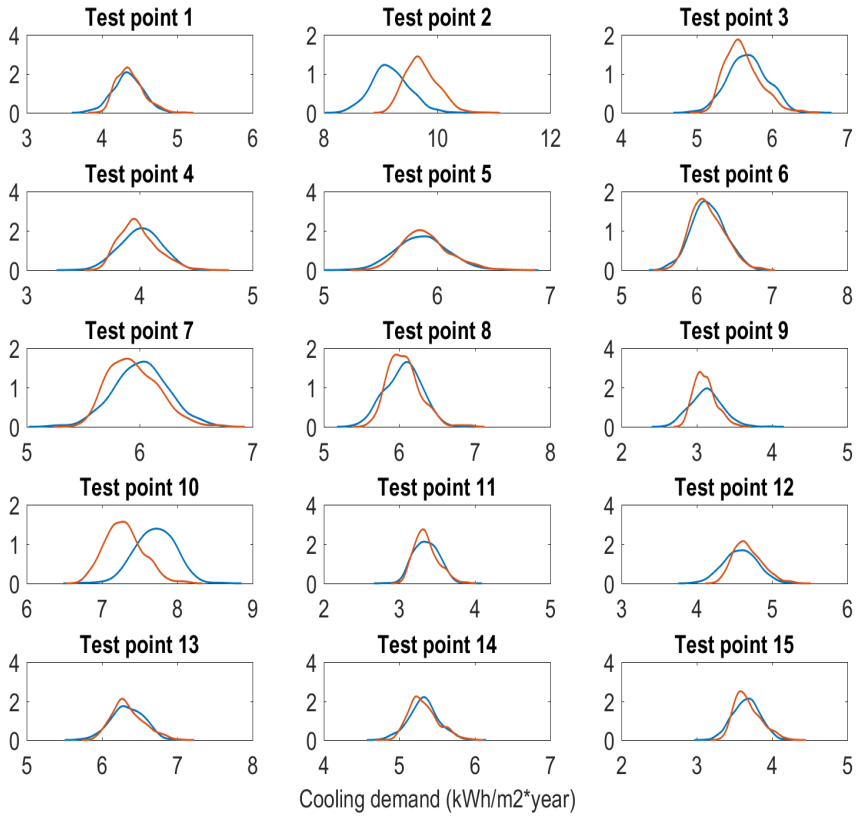


Figure 6.10: Uncertainty propagation comparison between GP predicted (blue) and S-BPS simulated PDFs (red) for cooling demand

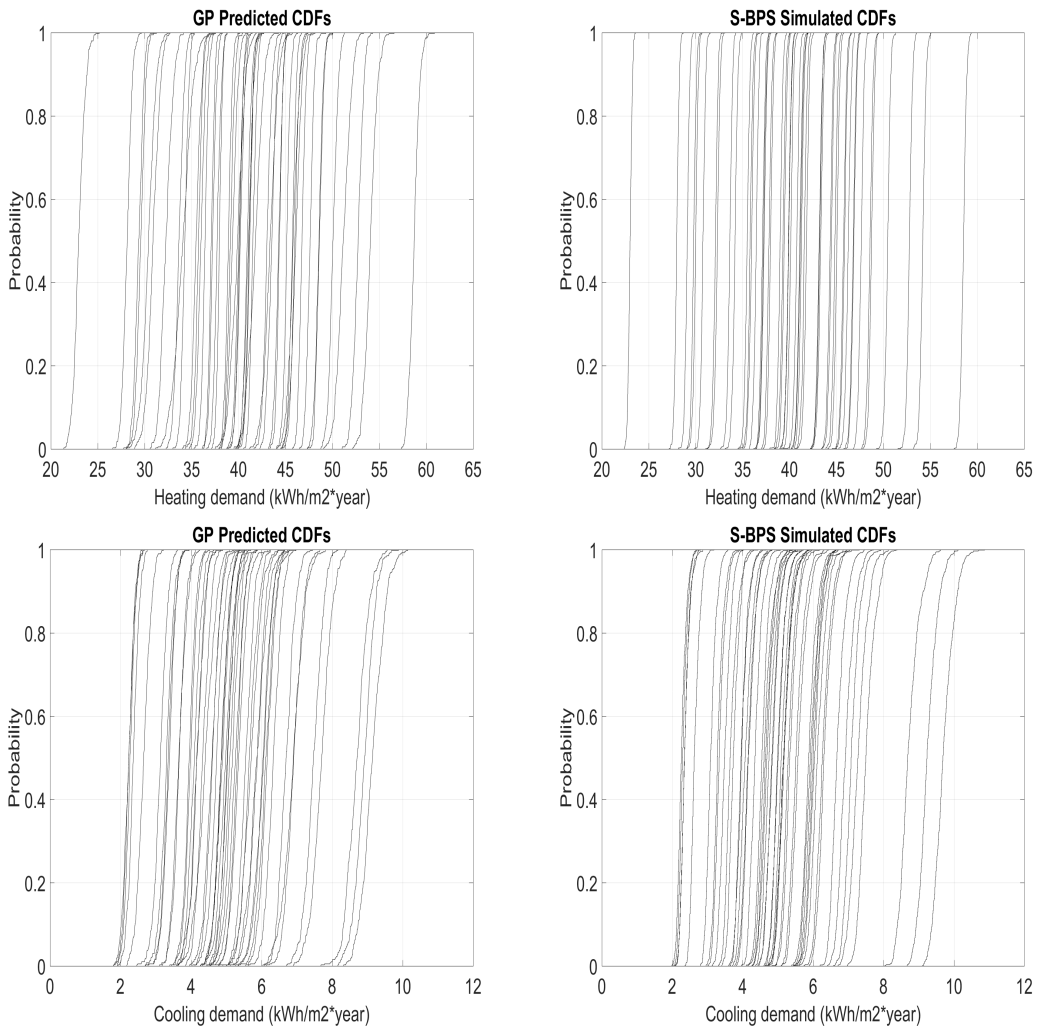


Figure 6.11: Comparison of GP predicted (left) and S-BPS simulated (right) CDFs for heating (top) and cooling (bottom) demand

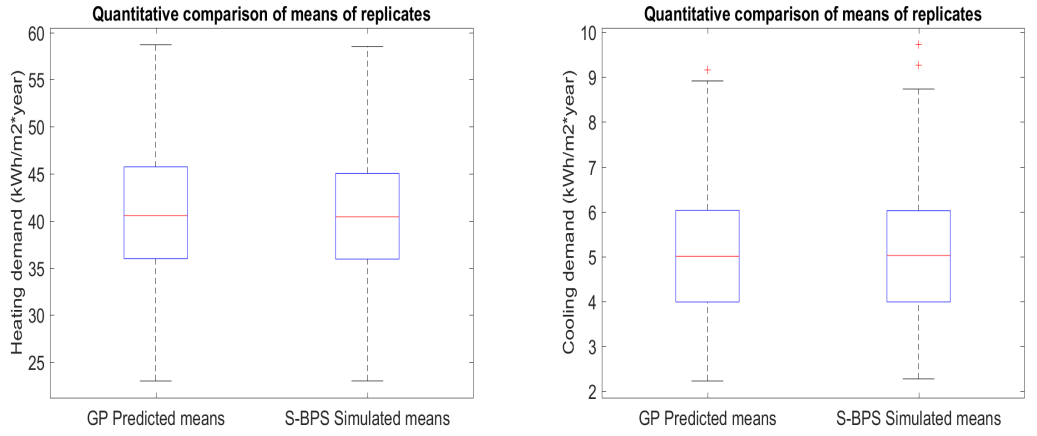


Figure 6.12: Comparison between S-BPS Simulated and GP predicted means of heating (left) and cooling (right) demand replicates in Figures 6.11

heat balance is dominated by internal heat storage and exterior heat transfer components than the stochastic internal gains due to a couple residing in the house. Finally, the quantitative comparison between the means of heating and cooling demand replicates presented in Figure 6.12 underline how accurate the GP emulator's predictions are. Having demonstrated that the GP emulator is an acceptably accurate means for accelerating stochastic building performance simulations, we proceed now to decompose the total uncertainty in predictions into its component parts by computing sensitivity indices for our extended case study.

6.2.4 Uncertainty decomposition using the GP emulator : Sensitivity analysis

The total uncertainty in energy demand prediction is decomposed using the individual and total effect indices of (4.47) and (4.48). The individual effect indices account for the impact of input variables and the total effect index includes that of occupants' behaviour. The index scale is between 0 and 1, with highly influential factors having values close to unity. Table 6.4 reports these indices, computed for both small (500) and large (20000) input sample sizes.

In the case of heating demand, and as discussed earlier, epistemic uncertainty due to wall insulation thickness PolyT and conductivity PolyC dominates the uncertainty in predictions, at the individual index values of 0.0771 and 0.8428 respectively. During heating season envelope heat losses (and thus impact of uncertainty in insulation thickness and conductivity) dominates the energy balance, much more so than occupants' behaviour, since occupants are relatively less likely to open windows (to avoid excessive heat losses) and to lower shades (thus reducing useful transmitted solar heat gains) during the heating season.

For cooling demands, the parameter BrickSA related to heat radiation components is dominant than aleatory uncertainty due to occupants' behaviour, as heat radiations from numerous exterior (and interior walls) components (for this multizone house) are more likely to be dominant than the occupants' interaction with the building envelope, as explained above in the previous section 6.2.3. Again, these quantitative analysis results highly corroborate with

both the qualitative analysis results (Figures 6.4 and 6.5) and the underlying emulation results (Figures 6.6 and 6.7), all indicating PolyC and PolyT being the most influential parameters for heating demand and PolyC and BrickSA for cooling demand predictions.

Table 6.4: Sensitivity indices quantifying the impacts of input parameters and occupants' behaviour on the heating and cooling energy demand predictions. S_1, S_2, S_3, S_4, S_5, S_6, S_7 and S_8 are the individual effects due to BrickC, BrickSA, PolyT, PolyC, ConcC, FoamC, WindowST and InfilRate respectively and S_{T_w} total effect due to occupants' behaviours

Output variable	Sample size	Sensitivity indices								
		S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_{T_w}
Heating demand	Small (500)	0.1353	0.1527	**0.1466**	**0.7158**	0.0739	0.0728	0.0790	0.1096	**0.0010**
	Large (20000)	-0.0299	-0.0172	**0.0771**	**0.8428**	-0.0257	-0.0301	-0.0290	0.0046	**0.0010**
Cooling demand	Small (500)	-0.2860	**0.2232**	-0.2672	**0.3099**	-0.2789	-0.2863	-0.2964	-0.2351	**0.0130**
	Large (20000)	-0.0083	**0.5032**	0.0065	**0.3714**	-0.0073	0.0035	**0.0162**	**0.0129**	**0.0130**

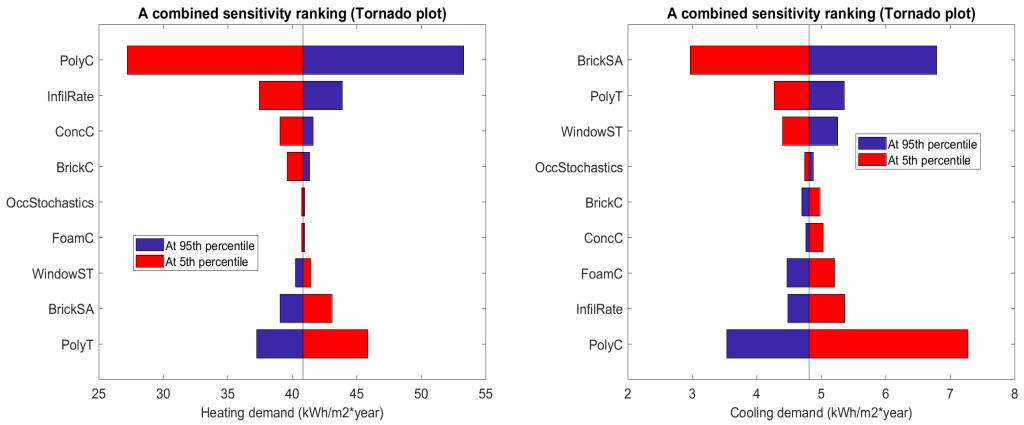


Figure 6.13: Sensitivity Tornado plots for heating (a) and cooling (b) demand

Figure 6.13 illustrates the combined sensitivity analysis of energy demands as a tornado plot. The tornado plots further reinforce the parameter screening (subsection 6.2.2) and sensitivity analysis (subsection 6.2.4; table 6.4) results showing that the occupancy stochastics, brick (BrickC) and concrete (ConcC) conductivities leading to negligible variation in energy demands. For both output variables, the input variables insulation conductivity (PolyC), infiltration rate (Infilrate), brick solar absorptance (BrickSA) and insulation thickness (PolyT) are the common variables that result into highest possible variation in energy demands. In fact, the tornado sensitivity analysis results are also consistent in their ranking of importance (with the ones in Figures 6.4 and 6.5) and as well as on how the input variables correlate with the output variable (Figures 6.6 and 6.7). The heating demand is positively correlated with the insulation conductivity and infiltration rate (5^{th} percentile falling left of median) and negatively with the brick solar absorptance and insulation thickness (5^{th} percentile falling right of median) (Plot a) of Figure 6.13). The polarity of correlation for these same input variables is reverse for cooling demand (Plot b) of Figure 6.13). The foam conductivity (FoamC) and window transmittance (WindowST) exhibiting moderately negative and positive respective impact on the cooling demand while negligible on the heating demand.

Figures 6.14 and 6.15 (top rows) show the convergence of indices values for the emulator evaluations at different sample sizes, ranging from 500 to 20,000. We consider 0.05 as the convergence threshold for the width of the 95% credible interval $CI_{indices}$, below or close to which the index value is deemed to

Table 6.5: Comparison of computational runs between Classical and EmUSA framework for the extended study

	Classical	Emulator based	
		Training	Sensitivity Analysis
Simulator runs	20×10^4	5×10^4	-
Emulator runs	-	-	0.05 to 2×10^4

be stable. In the case of heating demand (Figure 6.14), for the main effect index of wall insulation conductivity PolyC, the index value converges rapidly (at a small sample size of 3,000) near the threshold, while the other indices values approach the threshold (at values of around 0.2) at large sample sizes of 20,000. The bottom rows of the figures show the convergence of the actual credible intervals around indices values obtained via bootstrapping, demonstrating the trend for convergence at large sample sizes. For cooling demand, the significance of performing convergence study for sensitivity indices computation is prominent, as there is a change in dominance from wall insulation conductivity PolyC to brickwork solar absorptance BrickSA when the indices are computed at large sample sizes. From upper row in Figure 6.15, the convergence for all the parameters has been achieved quite rapidly at sample size of 5000. And from lower row in Figure 6.15, the change in dominance from PolyC to BrickSA is visible and stable at the sample size 6000 and onwards.

For the present extended case study example with 8 input parameters and 20,000 samples (large sample sizes at which indices converge as shown), the number of simulator runs that we need in our proposed Emulation based Uncertainty and Sensitivity Analysis (EmUSA) framework is a fraction (50,000 / 200,000 = 0.25) of that required by the standard approach (see Table 6.5), as the emulation framework only requires simulator runs to train GPs, which can then be employed to conduct sensitivity analyses at a low computational cost. Consequently, the HPC machine computation time reduces to a manageable c.1 day, in contrast to the more than half a weeks' time that is required by a classical approach (see Table 6.6). In this way, it is possible to complete an entire UQ study for the extended case of multizone house using our proposed EmUSA framework within a single day.

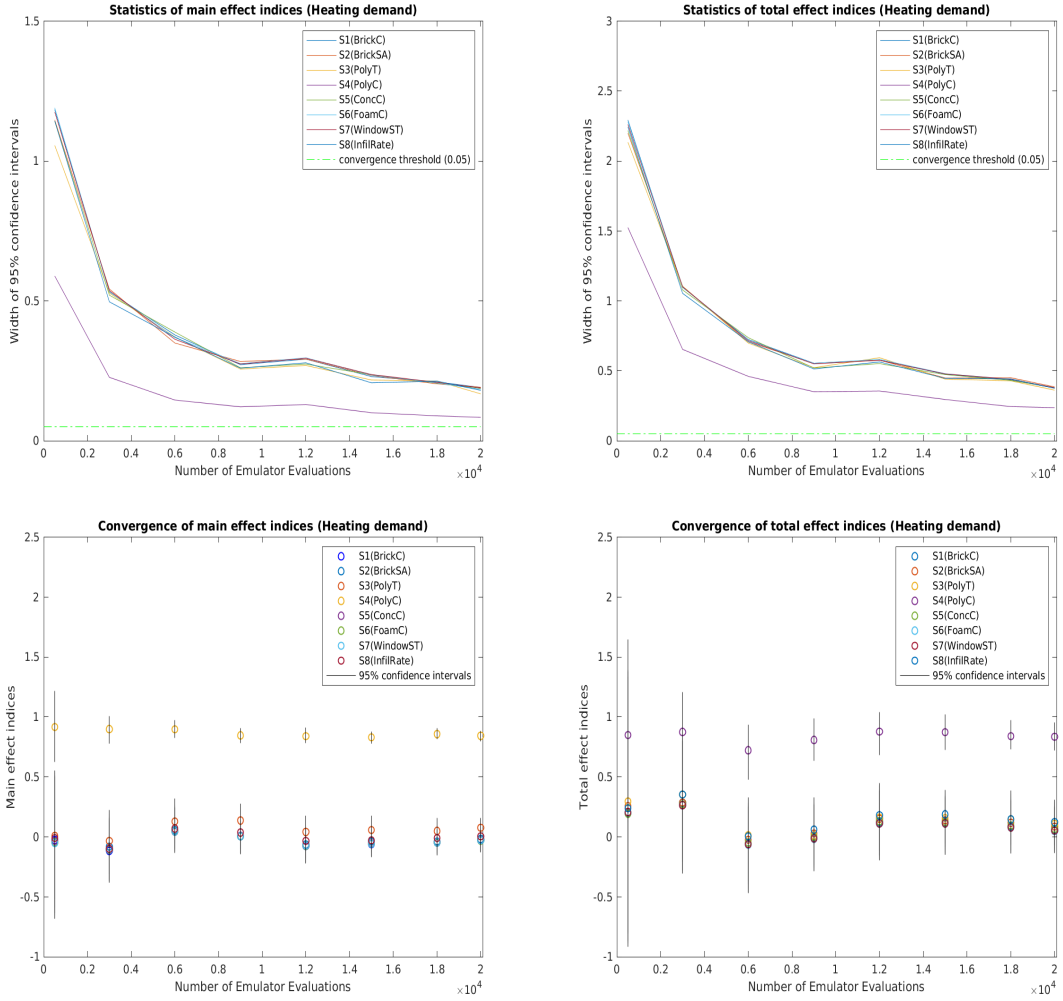


Figure 6.14: Convergence of main (left) and total (right) effect sensitivity indices of design and physical properties input variables for heating demand at distinct sample sets of increasing size. The top row shows the convergence of the width of index credible interval and bottom row shows the actual credible interval (Eq. 4.49).

Table 6.6: Comparison of computational expense on HPC between Classical and EmUSA framework for the extended study (One simulation run benchmark = 90 seconds)

	Classical	Emulator based	
		Training	Sensitivity Analysis
Simulator runs	4.17 days	1.04 days	-
Emulator runs	-	-	15 to 20 minutes

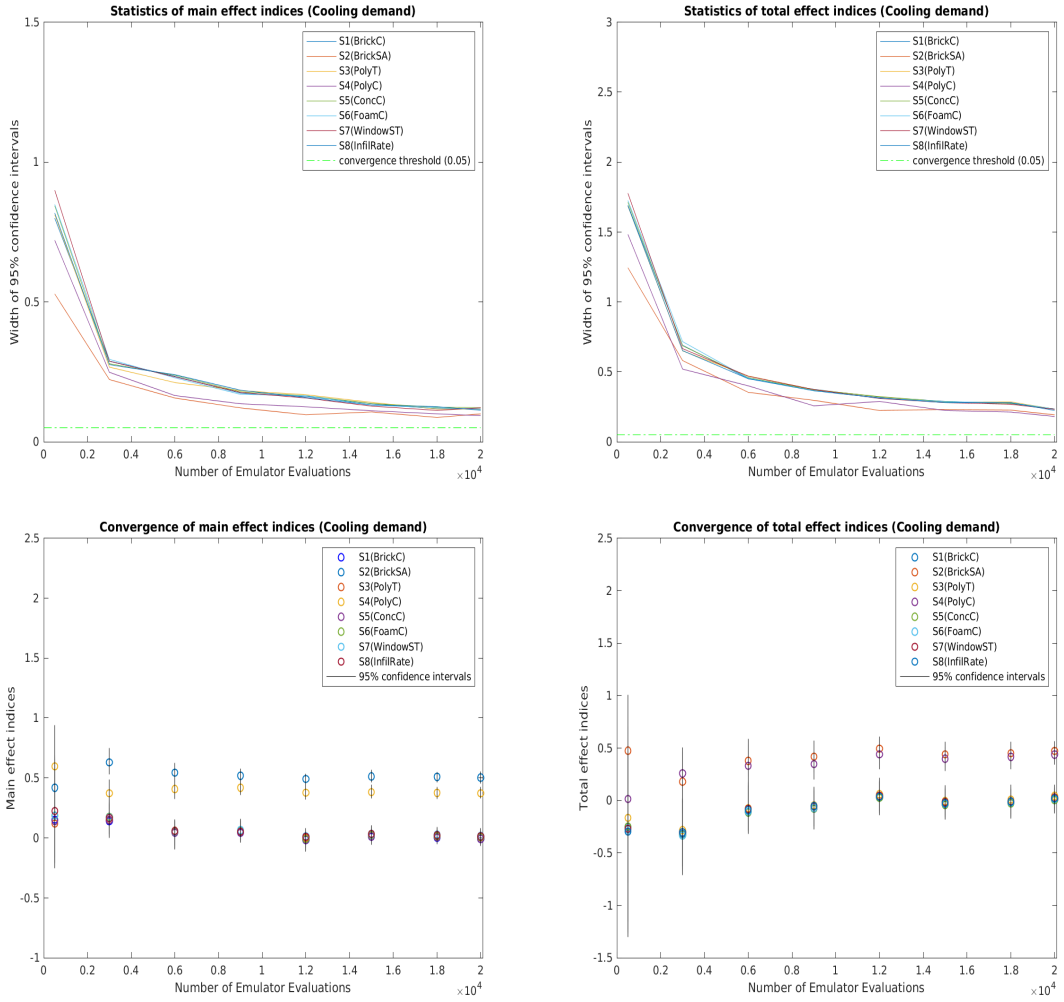


Figure 6.15: Convergence of main (left) and total (right) effect sensitivity indices of design and physical properties input variables for cooling demand at distinct sample sets of increasing size. The top row shows the convergence of the width of index credible interval and bottom row shows the actual credible interval (Eq. 4.49).

6.2.5 Computational savings in Sensitvity Analysis under EmUSA framework

The proof-of-principle (2-input model) and an extended study (59-input model, effectively only 8 most sensitive parameters after screening for an extensive UQ) have reported 96% and 75% savings, when compared to classical UQ framework [not employing a metamodel], respectively in the computation cost of performing sensitivity analysis using proposed EmUSA framework. The proof-of-principle study is focussed only on the modulation of heat conduction and shortwave irradiance due to the respective uncertainties in insulation thickness PolyT and window transmittance WindowST along with the stochastic occupants' behaviour with window, shade and lighting enabled. The extended study, on the other hand, enables modulation in convective and longwave radiative phenomena by a way of accounting for the uncertainties in infiltration rate InfilRate and brick solar absorptance BrickSA in addition to the heat conduction modulation due to uncertainties in envelope material conductivities.

Given the systematic and sequential enabling of increased complexities of underlying BPS model during these studies, the computational savings using proposed EmUSA framework falls by a little more than twenty percentage points for a factor of four increase in dimensionality of input parameter space from that of proof-of-principle study benchmark (i.e. 2-input case). This outcome on scaling of computational savings depending upon the model complexities may allow an uncertainty analyst to gauge the computational load required when the model complexity and / or dimensionality of input space varies on the lines of benchmarks assessed from these studies. The nature of scaling of computational savings, for instance a linear (e.g. forty percentage points fall) or an exponential (e.g. eighty percentage points fall) for 16 inputs (a factor of eight increase in dimensionality) case or the any other, is still a potential topic for further investigation when using proposed EmUSA framework.

6.3 Conclusions

The purpose of this extended study was to demonstrate the applicability and computational efficiency of EmUSA framework to support the quantification

and decomposition of both epistemic and aleatory uncertainties for the models of increasing complexity. From these results we conclude that:

- The pair of fitted GP emulators for this extended study successfully emulate the behaviour of the stochastic building performance simulator (S-BPS) at the chosen test design points, with predictivity coefficient $Q_2 > 0.95$, (falls short by small for heating demand variance GP).

- The proposed GP Emulator-based UQ framework, requiring simulator runs only to fit the emulators, is moderately computationally efficient (for this extended case); reducing the computational cost to a low fraction (0.25) of its classical counterpart, as simulator runs are required only for training the emulator which subsequently computes at a low cost.

- The effects of uncertainties in wall insulation conductivity PolyC and thickness PolyT on heating demand are more significant than those arising from stochasticity in occupants' behaviours; as these behaviours are relatively constrained during the heating season.

- In case of cooling demand, the effects of uncertainties arising from brickwork solar absorptance BrickSA and wall insulation conductivity PolyC dominate the stochasticity in occupants' behaviours, as the larger heated volume enclosed by the multizone building involving numerous heat storage and transfer components; resulting into acceleration of short- and long-wave radiation phenomena which dampen the occupants' interaction to restore their thermal comfort, leading to negligible or low impacts on cooling demand.

Chapter 7

Conclusions, Future works and Recommendations

This thesis attempts to make a contribution to the knowledge in an interdisciplinary field of Building Simulation and Uncertainty Quantification in terms of providing efficient means to effectively account for the uncertainties that lead to the *performance gap* between the simulated building performance and its measurement. To this end, this thesis develops a strategy to deal with the uncertainties arising from both epistemic and aleatory sources in the building energy performance predictions. Particularly, this thesis has attempted to answer following questions that specifically hinder a straightforward conduct of Uncertainty and Sensitivity Analysis for complex building energy performance simulation.

7.1 Answers to the Research Questions

***Q*. Are the existing metamodelling / emulation techniques sufficient enough to abstract the building energy performance phenomena and to accelerate the computation of uncertainties in energy predictions?** Now, there is a paradigm shift from deterministic to stochastic simulation of building energy performance with the advent of coupling stochastic occupants' behaviour [1] to the deterministic heat storage and transfer phenomena result-

[1] for instance, can also be related to the ever changing climate resulting into the uncertain

ing into a increasingly complex system of thermal energy flows in buildings. Increase in complexity, also comes at the cost of high computation expense. Since the computation of uncertainties involve Monte Carlo simulations of underlying computational S-BPS model, this requires the metamodel / emulator to be parsimoniously and plausibly structured but at the same time fast enough to accelerate the computation of uncertainties. **Objective I** addressed in Chapter 4 section 4.1 first establishes a selection strategy for the metamodelling technique that would be suiting the emulation needs of abstracting stochastic building performance. After evaluating these techniques against the laid out criteria, Gaussian Process Regression (GPR) stands out fulfilling the requirements of plausible abstraction and at the same time allowing the acceleration in quantification and decomposition of uncertainties in load predictions.

***Q*. How good is then the GPR to emulate the deterministic and stochastic heat transfer phenomena occurring inside the buildings?** GPR methodology allows to emulate the two superimposed phenomena by a way of distinctly representing the mean and variance component of the stochastic building performance response. The modelling procedure is flexible enough to configure occupants' variability into the emulator constructing procedure. Moreover, unlike other techniques, GPR provides uncertainty bounds due to the emulation procedure itself. **Objective II** addressed in Chapter 4 section 4.3 presents the developed mathematical framework behind the emulation of S-BPS response. The pair of metamodel is then validated (for its parsimony and plausibility) using the test simulation datasets by employing a predictivity coefficient validation measure (section 4.2.1). The corresponding computational framework (section 4.5) and methodology workflow (section 4.6) present the operating procedure of Gaussian Process Emulation based Uncertainty and Sensitivity Analysis framework, demonstrating its readiness for the future extensions.

***Q*. How good is then the proposed EmUSA framework in propagation of epistemic and aleatory uncertainties and decomposition of**

weather

the total uncertainty in energy predictions?

Using a relatively simple use case of a small monozone office building with 2-uncertain input parameters setting but enabling stochastic phenomena related to occupants' behaviours, the pair of fitted GP emulators is successful enough in emulating and capturing the behaviour of the stochastic building performance simulator with predictivity coefficient (> 0.95), (except for a small number of moderate exceptions). The framework is efficient in reducing the computational expense of Monte Carlo propagation and quantification of epistemic and aleatory uncertainties to a small fraction (0.0375) of its classical counterpart for this monozone office case. **Objective III** addressed in Chapter 5 using variance based decomposition technique (developed in section 4.4 for stochastic case) demonstrates with this proof-of-concept the decomposition of the total uncertainty in heating and cooling energy load predictions.

***Q*. How about preparedness and practical applicability of EmUSA framework for the uncertainty problems of increasing scope and complexity in building and urban simulation?**

Objective IV addressed in Chapter 6 demonstrates the fit-for-purpose of the new EmUSA framework to handle the enlarged scope (accounting for additional uncertain variables for instance, thermophysical properties of construction materials) and increased complexity (greater degrees of occupants' interaction freedom e.g. use of lights), although only at building scale. The framework has been proven to be readily extensible to incorporate a parameter screening step for the purpose of segregating the most influential input parameters from the non-influential ones. Then, the whole framework methodology of emulation, uncertainty propagation, quantification and decomposition has been applied to the practical case of multizone residential model considering the most influential parameters obtained after the parameter screening step. For this extended study, although the computational efficiency of the EmUSA framework has fallen from that of proof-of-principle study, as that is expected given the scope and complexity of the extended case but still results in reducing the total computational cost to a low fraction (0.25) of its classical counterpart.

This thesis was indeed technique rather than application focused. The

application focus is also highly important and this could be considered more fully in future work; likewise, clarifying the scope of applicability of developed technique in this broader real-world decision making context.

7.2 Future Works

This foundational research work on Emulation based Uncertainty Quantification for S-BPS can be of potential for following three (but not limited to) lines of research:

Acceleration in computation intensive tasks

Bayesian calibration [133, 134, 135, 140] and optimisation [136, 137] studies in BPS often suffer from a curse of high computation costs as they entail a detailed exploration of input space through Monte Carlo based parametrisation. And these studies assume underlying BPS being a deterministic one without the stochastic occupants' behaviours. Given the already expanded scope of the EmUSA framework to emulate both the deterministic processes and as well as the stochastic perturbations to it, the profound integration of this work therefore would possibly enable more comprehensiveness and rigorousness in future advanced studies on bayesian calibration and optimisation.

Adaptable framework for Temporal emulation in BPS

The scope of this thesis work has been constrained to simulate annual building energy demands. Given that the employed GP emulation technique already operates on multidimensional input space, it readily allows flexibility for its extension to consider time as another additional dimension (following and adapting Conti et al. [180]) and then make fast and accurate dynamic building performance predictions using "temporal" emulator. This is of particular importance when conducting sensitivity analysis of the effect of time-evolving and uncertain (current/short- and future/long-term) weather data on the building energy performance predictions.

Scaling up of uncertainties from building to urban scale

The current development of the framework is rigorous and comprehensive in

addressing the varying type (epistemic and aleatory) and nature (simple design parameters to complex thermophysical properties) of uncertainties, however only at building scale. For its enlarged scope to handle the building archetypes, EmUSA framework can be considered as a suitable candidate for the dynamic national housing stock energy simulation platform (e.g. EnHub developed in Sousa [181]) as EmUSA facilitates the use of Latin hypercube sampling to assign values to the parameters of given archetype and hence, it could then also emulate the energy dynamics of national housing stock. In order to address the challenge of accounting for uncertainties at urban scale, the EmUSA framework would then be required to account for the mutual shading effects of neighbouring buildings while developing the UQ framework for urban energy performance predictions (can be called EmUSA for Urban). This would require then the adaptation of EmUSA methodology for the simulation of urban energy flows and interactions between them (for instance, developing a similar emulator for urban energy micro-simulation platform CitySim - developed by Robinson et al. [182]). In addition to Urban Climate variability, the other uncertainties due to lack of information about existing buildings, household composition, installed HVAC systems, etc. may also play significant role in urban scale building energy performance predictions.

7.3 Recommendations

In terms of the wider applicability, the EmUSA framework can support, through a rigorous quantification and decomposition of uncertainties:

- Load and energy use prediction: More reliable load and energy use prediction, and the associated sizing of HVAC plant and renewable energy technologies.
- New building design: The more robust design of building envelopes and systems.
- Renovation design: The identification of parameters yielding the greatest potential performance improvements, identified through sensitivity analysis and the decomposition of uncertainties, and their corresponding design.

Appendix

Prior Informations used for GP emulation.

For the present work we use a Matern covariance function (see equation 4.3) with parameter $\nu = 3/2$ (as it is the most interesting case for machine learning that the Gaussian process to be less rough / dynamic and the obtained random / sample functions are smooth and differentiable [144]). Table 7.1 shows the distributions used for the hyperparameters of these functions. Table 7.2 displays the values for the hyperparameters in the mean and covariance functions. We use two distinct length scale parameter setting of covariance structure for respective two inputs - wall insulation thickness and window transmittance. For the log-variance emulators the uniform hyper-priors are used for the noise variance with small (0.001) value. The hyperpriors for the length scales and signal variances are log-Normal and uniform, respectively. We use a linear function for the mean prior mean functions for log-variance target variable. Unlike log-variance variables, we have fixed noise variances while fitting GP to mean variables as they are known following equation 4.20 from the predictions of GP fitted to log-variances. The length scales and signal variances both are assigned *loggaussian* priors. We use *constant*, *linear* and *squared* mean structures as prior mean functions for mean target variable.

Table 7.1: GP prior distribution settings for hyperparameters

Function	Hyperparameter	GP target variable			
		mean		log-variance	
		Heating	Cooling	Heating	Cooling
prior covariance	signal variance	loggaussian	loggaussian	uniform	uniform
	noise variance	fixed	fixed	uniform	uniform
	length scale	loggaussian	loggaussian	loggaussian	loggaussian

Table 7.2: GP prior value settings for hyperparameters

Hyperparameter	GP target variable			
	mean		log-variance	
	Heating	Cooling	Heating	Cooling
signal variance σ_Y^2	0.1	0.1	0.01	0.1
length scale l	[1/10, 1]	[1/10, 1/10]	[1/1000, 1/10]	[1/1000, 1/10]
prior mean $m(\mathbf{X})$	50	50	0.4	1
prior covariance $k(\mathbf{X}, \mathbf{X}')$	100	100	1	1

Bibliography

[1] Frédéric Haldi and Darren Robinson. Interactions with window openings by office occupants. *Building and Environment*, 44(12):2378–2395, 2009.

[2] Jacob Chapman. *Multi-agent stochastic simulation of occupants in buildings*. PhD thesis, The University of Nottingham, 2017.

[3] Christopher J Roy and William L Oberkampf. A comprehensive framework for verification, validation, and uncertainty quantification in scientific computing. *Computer methods in applied mechanics and engineering*, 200(25-28):2131–2144, 2011.

[4] Francesca Pianosi, Keith Beven, Jim Freer, Jim W Hall, Jonathan Rougier, David B Stephenson, and Thorsten Wagener. Sensitivity analysis of environmental models: A systematic review with practical workflow. *Environmental Modelling & Software*, 79:214–232, 2016.

[5] Wei Tian. A review of sensitivity analysis methods in building energy analysis. *Renewable and sustainable energy reviews*, 20:411–419, 2013.

[6] Anh-Tuan Nguyen and Sigrid Reiter. A performance comparison of sensitivity analysis methods for building energy models. *Building Simulation*, 8(6):651–664, 2015.

[7] Wei Tian, Yeonsook Heo, Pieter De Wilde, Zhanyong Li, Da Yan, Cheol Soo Park, Xiaohang Feng, and Godfried Augenbroe. A review of uncertainty analysis in building energy assessment. *Renewable and Sustainable Energy Reviews*, 93:285–301, 2018.

[8] United Nations. *UN-Habitat for a better Urban Future*, 2012 (accessed on March 24, 2020).

[9] European Union. *Consumption of energy in EU-28*, 2015 (accessed on March 24, 2020).

[10] European Environment Agency. *Share of renewable energy in gross final energy consumption in Europe*, 2016 (accessed on March 24, 2020).

[11] European Union. *Final energy consumption in the residential sector by type of end-use in EU-28*, 2015 (accessed on March 24, 2020).

[12] European Commission. *Europe 2020 : A Strategy for Smart, Sustainable and Inclusive Growth*, 2010. Communication from the Commission Publication Office.

[13] United Kingdom Government Legislator. *UK regulations: the Climate Change Act*, 2008 (accessed on March 25, 2020).

[14] European Commission. *Directive of the European Parliament on the energy performance of buildings*, 2010. Official Journal of the European Union.

[15] European Commission. *Directive of the European Parliament on energy efficiency*, 2010. Official Journal of the European Union.

[16] Office of Energy Efficiency and Renewable Energy. *About Building Energy Modeling*, 2019 (accessed on March 25, 2020).

[17] Torben Østergård, Rasmus L Jensen, and Steffen E Maagaard. Building simulations supporting decision making in early design–A review. *Renewable and Sustainable Energy Reviews*, 61:187–201, 2016.

[18] Joseph A Clarke, Cameron M Johnstone, Nicolas J Kelly, Paul A Strachan, and Paul Tuohy. The role of built environment energy efficiency in a sustainable UK energy economy. *Energy Policy*, 36(12):4605–4609, 2008.

[19] Chris van Dronkelaar, Mark Dowson, E Burman, Catalina Spataru, and Dejan Mumovic. A review of the energy performance gap and its underlying causes in non-domestic buildings. *Frontiers in Mechanical Engineering*, 1:17, 2016.

[20] Ingrid Allard, Thomas Olofsson, and Gireesh Nair. Energy evaluation of residential buildings: Performance gap analysis incorporating uncertainties in the evaluation methods. *Building Simulation*, 11(4):725–737, 2018.

[21] Joseph Andrew Clarke and JLM Hensen. Integrated building performance simulation: Progress, prospects and requirements. *Building and Environment*, 91:294–306, 2015.

[22] Jacob Chapman, Peer-Olaf Siebers, and Darren Robinson. On the multi-agent stochastic simulation of occupants in buildings. *Journal of Building Performance Simulation*, 11(5):604–621, 2018.

[23] Esfand Burman, Dejan Mumovic, and Judit Kimpian. Towards measurement and verification of energy performance under the framework of the European directive for energy performance of buildings. *Energy*, 77:153–163, 2014.

[24] Pieter De Wilde. The gap between predicted and measured energy performance of buildings: A framework for investigation. *Automation in Construction*, 41:40–49, 2014.

[25] Jean Lebrun Philippe Andre, Bernard Georges, Vincent Lemort, and Ion Vladut Teodorese. From model validation to production of reference simulations: how to increase reliability and applicability of building and HVAC simulation models. *Building Services Engineering Research and Technology*, 29(1):61–72, 2008.

[26] Marcus Keane Paul Raftery and James ODonnell. Calibrating whole building energy models: An evidence-based methodology. *Energy and Buildings*, 43(9):2356 – 2364, 2011.

[27] Anna Carolina Menezes, Andrew Cripps, Dino Bouchlaghem, and Richard Buswell. Predicted vs. actual energy performance of non-domestic buildings: Using post-occupancy evaluation data to reduce the performance gap. *Applied Energy*, 97:355 – 364, 2012.

[28] Frédéric Haldi and Darren Robinson. The impact of occupants' behaviour on building energy demand. *Journal of Building Performance Simulation*, 4(4):323–338, 2011.

[29] Darren Robinson and Frédéric Haldi. Modelling occupants' presence and behaviour–part I. *Journal of Building Performance Simulation*, 4:301–302, 2011.

[30] Houcem Eddine Mechri, Alfonso Capozzoli, and Vincenzo Corrado. Use of the ANOVA approach for sensitive building energy design. *Applied Energy*, 87(10):3073–3083, 2010.

[31] Clara Spitz, Laurent Mora, Etienne Wurtz, and Arnaud Jay. Practical application of uncertainty analysis and sensitivity analysis on an experimental house. *Energy and Buildings*, 55:459–470, 2012.

[32] Benjamin D Lee, Yuming Sun, Godfried Augenbroe, and Christiaan JJ Paredis. Towards better prediction of building performance: a workbench to analyze uncertainty in building simulation. In *13th International Building Performance Simulation Association Conference, Chambéry, France*, 2013.

[33] Michael D McKay, Richard J Beckman, and William J Conover. Comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 21(2):239–245, 1979.

[34] Il'ya Meerovich Sobol'. On the distribution of points in a cube and the approximate evaluation of integrals. *USSR Computational Mathematics and Mathematical Physics*, 7(4):86–112, 1967.

[35] Andrea Saltelli. Making best use of model evaluations to compute sensitivity indices. *Computer Physics Communications*, 145(2):280–297, 2002.

[36] Ilya M Sobol. Sensitivity estimates for nonlinear mathematical models. *Mathematical modelling and computational experiments*, 1(4):407–414, 1993.

[37] Ilya M Sobol. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and computers in simulation*, 55(1-3):271–280, 2001.

[38] Sebastian Burhenne. *Monte Carlo Based Uncertainty and Sensitivity Analysis for Building Performance Simulation*. PhD thesis, Karlsruhe Institute of Technology, 2013.

[39] Fanny Sarrazin, Francesca Pianosi, and Thorsten Wagener. Global sensitivity analysis of environmental models: convergence and validation. *Environmental Modelling & Software*, 79:135–152, 2016.

[40] Emmanuel Manceau, M Mezghani, I Zabalza-Mezghani, F Roggero, et al. Combination of experimental design and joint modeling methods for quantifying the risk associated with deterministic and stochastic uncertainties-An integrated test study. In *SPE Annual Technical Conference and Exhibition*. Society of Petroleum Engineers, 2001.

[41] Amandine Marrel, Bertrand Iooss, Sébastien Da Veiga, and Mathieu Ribatet. Global sensitivity analysis of stochastic computer models with joint metamodels. *Statistics and Computing*, 22(3):833–847, 2012.

[42] Nicola Stone. *Gaussian process emulators for uncertainty analysis in groundwater flow*. PhD thesis, University of Nottingham, 2011.

[43] Jacob Chapman. *Multi-agent stochastic simulation of occupants in buildings*. PhD thesis, The University of Nottingham, 2017.

[44] JA Clarke. Energy Simulation in Building Design, 2001.

[45] Drury B. Crawley, Linda K. Lawrie, Frederick C. Winkelmann, W.F. Buhl, Y.Joe Huang, Curtis O. Pedersen, Richard K. Strand, Richard J. Liesen, Daniel E. Fisher, Michael J. Witte, and Jason Glazer. EnergyPlus: creating a new-generation building energy simulation program. *Energy and Buildings*, 33(4):319 – 331, 2001.

[46] Darren Robinson, URS Wilke, and Frédéric Haldi. Multi agent simulation of occupants presence and behaviour. In *Proceedings of Building Simulation: 12th Conferernce of Interntional Building Performance Simulation Association*, pages 2110–2117, 2011.

[47] Jessen Page, Darren Robinson, Nicolas Morel, and J-L Scartezzini. A generalised stochastic model for the simulation of occupant presence. *Energy and buildings*, 40(2):83–98, 2008.

[48] 2005 ISO. *ISO 7730: Ergonomics of the thermal environment Analytical determination and interpretation of thermal comfort using calculation of the PMV and PPD indices and local thermal comfort criteria*, 2005.

[49] Frédéric Haldi and Darren Robinson. Adaptive actions on shading devices in response to local visual stimuli. *Journal of Building Performance Simulation*, 3(2):135–153, 2010.

[50] Christoph F Reinhart. Lightswitch-2002: a model for manual and automated control of electric lighting and blinds. *Solar energy*, 77(1):15–28, 2004.

[51] Scott Pigg, Mark Eilers, and John Reed. Behavioral aspects of lighting and occupancy sensors in private offices: a case study of a university office building. In *Behavioral Aspects of Lighting and Occupancy Sensors in Private Offices*, 1996.

[52] Jacob Chapman, Peer-Olaf Siebers, and Darren Robinson. Data-scare behavioural modelling and the representation of social interactions in building simulation. *Energy*. In preparation.

[53] Jared Langevin, Jin Wen, and Patrick L Gurian. Including occupants in building performance simulation: integration of an agent-based occupant behavior algorithm with Energy Plus. In *ASHRAE/IBPSAUSA Buidling Simulation Conference. Atlanta, GA*, 2014.

[54] Tianzhen Hong, Hongsan Sun, Yixing Chen, Sarah C Taylor-Lange, and Da Yan. An occupant behavior modeling tool for co-simulation. *Energy and Buildings*, 117:272–281, 2016.

[55] Bing Dong, Da Yan, Zhaoxuan Li, Yuan Jin, Xiaohang Feng, and Hannah Fontenot. Modeling occupancy and behavior for better building design and operation A critical review. *Building Simulation*, 11(5):899–921, 2018.

[56] Andreas Wagner, William OBrien, and Bing Dong. *Exploring Occupant Behavior in Buildings Methods and Challenges*. Springer, 2017.

[57] H Burak Gunay, William O'Brien, and Ian Beausoleil-Morrison. A critical review of observation studies, modeling, and simulation of adaptive occupant behaviors in offices. *Building and Environment*, 70:31–47, 2013.

[58] George EP Box. Robustness in the strategy of scientific model building. In *Robustness in statistics*, pages 201–236. Elsevier, 1979.

[59] Claude Elwood Shannon. A mathematical theory of communication. *Bell system technical journal*, 27(3):379–423, 1948.

[60] Andrea Saltelli, Marco Ratto, Terry Andres, Francesca Campolongo, Jessica Cariboni, Debora Gatelli, Michaela Saisana, and Stefano Tarantola. *Global sensitivity analysis: the primer*. John Wiley & Sons, 2008.

[61] JR Taylor. An introduction to error analysis: the study of uncertainties in physical measurements. *Science, Sausalito, CA*, 45:349, 1997.

[62] JM Hammersley and DC Handscomb. Monte Carlo methods. *Methuen's Monographs on Applied Probability and Statistics, London*, 1964.

[63] Robert E Shannon. Introduction to the art and science of simulation. In *Proceedings of the 30th conference on Winter simulation*, pages 7–14. IEEE Computer Society Press, 1998.

[64] Marc C Kennedy and Anthony O'Hagan. Bayesian calibration of computer models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(3):425–464, 2001.

[65] Andrea Saltelli, Stefano Tarantola, Francesca Campolongo, and Marco Ratto. Sensitivity analysis in practice: a guide to assessing scientific models. *Chichester, England*, 2004.

[66] GS Watson. A study of the group screening method. *Technometrics*, 3(3):371–388, 1961.

[67] MS Patel. Group-screening with more than two stages. *Technometrics*, 4(2):209–217, 1962.

[68] JE Jacoby and S Harrison. Multi-variable experimentation and simulation models. *Naval Research Logistics Quarterly*, 9(2):121–136, 1962.

[69] Bert Bettonvil and Jack PC Kleijnen. Searching for important factors in simulation models with many factors: Sequential bifurcation. *European Journal of Operational Research*, 96(1):180–194, 1997.

[70] Max D Morris. Two-stage factor screening procedures using multiple grouping assignments. *Communications in Statistics-Theory and Methods*, 16(10):3051–3067, 1987.

[71] Sarah C Cotter. A screening design for factorial experiments with interactions. *Biometrika*, 66(2):317–320, 1979.

[72] Terry H Andres and Wayne C Hajas. Using iterated fractional factorial design to screen parameters in sensitivity analysis of a probabilistic risk assessment model. In *Proceedings of the Joint International Conference on Mathematical Methods and Supercomputing in Nuclear Applications, Karlsruhe, Germany Eds. H. Keusters, E. Stein and W. Werner*, volume 25, pages 328–337, 1993.

[73] Max D Morris. Factorial sampling plans for preliminary computational experiments. *Technometrics*, 33(2):161–174, 1991.

[74] Francesca Campolongo, Jessica Cariboni, and Andrea Saltelli. An effective screening design for sensitivity analysis of large models. *Environmental Modelling & Software*, 22(10):1509 – 1518, 2007.

[75] Nicholas Metropolis and Stanislaw Ulam. The monte carlo method. *Journal of the American statistical association*, 44(247):335–341, 1949.

[76] Stefano Tarantola, William Becker, and Dirk Zeitz. A comparison of two sampling methods for global sensitivity analysis. *Computer Physics Communications*, 183(5):1061–1072, 2012.

[77] Lauwerens Kuipers and Harald Niederreiter. *Uniform distribution of sequences*. A Wiley-Interscience Publication, 2012.

[78] Russel E Caflisch. Monte carlo and quasi-monte carlo methods. *Acta numerica*, 7:1–49, 1998.

[79] William L Oberkampf and Christopher J Roy. *Verification and validation in scientific computing*. Cambridge University Press, 2010.

[80] Shankar Sankararaman and Sankaran Mahadevan. Separating the contributions of variability and parameter uncertainty in probability distributions. *Reliability Engineering & System Safety*, 112:187–199, 2013.

[81] Eduard Hofer, Martina Kloos, Bernard Krzykacz-Hausmann, Jörg Peschke, and Martin Woltereck. An approximate epistemic uncertainty analysis approach in the presence of epistemic and aleatory uncertainties. *Reliability Engineering & System Safety*, 77(3):229–238, 2002.

[82] K Durga Rao, HS Kushwaha, Ajit Kumar Verma, and Ajit Srividya. Quantification of epistemic and aleatory uncertainties in level-1 probabilistic safety assessment studies. *Reliability Engineering & System Safety*, 92(7):947–956, 2007.

[83] Shankar Sankararaman and Sankaran Mahadevan. Distribution type uncertainty due to sparse and imprecise data. *Mechanical Systems and Signal Processing*, 37(1-2):182–198, 2013.

[84] Emanuele Borgonovo. Measuring uncertainty importance: investigation and comparison of alternative approaches. *Risk analysis*, 26(5):1349–1361, 2006.

[85] Emanuele Borgonovo and S Tarantola. Moment independent and variance-based sensitivity analysis with correlations: An application to the stability of a chemical reactor. *International Journal of Chemical Kinetics*, 40(11):687–698, 2008.

[86] John W Dalle Molle and Douglas J Morrice. An information theoretic approach to computer simulation sensitivity analysis. In *Proceedings of 1993 Winter Simulation Conference-(WSC'93)*, pages 402–407. IEEE, 1993.

[87] Chang K Park and Kwang-Il Ahn. A new approach for measuring uncertainty importance and distributional sensitivity in probabilistic safety assessment. *Reliability Engineering & System Safety*, 46(3):253–261, 1994.

[88] Niklas Lüdtke, Stefano Panzeri, Martin Brown, David S Broomhead, Joshua Knowles, Marcelo A Montemurro, and Douglas B Kell. Information-theoretic sensitivity analysis: a general method for credit assignment in complex networks. *Journal of The Royal Society Interface*, 5(19):223–235, 2007.

[89] Moon-Hyun Chun, Seok-Jung Han, and Nam-IL Tak. An uncertainty importance measure using a distance metric for the change in a cumulative distribution function. *Reliability Engineering & System Safety*, 70(3):313–321, 2000.

[90] Emanuele Borgonovo. A new uncertainty importance measure. *Reliability Engineering & System Safety*, 92(6):771–784, 2007.

[91] Qiao Liu and Toshimitsu Homma. A new computational method of a moment-independent uncertainty importance measure. *Reliability Engineering & System Safety*, 94(7):1205–1211, 2009.

[92] Emanuele Borgonovo and Elmar Plischke. Sensitivity analysis: a review of recent advances. *European Journal of Operational Research*, 248(3):869–887, 2016.

[93] RI Cukier, CM Fortuin, Kurt E Shuler, AG Petschek, and JH Schaibly. Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. I Theory. *The Journal of chemical physics*, 59(8):3873–3878, 1973.

[94] Bradley Efron and Charles Stein. The jackknife estimate of variance. *The Annals of Statistics*, 9:586–596, 1981.

[95] Toshimitsu Homma and Andrea Saltelli. Importance measures in global sensitivity analysis of nonlinear models. *Reliability Engineering & System Safety*, 52(1):1–17, 1996.

[96] Andrea Saltelli and Ricardo Bolado. An alternative way to compute Fourier amplitude sensitivity test (FAST). *Computational Statistics & Data Analysis*, 26(4):445–460, 1998.

[97] Jon C Helton and Freddie Joe Davis. Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. *Reliability Engineering & System Safety*, 81(1):23–69, 2003.

[98] Jon C Helton, Jay Dean Johnson, Cedric J Sallaberry, and Curt B Storlie. Survey of sampling-based methods for uncertainty and sensitivity analysis. *Reliability Engineering & System Safety*, 91(10-11):1175–1209, 2006.

[99] Andrea Saltelli, Paola Annoni, Ivano Azzini, Francesca Campolongo, Marco Ratto, and Stefano Tarantola. Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. *Computer Physics Communications*, 181(2):259–270, 2010.

[100] Andrea Saltelli, Stefano Tarantola, and KP-S Chan. A quantitative model-independent method for global sensitivity analysis of model output. *Technometrics*, 41(1):39–56, 1999.

[101] Kevin J Lomas and Herbert Eppel. Sensitivity analysis techniques for building thermal simulation programs. *Energy and buildings*, 19(1):21–44, 1992.

[102] Michael Stendert De Wit. *Uncertainty in predictions of thermal comfort in buildings*. PhD thesis, TU Delft, Delft University of Technology, 2001.

[103] MS De Wit. Identification of the important parameters in thermal building simulation models. *Journal of statistical computation and simulation*, 57(1-4):305–320, 1997.

[104] Sten De Wit. Influence of modeling uncertainties on the simulation of building thermal comfort performance. In *Building Simulation*, volume 5. Citeseer, 1997.

[105] Iain Alexander Macdonald. *Quantifying the effects of uncertainty in building simulation*. PhD thesis, University of Strathclyde Glasgow, 2002.

[106] Thierry A Mara and Stefano Tarantola. Application of global sensitivity analysis of model output to building thermal simulations. *Building Simulation: An International Journal*, 1(4):290–302, 2008.

[107] EN ISO. 13790: 2008. *Energy performance of buildings-Calculation of energy use for space heating and cooling*, 2008.

[108] Janelle S Hygh, Joseph F DeCarolis, David B Hill, and S Ranji Ranjithan. Multivariate regression as an energy assessment tool in early building design. *Building and Environment*, 57:165–175, 2012.

[109] Yeonsook Heo. *Bayesian calibration of building energy models for energy retrofit decision-making under uncertainty*. PhD thesis, Georgia Institute of Technology, 2011.

[110] Yuming Sun. *Closing the building energy performance gap by improving our predictions*. PhD thesis, Georgia Institute of Technology, 2014.

[111] Qinpeng Wang. *Accuracy, validity and relevance of probabilistic building energy models*. PhD thesis, Georgia Institute of Technology, 2016.

[112] Qinpeng Wang and Godfried Augenbroe. Combined Sensitivity Ranking of Input Parameters and Model Forms of Building Energy Simulation. In *Proceedings of the 15th IBPSA Conference. San Francisco, CA, USA2017*, pages 2042–2049, 2017.

[113] Qinpeng Wang, Godfried Augenbroe, Ji-Hyun Kim, and Li Gu. Meta-modeling of occupancy variables and analysis of their impact on energy outcomes of office buildings. *Applied Energy*, 174:166–180, 2016.

[114] Per Heiselberg, Henrik Brohus, Allan Hesselholt, Henrik Rasmussen, Erkki Seinre, and Sara Thomas. Application of sensitivity analysis in design of sustainable buildings. *Renewable Energy*, 34(9):2030–2036, 2009.

[115] Alfonso Capozzoli, Houcem Eddine Mechri, and Vincenzo Corrado. Impacts of architectural design choices on building energy performance applications of uncertainty and sensitivity techniques. *International Building Performance Simulation Association*, 2009.

[116] Vincenzo Corrado and Houcem Eddine Mechri. Uncertainty and sensitivity analysis for building energy rating. *Journal of building physics*, 33(2):125–156, 2009.

[117] AT Booth, R Choudhary, and DJ Spiegelhalter. Handling uncertainty in housing stock models. *Building and Environment*, 48:35–47, 2012.

[118] D Garcia Sanchez, Bruno Lacarrière, Marjorie Musy, and Bernard Bourges. Application of sensitivity analysis in building energy simulations: Combining first-and second-order elementary effects methods. *Energy and Buildings*, 68:741–750, 2014.

[119] Yusuf Yildiz, Koray Korkmaz, Trkan Gksal zbalta, and Zeynep Durmus Arsan. An approach for developing sensitive design parameter guidelines to reduce the energy requirements of low-rise apartment buildings. *Applied Energy*, 93:337 – 347, 2012.

[120] Gloria Calleja Rodrguez, Antonio Carrillo Andrs, Fernando Domnguez Muoz, Jos Manuel Cejudo Lpez, and Yi Zhang. Uncertainties and sensitivity analysis in building energy simulation using macroparameters. *Energy and Buildings*, 67:79 – 87, 2013.

[121] Arthur Santos Silva and Enedir Ghisi. Uncertainty analysis of user behaviour and physical parameters in residential building performance simulation. *Energy and Buildings*, 76:381 – 391, 2014.

[122] A Ioannou and LCM Itard. Energy performance and comfort in residential buildings: Sensitivity for building parameters and occupancy. *Energy and Buildings*, 92:216–233, 2015.

[123] Douglas P Boyle, Hoshin V Gupta, and Soroosh Sorooshian. *Calibration of Watershed Models : Multicriteria calibration of hydrologic models*. Wiley Online Library, 2003.

[124] Jing Yang. Convergence and uncertainty analyses in Monte-Carlo based sensitivity analysis. *Environmental Modelling & Software*, 26(4):444–457, 2011.

[125] Jack PC Kleijnen and Jon C Helton. Statistical analyses of scatterplots to identify important factors in large-scale simulations, 1: Review and comparison of techniques. *Reliability Engineering & System Safety*, 65(2):147–185, 1999.

[126] Jon C Helton and FJ Davis. Illustration of sampling-based methods for uncertainty and sensitivity analysis. *Risk analysis*, 22(3):591–622, 2002.

[127] Curtis B Storlie, Laura P Swiler, Jon C Helton, and Cedric J Sallaberry. Implementation and evaluation of nonparametric regression procedures for sensitivity analysis of computationally demanding models. *Reliability Engineering & System Safety*, 94(11):1735–1763, 2009.

[128] Parag Rastogi. *On the sensitivity of buildings to climate: the interaction of weather and building envelopes in determining future building energy consumption*. PhD thesis, Ecole Polytechnique Fédérale de Lausanne, 2016.

[129] Min-Yuan Cheng and Minh-Tu Cao. Accurately predicting building energy performance using evolutionary multivariate adaptive regression splines. *Applied Soft Computing*, 22:178–188, 2014.

[130] Athanasios Tsanas and Angeliki Xifara. Accurate quantitative estimation of energy performance of residential buildings using statistical machine learning tools. *Energy and Buildings*, 49:560–567, 2012.

[131] Joshua Hester, Jeremy Gregory, and Randolph Kirchain. Sequential early-design guidance for residential single-family buildings using a probabilistic metamodel of energy consumption. *Energy and Buildings*, 134:202–211, 2017.

[132] Xi Chen, Hongxing Yang, and Ke Sun. Developing a meta-model for sensitivity analyses and prediction of building performance for passively designed high-rise residential buildings. *Applied energy*, 194:422–439, 2017.

[133] Matthew Riddle and Ralph T Muehleisen. A guide to Bayesian calibration of building energy models. In *Building Simulation Conference*, 2014.

[134] Qi Li, Godfried Augenbroe, and Jason Brown. Assessment of linear emulators in lightweight Bayesian calibration of dynamic building energy models for parameter estimation and performance prediction. *Energy and Buildings*, 124:194–202, 2016.

[135] Jun Yuan, Victor Nian, Bin Su, and Qun Meng. A simultaneous calibration and parameter ranking method for building energy models. *Applied Energy*, 206:657–666, 2017.

[136] Bryan Eisenhower, Zheng ONeill, Satish Narayanan, Vladimir A Fonoberov, and Igor Mezić. A methodology for meta-model based optimization in building energy models. *Energy and Buildings*, 47:292–301, 2012.

[137] Liesje Van Gelder, Hans Janssen, and Staf Roels. Probabilistic design and analysis of building performances: methodology and application example. *Energy and Buildings*, 79:202–211, 2014.

[138] Michael Wetter and Elijah Polak. A convergent optimization method using pattern search algorithms with adaptive precision simulation. *Building services engineering research and technology*, 25(4):327–338, 2004.

[139] Somayeh Asadi, Shideh Shams Amiri, and Mohammad Mottahedi. On the development of multi-linear regression analysis to assess energy consumption in the early stages of building design. *Energy and Buildings*, 85:246–255, 2014.

[140] Massimiliano Manfren, Niccolò Aste, and Reza Moshksar. Calibration and uncertainty analysis for computer models–a meta-model based approach for integrated building energy simulation. *Applied energy*, 103:627–641, 2013.

[141] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The elements of statistical learning data mining, inference, and prediction*. Springer Heidelberg, 2009.

[142] Jerome H Friedman. Multivariate adaptive regression splines. *The annals of statistics*, 19(1):1–141, 1991.

[143] Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. *An introduction to statistical learning*, volume 112. Springer, 2013.

[144] Carl Edward Rasmussen and Christopher KI Williams. *Gaussian processes for machine learning*, volume 1. MIT press Cambridge, 2006.

[145] Vladimir Vapnik, Steven E Golowich, and Alex J Smola. Support vector method for function approximation, regression estimation and signal processing. In *Advances in neural information processing systems*, pages 281–287, 1996.

[146] B Yildiz, JI Bilbao, and AB Sproul. A review and analysis of regression and machine learning models on commercial building electricity load forecasting. *Renewable and Sustainable Energy Reviews*, 73:1104–1122, 2017.

[147] Kadir Amasyali and Nora M El-Gohary. A review of data-driven building energy consumption prediction studies. *Renewable and Sustainable Energy Reviews*, 81:1192–1205, 2018.

[148] Torben Østergård, Rasmus Lund Jensen, and Steffen Enersen Maagaard. A comparison of six metamodeling techniques applied to building performance simulations. *Applied Energy*, 211:89–103, 2018.

[149] Lai Wei, Wei Tian, Jian Zuo, Zhi-Yong Yang, YunLiang Liu, and Song Yang. Effects of building form on energy use for buildings in cold climate regions. *Procedia Engineering*, 146:182–189, 2016.

[150] Hyunwoo Lim and Zhiqiang John Zhai. Comprehensive evaluation of the influence of meta-models on Bayesian calibration. *Energy and Buildings*, 155:66–75, 2017.

[151] Ruichen Jin, Wei Chen, and Timothy W Simpson. Comparative studies of metamodelling techniques under multiple modelling criteria. *Structural and multidisciplinary optimization*, 23(1):1–13, 2001.

[152] Parag Rastogi, Mohammad Emtiyaz Khan, and Marilyne Andersen. Gaussian-Process-Based Emulators for Building Performance Simulation. In *Proceedings of the Fifteenth IBPSA Conference*, pages 1701–1709, 2017.

[153] Hai-xiang Zhao and Frédéric Magoulès. A review on the prediction of building energy consumption. *Renewable and Sustainable Energy Reviews*, 16(6):3586–3592, 2012.

[154] Yixuan Wei, Xingxing Zhang, Yong Shi, Liang Xia, Song Pan, Jinshun Wu, Mengjie Han, and Xiaoyun Zhao. A review of data-driven approaches for prediction and classification of building energy consumption. *Renewable and Sustainable Energy Reviews*, 82:1027–1047, 2018.

[155] Soteris A Kalogirou. Artificial neural networks in energy applications in buildings. *International Journal of Low-Carbon Technologies*, 1(3):201–216, 2006.

[156] T. Olofsson and S. Andersson. Long-term energy demand predictions based on short-term measured data. *Energy and Buildings*, 33(2):85 – 91, 2001.

[157] Won-Yong Lee, John M House, and Nam-Ho Kyong. Subsystem level fault diagnosis of a building's air-handling unit using general regression neural networks. *Applied Energy*, 77(2):153–170, 2004.

[158] Aurélie Foucquier, Sylvain Robert, Frédéric Suard, Louis Stéphan, and Arnaud Jay. State of the art in building modelling and energy performances prediction: A review. *Renewable and Sustainable Energy Reviews*, 23:272–288, 2013.

[159] Young-Jin Kim, Ki-Uhn Ahn, CS Park, and In-Han Kim. Gaussian emulator for stochastic optimal design of a double glazing system. In *Proceedings of the 13th IBPSA Conference, August*, pages 2217–2224, 2013.

[160] JE Yan, Young-Jin Kim, Ki-Uhn Ahn, and Cheol-Soo Park. Gaussian process emulator for optimal operation of a high rise office building. In *Proceedings of 13th International Building Performance Simulation Association Conference*, pages 2225–2231, 2013.

[161] Siamak Safarzadegan Gilan and Bistra Dilkina. Sustainable Building Design: A Challenge at the Intersection of Machine Learning and Design Optimization. In *AAAI Workshop: Computational Sustainability*, 2015.

[162] Michael James Wood. *An exploration of building design and optimisation methods using Kriging meta-modelling*. PhD thesis, University of Exeter, 2016.

[163] Filippo Monari. *Sensitivity analysis and Bayesian calibration of building energy models*. PhD thesis, University of Strathclyde, 2016.

[164] Anthony OHagan. Bayesian analysis of computer code outputs: a tutorial. *Reliability Engineering & System Safety*, 91(10):1290–1300, 2006.

[165] Leonardo S Bastos and Anthony OHagan. Diagnostics for Gaussian process emulators. *Technometrics*, 51(4):425–438, 2009.

[166] Younus Al-Taweel. *Diagnostics and Simulation-Based Methods for Validating Gaussian Process Emulators*. PhD thesis, University of Sheffield, 2018.

[167] Bruce Ankenman, Barry L Nelson, and Jeremy Staum. Stochastic kriging for simulation metamodeling. *Operations research*, 58(2):371–382, 2010.

[168] Jarno Vanhatalo, Jaakko Riihimäki, Jouni Hartikainen, Pasi Jylänki, Ville Tolvanen, and Aki Vehtari. GPstuff: Bayesian modeling with Gaussian processes. *Journal of Machine Learning Research*, 14(Apr):1175–1179, 2013.

[169] Paul W. Goldberg, Christopher K. I. Williams, and Christopher Bishop. *Regression with Input-Dependent Noise: A Gaussian Process Treatment*, volume 10. MIT Press, advances in neural information processing systems edition, January 1998. Advances in Neural Information Processing Systems.

[170] Kristian Kersting, Christian Plagemann, Patrick Pfaff, and Wolfram Burgard. Most Likely Heteroscedastic Gaussian Process Regression. In *Proceedings of the 24th International Conference on Machine Learning*, ICML '07, pages 393–400, New York, NY, USA, 2007. ACM.

[171] Wenjing Wang and Xi Chen. The effects of estimation of heteroscedasticity on stochastic kriging. In *Proceedings of the 2016 Winter Simulation Conference*, pages 326–337. IEEE Press, 2016.

[172] Alexis Boukouvalas and Dan Cornford. Learning heteroscedastic Gaussian processes for complex datasets. *Technical report*, 2009.

[173] David Williams. *Probability with Martingales*. Cambridge University Press, 1991.

[174] Francesca Pianosi, Fanny Sarrazin, and Thorsten Wagener. A Matlab toolbox for global sensitivity analysis. *Environmental Modelling & Software*, 70:80–85, 2015.

[175] GEB Archer, Andrea Saltelli, and IM Sobol. Sensitivity measures, ANOVA-like techniques and the use of bootstrap. *Journal of Statistical Computation and Simulation*, 58(2):99–120, 1997.

[176] Ning Quan, Jun Yin, Szu Hui Ng, and Loo Hay Lee. Simulation optimization via kriging: a sequential search using expected improvement with computing budget constraints. *IIE Transactions : Advances in Simulation Optimization and Its Applications*, 45(7):763–780, 2013.

[177] International Organization for Standardization. *Ergonomics of the Thermal Environment: Analytical Determination and Interpretation of Thermal Comfort Using Calculation of the PMV and PPD Indices and Local Thermal Comfort Criteria*. International Organization for Standardization, 2005.

[178] JA Clarke, PP Yaneske, and AA Pinney. The harmonisation of thermal properties of building materials. *BRE, UK*, 1990.

[179] Fernando Domínguez-Muñoz, Brian Anderson, José M Cejudo-López, and Antonio Carrillo-Andrés. Uncertainty in the thermal conductivity of insulation materials. *Energy and Buildings*, 42(11):2159–2168, 2010.

[180] Stefano Conti, John Paul Gosling, Jeremy E Oakley, and Anthony O'Hagan. Gaussian process emulation of dynamic computer codes. *Biometrika*, 96(3):663–676, 2009.

[181] Gustavo Sousa. *Analysis of strategies to decarbonise the UK housing stock*. PhD thesis, The University of Nottingham, 2017.

[182] Darren Robinson, Frédéric Haldi, Philippe Leroux, Diane Perez, Adil Rasheed, and Urs Wilke. CitySim: Comprehensive micro-simulation

of resource flows for sustainable urban planning. In *Proceedings of the Eleventh International IBPSA Conference*, pages 1083–1090, 2009.