博士論文

A study of the uncertainty considered decision-making framework using sensitivity analysis and Gaussian process in the early stage of architecture environmental design

(感度分析とガウス過程による不確実性を考慮した
建築環境設計初期段階における
意思決定フレームワークに関する研究)

高　雲

Contents

List of Figures

List of Tables

Chapter 1

Introduction and Literature Review

This chapter narrates the overall story of this research. Research background is described in the first section. By placing emphasis on the uncertainty problems and the obstacles to solving them, the importance of this research is explained. The second section states the research purpose, by answering the questions that what can be done by a researcher for architects and occupants when facing uncertainty problems. The third section consists of the reviews of available tools and literature from different disciplines, mainly AEC[1] and CS[2], which declares the state of art and shows the gap. Rather than reviewing detailed procedures, this section is introduced in a wide perspective, identifying the research fields and interests. The composition of the whole thesis is described in the last section. The main targets of the following chapters and how they are related to the research purpose are explained.

[1] Architecture, Engineering and Construction
[2] Computer Science

1.1 Background

Buildings are responsible for more than 40 percent of global energy use and one third of greenhouse gas emissions globally.[23] In Japan, since the oil crisis in the 1970s, the energy consumption by building sector considerably increased by about 250%, which accounts for 34.5% of the total energy use in the country.[40] In the report titled *International energy outlook 2017*, U.S. Energy Information Administration estimated that, with the progress of urbanization, the energy use of building department will continuously increase by about 30% in the coming 20 years.[43] During the building life cycle, the operation stage overwhelm others in energy consumption[16], which indicates the largest potential in energy saving on the other hand. In both residential and commercial buildings, HVAC and lighting, which maintain the physical comfort of the indoor environment, consume the most energy. Moreover, comparing to others, the energy consumption of HVAC and lighting is more closely related to architecture design. With the popularization of LED, the energy consumption by lighting become less serious than before. How to reduced the energy consumption of HVAC system will be the key to building energy saving.

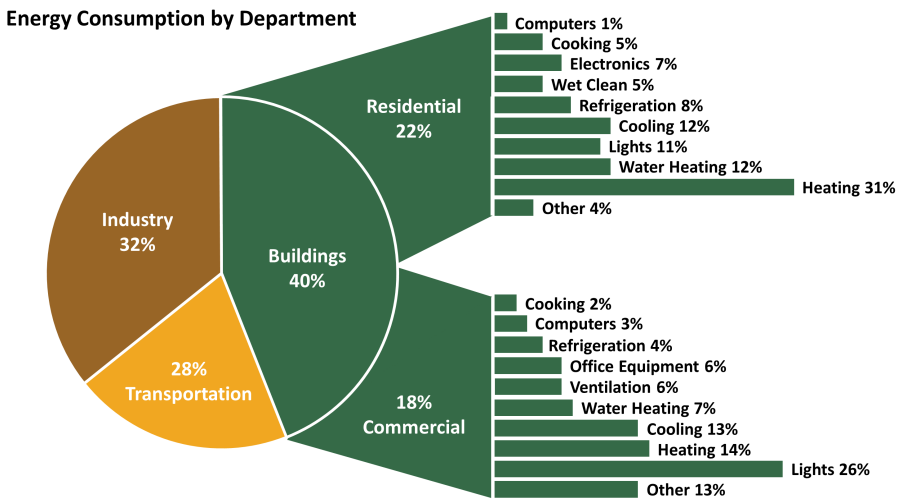


Figure 1-1: The partitions of energy consumption by different department
(Data source: Laurent Pilon's Research Group, UCLA | 2011)

1.1.1 The energy saving potential of architectural design in early stage

It has been proved that the architecture design, especially in the early stage, has a significant impact on the energy performance[35]. As illustrated in figure 1-2, the cost of changing design is low in the vary early stage and will get higher and higher in the later stage, while the effect gets lower.[41] However, the most important decisions are made in the early design stage by architects[13], usually with some rules of thumbs. With the energy codes getting more and more strict in recent years, it has been appointed out by some researchers that

guild lines or rule of thumbs are not enough to ensure the energy performance[53] to meet these codes. Some researchers suggested that Integrated Design Process (IDP) would be required to achieve ultra-low-energy design.[55] Quantitative energy analysis should be carried out in the early stage. Many researchers have made a lot efforts to integrate quantitative energy performance assessment smoothly into the early design stage, which is explained with details in the Literature review section. In 2008, the

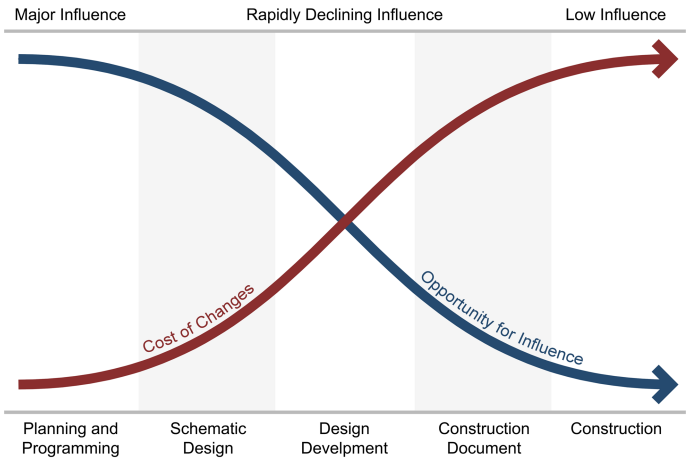


Figure 1-2: Cost and opportunity to to make changes is during programming
(Data source: Whole building design guide | 2016)

1.1.2 The gap between assessed and real energy performance

Not only the academic but also the industry has understood the importance of energy performance assessment in the early design stage. Computer simulations are widely used in energy performance assessment. However, the gap between the assessed energy performance, based on the simulations, results and the real energy consumption, in operational stage, is widely acknowledged.[46]

To study the gap between assessed and real energy performance, many researchers put their eyes on LEED[67] rating system. LEED BD+C[62] adopts the energy performance assessment method from ASHRAE Standard 90.1[61] appendix G, a simulation-based method which is comprehensive and highly completed, as one way to get energy credits. On the other hand, LEED is worldwide popular and successful in business. With the number of LEED certified buildings getting larger, LEED became a good study subject and therefore draw attentions from researchers.

In 2008, funded by the USGBC[3], the mother organization of LEED, NBI[4] carried out a survey on the measured energy performance of the LEED certified buildings.[10] 121 occupied LEED buildings were included

[3]U.S. Green Building Council
[4]New Buildings Institute

in this study. The designed and measured Energy Use Intensities[78] was compared, as well as the proposed and measured saving (figure 1-3). The researchers concluded that though, on average, LEED buildings are delivering anticipated savings, the measured performance displays a large degree of scatter. The also suggested that the baseline used in LEED was not aggressive enough as anticipated. In another word, the designed EUIs were underestimated. Though the researchers did not calculate the Pearson correlation coefficient, it can be estimated to be around 0.5.

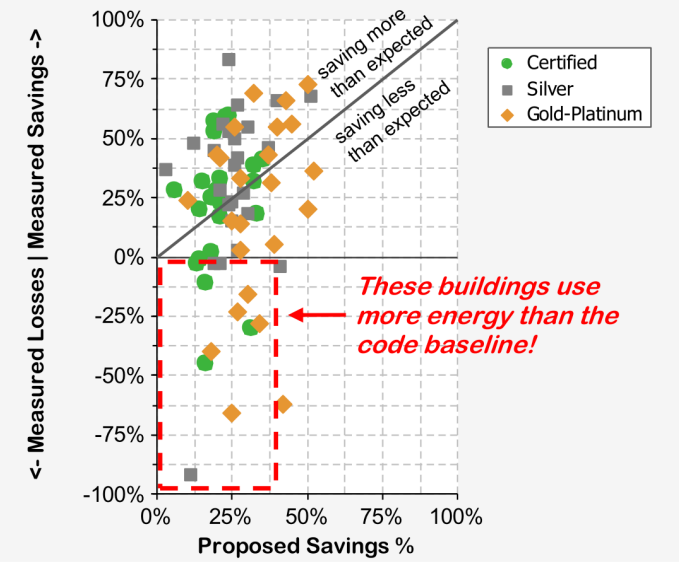


Figure 1-3: Measured versus Proposed Savings Percentages
(New Buildings Institute | 2008)

Other researchers also did study of the real energy performance of LEED and come to similar conclusions.[12, 14] It can be inferred that the gap between assessed and real energy performance, no matter it is LEED certified or not, should be non-negligible.

1.1.3 The reason that this gap matters

First of first, the real energy performance of buildings is certainly the most important.

It is interesting that when I did the literature review of uncertainty in energy performance, Energy Performance Contracting (EPC) was a extremely high-frequent keyword. After all, people become serious when it comes to money. A EPC is made between the building owner/user and a contractor, which is usually a Energy Saving Company (ESCO). The contractors offer energy design supports or retrofit plans to the owner to

improve the energy performance of the designs or existing buildings. Energy performance objectives would be decided when contracting. The contractors will be paid based on the measured energy performance after the contract is finished. Briefly speaking, the building owners/users will pay the contractors if the measured energy performance met the objectives in the contracts. Otherwise the contractors should compensate the owners/users in some cases. Before the first year of operation, there is no measured data to support decision making, so the objectives would be decided based on the assessed energy performance, usually simulations. In this case, the gap between assessed and measured energy performance dose not only represent a failure of design, but also means financial risks for the ESCOs. The EPCs is now wildly accepted and adopted all around the world. JAESCO[5] has defined the EPC as one of the features of an ESCO.[45] In Europe, eu.ESCO[6] declares that its accelerate the adoption of EPC to assist the EU members to achieve their energy efficiency targets.[68] EPC is also included in the EU Energy Efficiency Directive 2012.[51] In Australia, there are kinds of finance innovation for retrofit projects, which is called Environmental Upgrade Agreement (EUA)[18].

Since the beginning of this century, building energy codes has been adopted in many countries, such as *GB 50189*[32] in China, Title 24 in California, etc. Sustainable building certification systems, such as LEED, CASBEE were also developed and widely adopted. However, these codes are actually design standards, which specify the building specs, like the insulation of envelop, or specify the calculation methods of energy performance and set a threshold. In the last decade, more codes and certifications put more emphasis on the measured energy performance. In LEED v4, the most recent version, the building-level energy metering is required in LEED BD+C, which is a newly added part compared to the 2009 version. This building-level energy metering is also required when applying for the certification of LEED O+M.

In Japan, the Zero Energy House (ZEH) is gradually becoming mandatory[40]. In April 2014, the council decision approved 'The 4^{th} Strategic Energy plan'[7], in which it has been stipulated that, by 2020, all the newly constructed standard residential houses should be ZEH and, by 2030, all the newly constructed residential buildings should be ZEH in average. The certification of ZEH is based on the balance of measured energy consumed and generated onsite. Therefore, the gap between calculated and measured energy consumption would means failing to meet the codes. The Zero Net Energy (ZNE) in US and Europe shares the similar goals to ZEH in Japan and uses similar certification methods as well. In September, 2008, California Public Utilities Commission declared 2 major goals in California long term energy efficiency strategic plan: (1) Before 2020 all new residential buildings and (2) before 2030 all new commercial buildings should be ZNE[8]. The European Union has also established similar goals. Before 2018, all new buildings that are owned and occupied by public authorities will be 'nearly ZNE'. By 2020 all new buildings will be 'nearly ZNE'[23].

[5] Japan Association of Energy Service Companies
[6] European Association of Energy Services Companies
[7] 第 4 次エネルギー基本計画

1.1.4 Uncertainties in the inputs of simulations

Researchers has pointed out that this gap between assessed and real energy performance comes from the uncertainties of the input information. The uncertainties in the results are introduced by the errors in some common assumptions which is normally used in the energy performance simulations. These assumptions usually consist of several hard-to-measure factors[27], including building-related (e.g. air-tightness level) and occupant-related (e.g. AC setpoints) ones. The values of factors in these assumptions are usually from codes and design handbooks, such as ASHRAE 90.1, or from experience. These assumed values could hardly reproduce the reality in every building and every scenario. A simulation done with single assumed values are called deterministic simulation by many researchers.

The derivation of the uncertainties is actually lack of information. Various kinds of information are not accessible theoretically.

A model used in the energy performance simulations is a simplified reflection of the real building in virtual space. There is always a gap between a computational model and the real building, especially the geometries. The propensities of the building parts used in calculations are always, more or less, different from the real ones. Thermal bridges also have impact on the insulation level of the facade and furthermore influence the energy performance. However, thermal bridges can hardly be quantitatively analyzed until the construction of the facade is decided, which is not accessible in the early stage. Usually, in the early stage, the properties of the glass used on openings are considered. However, the influence from the window frame, which is sometimes stronger than that from the glass, is not well considered. Thermal bridge and window frames issues could be studied using THERM and WINDOW software from LBNL[cite]. However, detailed construction drawings are required, and it takes time to do the analysis. Similarly, air tightness also enlarge the gap between assessed and real energy performance.

The most unpredictable factors in the design stage are those occupant-related ones. Even the usage of a building is known, how people will use this building is still hard to predict, as every thing is so dynamic. Compared to residential buildings, office buildings have less uncertainties. However, the exact schedules are still not known, especially in East Asia, as we have ‘overtime culture’ here. The cooling and heating setpoints are not as dynamic as that in a residential building. Nevertheless, due to the uneven temperature distribution and the error in sensors, the real temperature is high-probably different from the setpoint. Computers are indispensable in a modern office. However, the heat generation of computer varies dramatically with its hardware level and works executed. What more, the density of the computers are not predictable. Uncertainties in lighting are similar issues. Natural ventilation is also another factor that with high uncertainties, as it is related to both the habit of the users and the weather.

The climate data in the energy performance simulations also introduces uncertainties into the results. Normally, Typical meteorological year (TMY)[71] data is used, which is close to (not exact) the average values of observation years. A month, whose data is most close to the average of all the same months, will be used as the typical meteorological month. For example, in the case that the climate data is from 1985 to 2005, if

the data of the March in 1990 is most close to the average of all Marches of the 21 years, then it will be used in the TMY data. As it is a selected data, the gap from the real climate cannot be ignored.

The climate change is another reason that TMY data fails to reproduce the current. TMY data are usually from the records decades ago. Duo to the global warming, the air temperature and solar radiation has risen significantly in recent years. The climate instability also increases the uncertainties in the climate.

Table 1-1 shows the factors including uncertainties.

Table 1-1: Factors including uncertainties

Building-related	**Operation-related**	**Climate-related**
Simplified computational model	AC setpoints	TMY data
Air tightness	Occupancy period	Climate change
Thermal bridge	Internal heat gain intensity	
Window frames	Natural ventilation	
COP	Lighting	

As a result of lack of information, in the design stage, it is almost impossible to reduce the uncertainties in the simulation inputs. Rather than fill the gap, it is more practical to show the possible gap clearly to the designers. Theoretically, it is not difficult to take uncertainties into consideration in design stage. Monte Carlo method, adapted by many researcher in the uncertainty related works, is a good tool to randomly generate various of scenarios based on reasonable assumption. However, the calculation cost would become thousands of times more than usual. Extremely high calculation cost is a very serious problem.

In this research, we mainly focused on the uncertainties comes with the operation-related factors, for the reason that they have much higher impact on the energy performance, especially the cooling and heating setpoints, than the building-related factors, which has been illustrated in section 5.5 and 6.3. On the other hand, though those building-related factors are uncertain in the early stage, they can be well studied in the later design stages. The influence from the uncertainties in the building-related factors can be eliminated or reduced to some extend with proper design for details. The uncertainties in the TMY climate data and the climate change are also considered in this research, but not the main targets.

Absolutely, even with very precise input information and careful simulation configurations, there are always errors, more or less, in the calculation results cased by simulation tools. However, we were not able to calibrate or quantify the uncertainties caused by simulation tools, neither could designers in the early stage. Therefore, they are not in consideration in this research.

1.1.5 Other problems existing in the early design stage

Besides the uncertainty issues, there are also some other problems in the early stage of architecture design.

In many projects, the design team still adapt the traditional trial-and-error process (figure 1-4) to perform the architecture environmental design. The designers do the environmental analysis, mainly using simulations. Based on the results, the designers adjust their design, then do the analysis again. A lot of times was expended on adjusting the models for simulations and waiting for the calculations to finish. This process is quite low-efficient, resulting in that the design possibilities have been tested are quite limited and the final decisions can hardly be said as optimal. Moreover, in many cases, the designers do not do the analysis themselves but trust to the engineers, which furthermore reduce the efficiency. The communication cost within the design team is not negligible. The engineers need to spend a lot of time to understand the modified design that the designers proposed, while designers spend time to understand the analysis results and suggestions from the engineers. Each time the designers negotiate with the engineers, a lot of time was wasted on understanding each other as the designers and engineers have different education backgrounds and professional knowledge. As it is difficult to execute simulations in meetings, both designers and engineers are lack of evidence to support their ideas.

On the other hand, the communication between the design team and the clients is also another problem. Usually a meeting is needed when the design team try to propose some modifications to the clients in design and get feedback, results in very high communication cost. In the 'pure' architecture design field, parametric modelling, mainly using Rhinoceros + Grasshopper, has been introduced into the industry to reduces communication cost, as the designers can propose a large bunch of cases to the clients at the same time or adjust the design efficiently in the meeting site. However, there have not been a similar and mature solution for the environmental design currently.

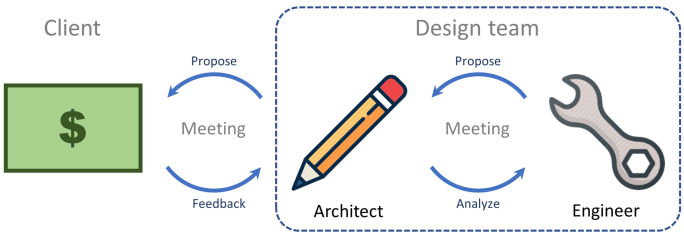


Figure 1-4: The trial and error process

With the popularization of the parametric modeling, optimization algorithms, such as GA[8], PSO[9] and SAT[10], have been introduced into the design process. However, we think that these optimization algorithms might not be suitable for the early stage of design, for 3 reasons. First, though optimization methods can

[8] Genetic algorithms
[9] Particle swarm optimization
[10] Simulated Annealing Terminology

powerfully find the optimal in a specific scenario, as there are uncertainties in both the inputs and outputs if the objective functions, the results might not be optimal in other scenarios. Second, unlike rockets or aeroplane engineering, building industry is not something that asks for high accuracy or extremely optimal solutions, especially in the early stage. We think that, rather than finding optimal solutions, widely exploring the decision space, looking for more design possibilities, is more important in the early stage of architecture design. Third, as there are many factors very difficult to quantify, such as aesthetic or historical issues, many aspects of an architecture cannot be taken into consideration in the optimization process, leading to a results that the optimized solutions are not acceptable to the designers. By the way, the calculation cost of the optimization process is also an obstacle.

Another problem is how to make decisions. The analysis results themselves are not able to tell the designers how to improve the design. The designers need specialized knowledge to understand the analysis results and find better solutions. On the other hand, even with the decision space has been well explored after millions of simulations executed, it is still very difficult for the designers to make decisions when facing a large amount of raw data. How make full use of the analysis results and the simulation data, to support the decision-making by designers in the early stage, is also a valuable topic of study.

Moreover, 'looped works' problems were also found in the current design process. Due to the interactions between parameters, the decisions made in early phases will be influenced by those made in latter phases and become no more optimal. For example, in order to reduce the construction cost, the client decides to replace the low-E glass with normal glass material, then the design of openings on facades optimized with the assumption that low-E glass is used should be restudied from the very beginning. How to analyze the interactions between parameters and reasonably separate the design process into sub-phases, to avoid this kind of 'looped works' is another problem waiting to be solved.

To conclude, the problems existing in the current design process in the early stage are,

- The uncertainty issues are not well considered
- The lack of calculation power
- The low-efficient trial-and-error process
- The high communication cost
- The insufficient exploration of the decision space and design possibility
- The lack of supports for decision-making by utilizing the analysis data
- The 'looped works' due to the interactions between parameters

1.2 Research purpose

The ultimate goal is certainly to fill the gap between the assessed and real energy performance. Limited by the situations and our abilities, this research focuses only on the early design stage, out of the whole the life cycle of a building. The purpose of this research to **find a high-efficient decision-making framework for the early stage of architecture environmental design** that is able to answer these 5 questions at the same time.

1) How to keep the calculation cost at a low level?
2) How to take uncertainties into consideration in the early design stage?
3) How to fully explore the design possibilities?
4) How to utilize the data to support decision-making and benefit communications?
5) How to analyze the interactions and avoid 'looped' work?

Architecture design is a kind of customized service, rather than standard industry mass production. Every project is unique, duo to the site, scale, the wishes of clients, the ideas of designers and so on. As a result, it is very difficult to deduce a universal solution of how to deal with uncertainties, which is clear, practical and quantifiable at the same time. Such universal solutions finally becomes no more than rules of thumb, which have been proved to be useless in extremely-low energy building design. Another way is to try to list as many as situations as possible and make all solutions conditioned with details in a tree structure, originating them into a handbook. However, it is very difficult to list all the possible, the situations are always discrepant, more or less, from the listed ones. This handbook-method is lack flexibility to deal with changing situations in real projects.

Therefore, this research aims at proposing an efficient framework that can help designers find suitable solutions for different projects in specific situations, rather than solutions themselves. In another word, the main target of this research is to propose and demonstrate methodologies.

In this research, we concentrate on how to incorporate the uncertainty issues into the early stage. Though considering uncertainty is the premise of this research, neither uncertainty itself, nor how to model the uncertainty in the inputs of simulations, is included in the purposes of this research.

1.3 literature review

1.3.1 Uncertainty

Many researchers have already noticed the uncertainty issues and taken uncertainties into consideration into their researches on building environmental topics. However, when researching with the keyword 'uncertainty' in the under topics of building, energy and environment, the amount of results is still overall few.

Quantification and decomposition the uncertainties in energy performance or thermal comfort have been studied by some researchers. Harter et al.[69] proposed a method to evaluate the project-specific uncertainty in LCEA[11]. They defined multiple LODs[12] and configured different ranges of parameters for different LODs. Variance-based method was used to quantify the uncertainties caused by different parameters. Chen el al.[44] studied the influence from meteorological, urban, building, system and occupant uncertainties on the thermal comfort in a natural-ventilated Building. A wide range of uncertainties have been considered in their research. Monte Carlo method was used to get the distribution of the thermal comfort and the risk of overheat was checked. Faggianelli et al.[46] studied the uncertainty issues for ESPC. They firstly picked up the parameters including uncertainties, the attached 5%-10% uncertainties to them using normal distribution. Monte Carlo methods was used to get the distribution of energy performance. The authors also compared different sampling methods and sample sizes. Van Gelder, Janssen and Roels[30] separated the parameters into design options and scenarios. For each design option, they used a Monte Carlo loop to get a converged distribution of results.

Optimization with uncertainties was also studied by researchers. Bamdad et al.[54] used 3 scenarios, "low", "base" and "high" and set 3 objective functions correspondingly. The executed multi-objective optimization using ACOMV[13]. Sun et al.[38] did an optimization of system design for ZNE buildings under uncertainties. They introduced uncertainties into physical, building and scenario parameters and assessed the system performance under uncertainties using Monte Carlo simulations. Boundaries of certain confident intervals, based on the threshold set by the user, was used in the optimization. In the field of computer science, the theory of Robust optimization[33] has been well developed. Robust optimization used both the mean and variance of a distribution as objectives, or sometimes uses the boundary of a certain confident interval. The works of Bamdad and Sun are close to the idea of Robust optimization. However, the theory of Robust optimization hasn't been formally introduced into the field of building energy, especially for the design stage.

Other researches also studied the uncertainties caused by different calculation models of water source heat pump[72], as well as the that caused by the computer models used in simulations[29]. The uncertainties in the measurement were also studied.[63][60]. The researches are beyond the interests of this research and my knowledge.

[11] Life cycle energy assessment
[12] Level of development
[13] Ant Colony Optimization for Mixed-Variable

1.3.2 Climate change issues

Besides some crazy presidents, people believe that the climate change is in progress and gradually show the impact on our life. With the air temperature and solar radiation rising, the cooling load will also rises in most area in the world.

Many efforts have been made world widely to predict the development of climate change by researchers. IPCC[14] is the most important organization that contribute to the studies about the climate change and integrate the attributes from researchers all over the world.

In 2000, IPCC established 6 families of scenarios, assuming different green gas emission in the future, in the their report 'Special Report on Emissions Scenarios' (SPES).[6] As shown in table 1-2, the scenario families are divided based on the economy and globalization. 'A' families means putting more emphasis in economy development, while 'B' mean considering more about the environmental issues. '1' means globalization and '2' means regionalization. Therefore, 'B1' is the best environment-friendly scenario, in which the assessed warming is $1.1 - 2.9°C$. Conversely, from the aspect of environment protection, the 'A2' can be said as the worst scenario, which this world is now heading to. 'A1' includes 3 sub-scenarios. The difference between them are the energy source balance. In the 'A1FI', a fossil Intensive scenario, the world relies on the fossil energy sources. 'A1B' means a balanced scenario that uses all energy sources. 'A1T' mean a more environment-friendly scenario that puts more emphasis on non-fossil energy sources.

Table 1-2: SPES families

	More economic focus	**More environmental focus**
Globalization (homogeneous world)	A1FI, A1T, A1B rapid economic growth $1.4 - 6.4°C$	B1 global environmental sustainability $1.1 - 2.9°C$
Regionalization (heterogeneous world)	A2 regionally oriented economic development $2.0 - 5.4°C$	B2 local environmental sustainability $1.4 - 3.8°C$

IPCC also support the research institutes all over the world with simulating the climate change using GCMs[15] and RCMs[16] under all SPESs. The models, such as the widely used HadCM3 from Hadley Centre, UK and CGCM2.3.2 from Meteorological Research Institute, Japan, and their simulated results were collected and published by IPCC.

Based on these simulation results, many tools that can generate weather files for the future have been developed. Climate Change World Weather Generator (CCWorldWeatherGen)[9] is an MS Excel based free

[14] Intergovernmental Panel on Climate Change
[15] General Circulation Models
[16] Regional Circulation Models

tool that utilizes the results of HadCM3[80] to morph a EPW file and generate a new EPW file for the future climate. Meteonorm[73] is a commercial software that runs on MS Windows. It use the average values from several models available in IPCC AR4[75] to generate new EPW files. Several scenario can be selected in Meteonorm. WeatherShift is an online tool that gets support from USGBC. The data from WeatherShift is very expensive.

In the IPCC AR5, IPCC established a new set of scenarios named Representative Concentration Pathways (RCPs), which superseded the SRESs. A RCP is a time series of the variability of the GHG emission and density in the future. RPCs includs RCP 1.9, 2.6, 3.4, 4.5, 6, 7, 8.5. The higher the index is, the worse the scenario is. Researchers have gradually converted their works from SRESs to RCPs.

Most researchers used weather files generated with IPCC data in their researches. But some researchers looked for their individual ways to predict the climate change. Zhu el at.[42] analyzed the historical climate records in Shanghai with the long- /short-term climate periodicity analysis, then proposed a Dual-Periodic Time Series Model is established to predict the future monthly temperatures in Shanghai. They also compared the predicted results with the observations and the GCM outputs under RCP 4.5. Their results were much closer to the observations than the GCM outputs. Park et al.[50] trained a LSTM[17] network, a kind of RNN[18], with the historical data of 36 years recorded hourly, to predict the climate change.

In building energy field, researchers concentrates more on the influence of the climate changes on the cooling and heating loads. Wang el at.[52] generate the EPW files of 2020, 2050 and 2080, using two IPCC climate models, HadCM3 under A2 scenario and CESM under RCP 2.6, 4.5 and 8.5. They did simulations of a office building using *EnergyPlus* with these weather file and discuss the variation whole building energy source use. Rey-Hernández el at.[58] discussed the impact of climate change on a constructed zero-energy zero-carbon building in Spain. They generate the weather files of 2020, 2050 and 2080 using CCWorld-WeatherGen. An EnergyPlus model of the objective building was made and simulated with these weather files. They concluded that, due to the climate change, this zero-energy zero-carbon building will become no more zero.

However, there is still a gap from these researches to our purposes, as we aim at benefiting the architecture design with our researches. The research of Chen[66] from Maelab is good example to illustrate our ideas. In her master thesis, Chen generated 5000 cases using Monte Carlo and simulated them under climate of 2020 to 2080 with a 10-year-wide step. She compared the top 5% cases under different climate, as well as the top 5% the cumulative results. It has been found that, the top cases in 2080, compared to those in 2020, have larger shadings but less insulation. This research studied the influence of the climate change on the optimal of building parameters, which can be guidelines in the design stage.

[17] long short-term memory
[18] recurrent neural network

1.3.3 Sensitivity and interaction analysis

In the last decades, the statistic researchers developed a lot of sensitivity analysis skills, which can be roughly divided into 3 categories, OFAT[19], variance based and linear regression based. The representative methods of these 3 types are respectively Morris method, Sobol indices and multi-variable linear regression. All these 3 types of sensitivity analysis methods have been introduced into the field of building energy by researchers, to quantify the impact of different parameters on the energy performance, and then reduce the number of parameters and calculation cost.

Chen et al.[44] introduced MARS[20], a kind of linear regression based method, into their research to identify the sensitivities of parameters. The weight coefficient of each parameter in the regressed formula indicated its sensitivity. Heiselberg et al.[11] introduced Morris method into the design of sustainable buildings. They screened the parameters by their importance and dropped less important parameters in the latter stage. They commented that sensitivity analysis should be performed in the early stage when important factors were not decided. Østergård et al.[49] described an approach to explore the multi-dimensional design space. They used Morris method to screen parameters. The importance of each parameter helped designers to make decisions in the latter stage. Nguyen et al.[36], Rivalin et al.[59] and Ganon et al.[56] compared several sensitivity analysis technics applied to building energy performance assessment. Linear regression was reported not good enough by most researchers. Morris method showed both good efficiency and enough reliability. Variance-based methods, though had the best reliability, was very calculation costing. In this research, we tested both Sobol indices and Morris method, but not the linear regression as it cannot be further utilized in interaction analysis, in chapter 4.

Interaction analysis is actually higher-order sensitivity analysis. The development of the skills of interaction analysis is actually quite mature. However, though researchers have touched the higher-order analysis, the interaction analysis has been introduced into the research field of building energy. Garcia et al.[28] analyzed the sensitivities of 24 parameters over the energy performance of an office building using Morris method and *EnergyPlus* simulations. Interaction analysis, based on the second-order effect, was also performed using Expanded Morris method. They commented that higher-order effects could help better understanding the results of sensitivity analysis. Regretfully, They did not go a further step into the utilization of the interactions between parameters.

As sensitivity analysis is a kind of statistic technic, the sampling method is very important. Some researchers in the field of building energy also compared different sampling method. Faggianelli et al.[46] performance uncertainty analysis using Morris method. In their research, they also studied the influence from sampling method and sample size. They results showed that, in the case that the sample size was small, LHS outperformed the common Monte Carlo in Morris method analysis. When the sample size passed 2000, in their case study, LHS and common Monte Carlo method showed no essential differences. We also talked about the sampling method in section 4.2.

[19] One factor at one time

[20] Multivariate adaptive regression spline

1.3.4 Meta-modeling

Making meta-models, or regression in another word, based on the simulation data, is a very effective way to reduce calculation cost. Recently, meta-modeling is sometimes also called machine learning in some contexts. But in an academic context, we think it is better to simply say meta-modeling. A trained meta-model can predict thousands of results in a moment. Replacing simulations with trained meta-models can accelerate the whole process dramatically. Meta-modeling has certainly been introduced into the researches in building energy field. Some researchers trained meta-models before sensitivity analysis and then get the data needed to perform sensitivity analysis. Some other researchers performed sensitivity analysis first and did meta-modeling with screened parameters. Meta-modeling was also used by some researchers as a solution to reduce time cost of optimization processes.

Hygh et al.[17] presented a Monte Carlo framework to developed a multivariate linear regression model based on 27 parameters. The coefficient of each parameter in the regressed model could be used as the sensitivity. Østergård et al.[49] also made metamodels with multivariate linear regression. With this meta-model, they made a 'what-if' dashboard that can give real-time feedback to the designer. Van Gelder, Janssen and Roels[30] used MARS to make meta-models, which was also used as sensitivity analysis method by other researchers. Gossard et al.[22] trained an ANN[21] with the simulation results of energy performance of the building. They used this trained ANN as a part of objective function in GA to accelerate the optimization process. Asadi et al.[25] also combined ANN and GA in their optimization process and applied them in retrofit projects. Faggianelli et al.[46] used Polynomial chaos expansion as the meta-modeling method in their research of quantification of the uncertainties.

Rivalin et al.[59], Wei et al.[39] and Østergård et al.[57] compared several techs of meta-modeling applied to building energy performance assessment. Gaussian process was reported to be the most robust and easy-to use, neural network and MARS also had good performance. In some cases, Polynomial chaos showed better accuracy than Gaussian process did.

1.3.5 Assistance for decision-making

To be honest, we do not think that decision making in the early design stage is a research topic, as it cannot be concluded with into theorems or practical paradigms. 'Decision-making' itself is a kind of art. As researchers, what we can do is to look for better ways to assist designer when they make decisions, rather than tell them how to. The literature on the topic of decision-making in early design stage is quite little. But still some efforts have been done by researchers who have knowledge of both architecture design and environmental engineering.

Schlueter and Thesseling[13] developed a plug-in for Autodesk Revit called Design Performance Viewer (figure 1-5), which could evaluate the energy performance of the building, using the information from BIM.

[21] Artificial nerual network

Once the designers modified the model, they could get nearly real-time feedbacks. Though modeling and analyzing was integrated, designers still had to work in the trial-and-error process. It was difficult to compare a lot of alternatives simultaneously.

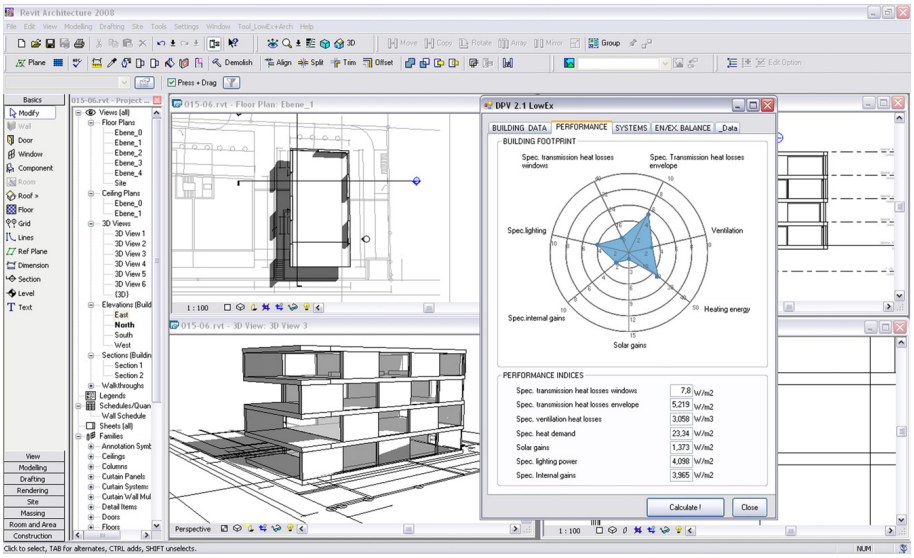


Figure 1-5: Modeller with tool interface of Design Performance Viewer (Schlueter and Thesseling | 2009)

With meta-modeling trained using MARS, Østergård et al.[49] made a 'what-if' dashboard (figure 1-6) to help the designers making decisions. Designers can input their ideas about how to modefy the desing, like 'if we enhance window-frac-ratio by 10%', and get the real-times feedback about the variations in sensitivities and outputs of energy performance, thermal comfort and daylight.

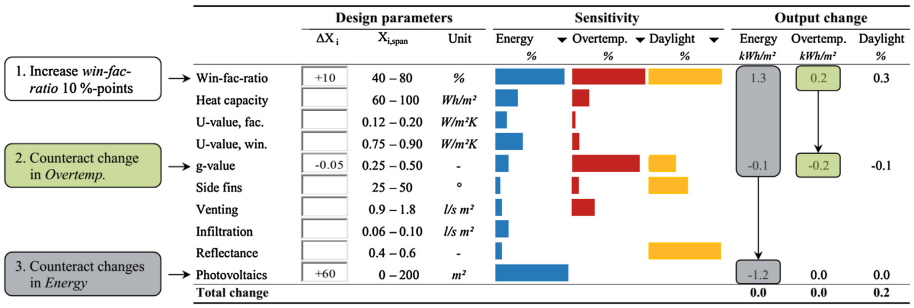


Figure 1-6: 'What-if' dashboard (Østergård et al. | 2009)

From the researches above we could find that dashboards, able to visualize the building parameters and the energy performance, are usable when designers make design decisions. Therefore, we also studied the dashboard tools currently available.

Insight 360[31] (figure 1-7) is a building energy performance assessment tool published by Autodesk. This designer-oriented dashboard tool is intuitive and easy to use. Good data visualization helps the designer understand the energy performances well. Parameters can be modified, and the feedback is real-time. However, there are still some limitations. The parameters and their range that can be studied are limited to the preset. Only the average value of energy consumption of all case is displayed. Other criteria, such as thermal comfort, are not included or extendable.

Designer Explorer 2[64] (figure 1-8) is another tool under the concept of design dashboard. It consists of 2 parts, Grasshopper3D components and online data visualization. Designers can do the analysis in Grasshopper3D and output the results into a CSV file. The online data visualization tool visualizes the data and makes dashboards once this CSV file is uploaded. This process takes quite a long time to execute the simulations if the designer wants to explore the possibilities of design widely.

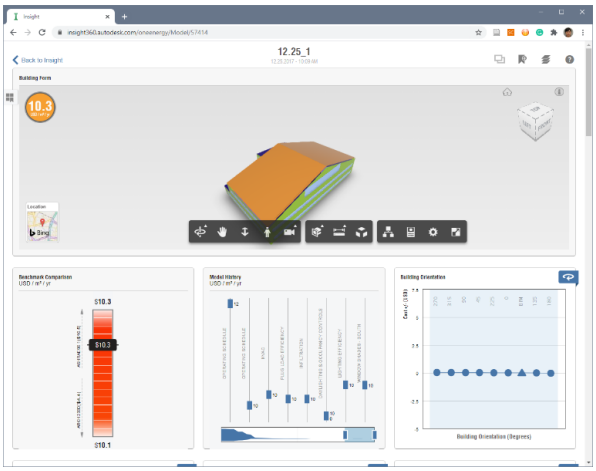


Figure 1-7: Autodesk Insight 360
(Autodesk | 2019)

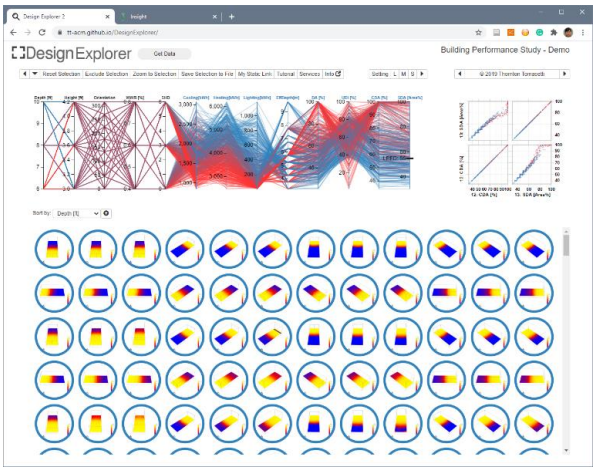


Figure 1-8: Design Explorer 2
(CORE studio | 2019)

Decision-making with uncertainty considered has been discussed by researchers. Van Gelder, Janssen and Roels[30] proposed and demonstrated a Probabilistic design process for building performance that considering uncertainties in both design options and scenarios. MARS was used to building the meta-models. Sensitivity analysis is used to screen the parameters. 640 design options was generated with the left parameters. Each design option was tested in Monte Carlo loops until the distribution of output converged. The effectiveness and robustness of all the design options were compared. The method proposed in this research is somehow close to our targets. But, even with meta-models, it still takes times to get converged distribution for each design option. On the other hand, 640 options are too few for the early design stage.

1.3.6 BIM based tools

We have made several interviews with professional architects. They said that they prefer to stick to BIM tools from the very beginning, even though they know that Grasshopper 3D might be more convenient, so that they can push the work more smoothly to latter stage. Several BIM based tools, aiming at the early design stage, have been developed by researchers. The DPV has by Schlueter and Thesseling[13] has already been introduced in the last subsection. Asl et al.[37] integrated BIM and multi-objective optimization. They extracted the information from the BIM model to carry out energy, daylight and structure analysis. The whole process was highly integrated, nevertheless, the calculation was still very costly. Jalaei1 and Jrade[34] integrated BIM and LEED certification system for the conceptual stage. This integrated tool could generate LEED certification point using the information from BIM and help the designer select proper materials. However, LEED certification required a highly completed design. A lot of information necessary for LEED was lacked in the conceptual stage. As a result, it was still a question that whether LEED provides good criteria in the early stage.

1.3.7 A brief summary of the state of the art

Through the literature review, we found that, with proper assumed distributions of the parameters, Monte Carlo method is usually used to reflect the uncertainties in parameters. Using multiple sets of parameters generated randomly, researcher can get the distributions of criteria. However how to model the uncertainties is still not well-known or well-studied. Variance-based methods, such as Sobol indices, have been used as uncertainty analysis method to decompose the uncertainties in results by parameters. Some researchers have tried to introduces uncertainty issues into optimization process, by configuring several scenarios or using special assessment and screening strategies.

The methods of sensitivity analysis are no different from those of uncertainty analysis, besides the linear regression based ones. Sensitivity analysis has been introduced into the environmental design and proved to be effective. Interaction analysis has been touched, but further utilization of the interactions between parameters has not yet been made.

Several kinds of meta-modeling methods have been proved practical in environmental design. However, the characteristic of Gaussian process that predicts the variances of outputs has not been well utilized.

Dashboards would be good tools that helps designers make decisions, while the benefits from BIM are quite limited in the early stage.

The state of the art is that knowledge and technics from computer science and statistic have been introduced into building energy field, but still fragmented. A framework is necessary to integrate these knowledge and technics to help the designers make decisions in the early stage of architecture environmental design.

1.4 The composition of the thesis

This thesis consists of 7 Chapters.

In the 1^{st} chapter, the background of this research is introduced, including the uncertainty issues and other problems existing in the architecture design industry, followed by the research purpose which is to propose an efficient framework that us able to solve these problems. Literature review has been carried out to see what has been done by researchers and what is insufficient in the state of art. The 2^{nd} chapter introduced the proposed decision-making framework proposed for the early stage of environmental design, which integrates sensitivity and interaction analysis, meta-modeling using GPR, as well as interactive data visualization.

In order to mass produce simulation models and execute simulations efficiently, we developed parametric modeling tools called, *EPPiX* for *EnergyPlus* and GH2FD for *FlowDesigner*. Besides, we also developed a tool called *Ultimate EP Executor* that is able to make full use of the CPU power and run *EnergyPlus* simulations in parallel. The system designs of these tools are described in Chapter 3 with UML[22]

Chapter 4 introduces the statistic technics used in this research. This chapter begins with explaining the importance of reducing the number parameters. Different sampling methods was described and compared. The following sections introduced Sobol indices and Morris method of sensitivity and analysis, and their expansions for interaction analysis, Sobol second-order indices and Expanded Morris method. All these methods are demonstrated and compared. A original mathematical method called PCP is also introduced.

The regression technics used in this research is introduced in Chapter 5. We firstly introduced 2 models derived from R^2, which are specialized for data analysis with uncertainties. Following, the principles of Gaussian process regression, the characteristic of GPR dealing with uncertain data was explained and demonstrated. To perform decent training in the situation that the uncertainties are vary high, we proposed a original sampling and training method called STMN. The covariance between criteria and coregionalized GPR are also briefly introduced in the last section.

A demonstration of the application of the proposed framework in early design stage is made in Chapter 6, with an imaginary office building located in Tokyo. After sensitivity and interaction analysis, the number of parameters is largely reduced. A database is generated with predictions from trained GPR models. Dashboards are made to interactive visualize the data in the database.

In the last chapter, we made conclusions of the achievements and limitations of this research, as well as the suggestions of further work.

[22]Unified Modeling Language

Chapter 2

The decision making framework

The 2^{nd} chapter begins with the mode of thought issues. The concept of *Antifragile*, a robust way of decision making, is explained with the example of *Go* game. We also suggest designers to avoid 'premature optimization' in the early stage of design. A pair of new introduced concept, targeted and non-targeted parameters, is explained in the second section. The introduction to the decision-making framework is proposed in the third section. The purpose and technologies of each step in the framework are illustrated with details.

2.1 Modes of thought

Before proposing the framework, I would like to make an introduction to the modes of thought that we suggest designer to hold when facing uncertainties in the early stage of architecture environmental design.

2.1.1 Optimized framework, not optimized results, in the early stage

'Premature optimization is the root of all evil' is a famous saying by Donald Knuth, a computer scientist from Stanford university, in his paper *Structured Programming with go to Statements*[1] published in 1972. Prof. Knuth said 'We should forget about small efficiencies, say about 97% of the time. Yet we should not pass up our opportunities in that critical 3%.', which mean that we should concentrate in the critical part but do not waste time on studying details of other parts. Some researchers explain Knuth's words as '**optimizing before we know that we need to**'[26], as in the early stage we can hardly correctly predict which part is critical or needs optimization.

The problem is how can we know which part belongs to the critical 3%? Experienced specialists could roughly judge which part is important with their experience. However, as we explained in the last chapter, the architecture environmental issues are so complicated, as an architecture is strongly influenced from site, climate, occupancy, etc., which are very dynamic. It is very difficult to confirm the importance of each parameter by rules of thumbs. On the other hand, the interactions between parameters could be non-predictable, which could also cause troubles. In the early stage of architecture environmental design, in order to find the critical 3%, we suggest to:

- Explore as many design possibilities as possible, (do full-set exploration)
- Quantify the sensitivities of parameters,
- Quantify the interactions between parameters.

As a result, an optimized and framework is needed to efficiently explore the design possibilities and analysis sensitivities and interactions.

Though Prof. Knuth talked about the software engineering, these words also inspires the architecture environmental design in early stage. By the way, there is actually an interesting truth that software engineering has learned a lot from architecture design. The establishment of the system design theory in the software engineering filed is inspired by the pattern language[cite], which is initially introduced in 1977 by Christopher Alexander, a mathematician teaching in architecture department of U.C. Berkeley, in his book *A Pattern Language: Towns, Buildings, Construction*[2].

2.1.2 Antifragile, a robust way of decision making

Briefly speaking, when facing uncertainties, rather than finding a best solution in a certain specific circumstance, trying to avoid failures under different conditions could be a better way.

Antifragile is a word invented by Nassim Nicholas Taleb, a financial researcher, in his book *Antifragile: Things That Gain from Disorder*[20]. Financial investment is an good example, as the investor faces huge uncertainties, even the possibility of a black swan event, when making decisions. The core concept of this book can be concluded as 'Try to evaluate the risks in different situations. Diversify investments into different sectors. Even in the case that the worst situation occurred in a sector, the lose can be covered by the profits from other sectors.'

Go[1] game is another inspiring example of decision making under uncertainties. There are countless possibilities in a game flow. The value of a step varies, especially according to how the opponent plays, during the game flow. In the tradition of *Go* game in Japan, *Honte*[2] is a very important concept. Unlike the quick games in China and Korea, there was no time limitation in games during the Edo period. Compared to the aggressive strategies preferred by Chinese and Korean players, most Japanese players would spend long hours on thinking as many situations as possible and evaluating the risks caused by this step in latter phases, to avoid showing weak point to the opponent. Robust decisions is preferred by Japanese players, which gradually became the concept of *Honte*. Takemiya 9dan[3], the first international champion in the world, described *Honte* as 'a step without worries in latter phases'.[4]

In the environmental design, a designer should also consider different situations to make sure the building perform decently in most cases. However, over-conservative strategies lead to high construction cost, boring building space and mediocre energy performance. Showing the probability distribution of all possible results could be a key solution to this issue. So that the decision maker can weigh the profits and risks.

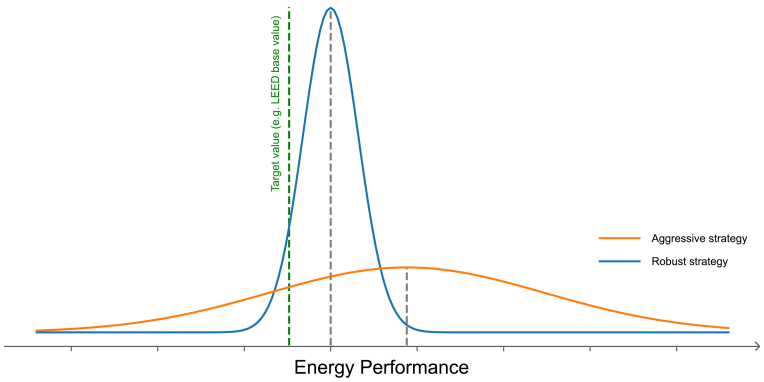


Figure 2-1: Schematic distributions of different strategies

[1] 碁，囲碁
[2] 本手
[3] 武宮正樹九段

2.2 Targeted parameters and non-targeted parameters

In order to make this research more understandable, we defined 2 pairs concepts, **Targeted parameters** and **Non-targeted parameters**, **Variety** and **Fluctuation**.

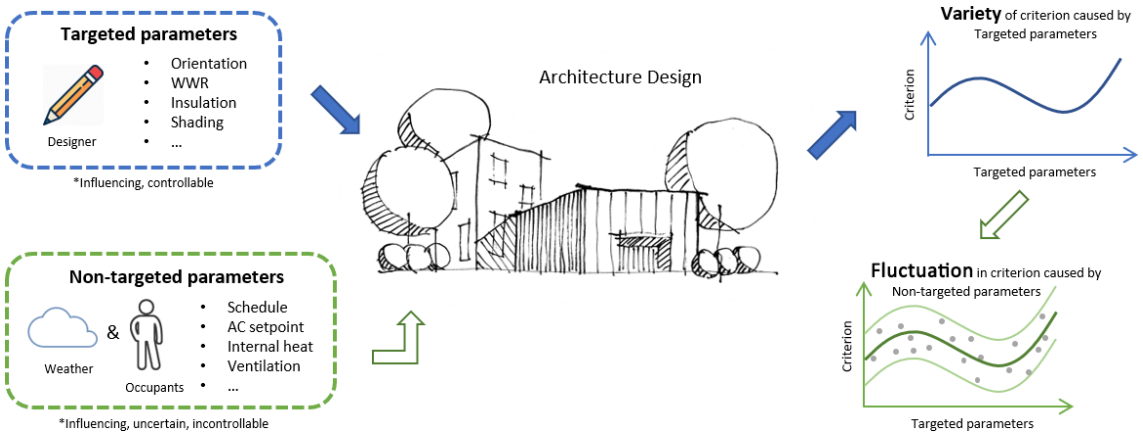


Figure 2-2: Target and Non-targeted parameter, Variety and Fluctuation in criterion

2.2.1 Target parameters and Variety

Target parameters (TP), which are actually an alternative way to say *parameters of interest*, means those parameters that designers want to study in the design stage. These parameters should have influence on the energy performance. In the meantime, they should be able to be decided or controlled by designers. In the early stage of environmental design, TPs are mainly building-related, which could be window to wall ratio (WWR), insulation, etc. Admittedly, there are also gaps between the designed and measured values of building-related parameters, so that there are also uncertainties. However, these uncertainties are almost negligible for 2 reasons. First, the uncertainties in building-related parameters are much smaller than those in occupancy-related and climate-related parameters. Second, the sensitivities of building-related parameters over energy performance are much smaller than those of operation-related and climate-related parameters, which can be found from section 6.3. Therefore, the influence of the uncertainties in building-related parameters on energy performance extremely small. Therefore, the uncertainties in building-related parameters are not considered in this research.

The energy performance and thermal comfort will certainly change with the values of TPs. In this research, we call the changing of a criterion caused by the changing of TPs '**Variety**', which is actually what we want to study in the design stage.

2.2.2 Non-target parameters and Fluctuation

There are also many other parameters influence the energy performance and thermal environment, but their values cannot be confirmed in the early stage. Meanwhile, the designers can hardly decide or control the values of these parameters, such as climate, operation, etc. Uncertainties of these parameters will propagate in energy performance calculation[15], and introduce fluctuation into the analysis results. Theses uncertain parameters are called **Non-targeted parameters (NTP)** in this research. They should also be included in the assessment, as well as their sensitivities over the energy performance and interactions with the TPs.

Monte Carlo method is usually used to reproduce the uncertainties of the NTPs. As their values are randomly decided, the calculation results of a criterion will also be unpredictably fluctuated. In this research, the word 'uncertainty' is defined as 'the characteristic of a NTP that its real value cannot be precisely predicted', while the variability of a criterion caused by the uncertainties of NTPs is called '**Fluctuation**'. Though the 'Fluctuation' is not a target in design stage, designers should also consider it to make robust decisions.

2.2.3 Wording issues

Table 2-1 is a brief conclusion of these 2 pairs of concepts.

Table 2-1: Target and Non-targeted parameters, Variety and Fluctuation

Concept	Definition
Targeted parameter (TP)	Parameters that we want to study
Non-target parameter (NTP)	Parameters that are not study targets but should be considered
Variety	The changing of a criterion caused by the changing of TPs
Fluctuation	The variability of a criterion caused by the uncertainties of NTPs

In this research, we mainly used the words targeted and non-targeted parameters. Besides, we mainly used the words **criteria** to express the criteria in environmental design, like cooling/heating load, as well as their values. In the meantime, we also used other ways of wording according to the context, to make this congruent. Table 2-2 is a summary of wording in this thesis.

Table 2-2: Wording for same concepts in different contexts

Context	Concept		
This research	Targeted parameter	Non-target parameter	Criterion
Common	Parameter of interest	?	Criterion
Simulation	Input	Input/Scenario	Output/Result
Other	Design option	Scenario	Result
Statistic	Sample		Result
Function	Independent variable		Dependent variable
Calculation	Parameters/Factors/Variables		Result
Machine learning	Feature		Target

2.3 The proposed framework

In this research, we proposed a decision-making framework with uncertainty considered, for the early stage of design. Figure 2-3 illustrates the flow chart of the framework. As illustrated in the figure, this framework consists of four phases,

- 1. Pre-process
- 2. Dimension reduction
- 3. Meta-modeling
- 4. Post-process

2.3.1 Pre-process

The purpose of the pre-process is to **decide what to study**.

A design team, including architects, engineers and specialists from other disciplines, firstly set the target of the project, decide the types of criteria, like cooling/heating load, thermal comfort, natural lighting, etc. Then, according to the criteria, as well as the site and climate, the design team do the conceptual design. The overall building mass, the master plan, as well as the most important thing, the environmental design strategies, should be decided before this step.

Though the types of parameters influencing the energy performance are similar in different projects, they still vary according to the environmental design strategies adopted by the designers. The design team, including architects, engineers and specialists from other disciplines, firstly decide the TPs to study quantitatively

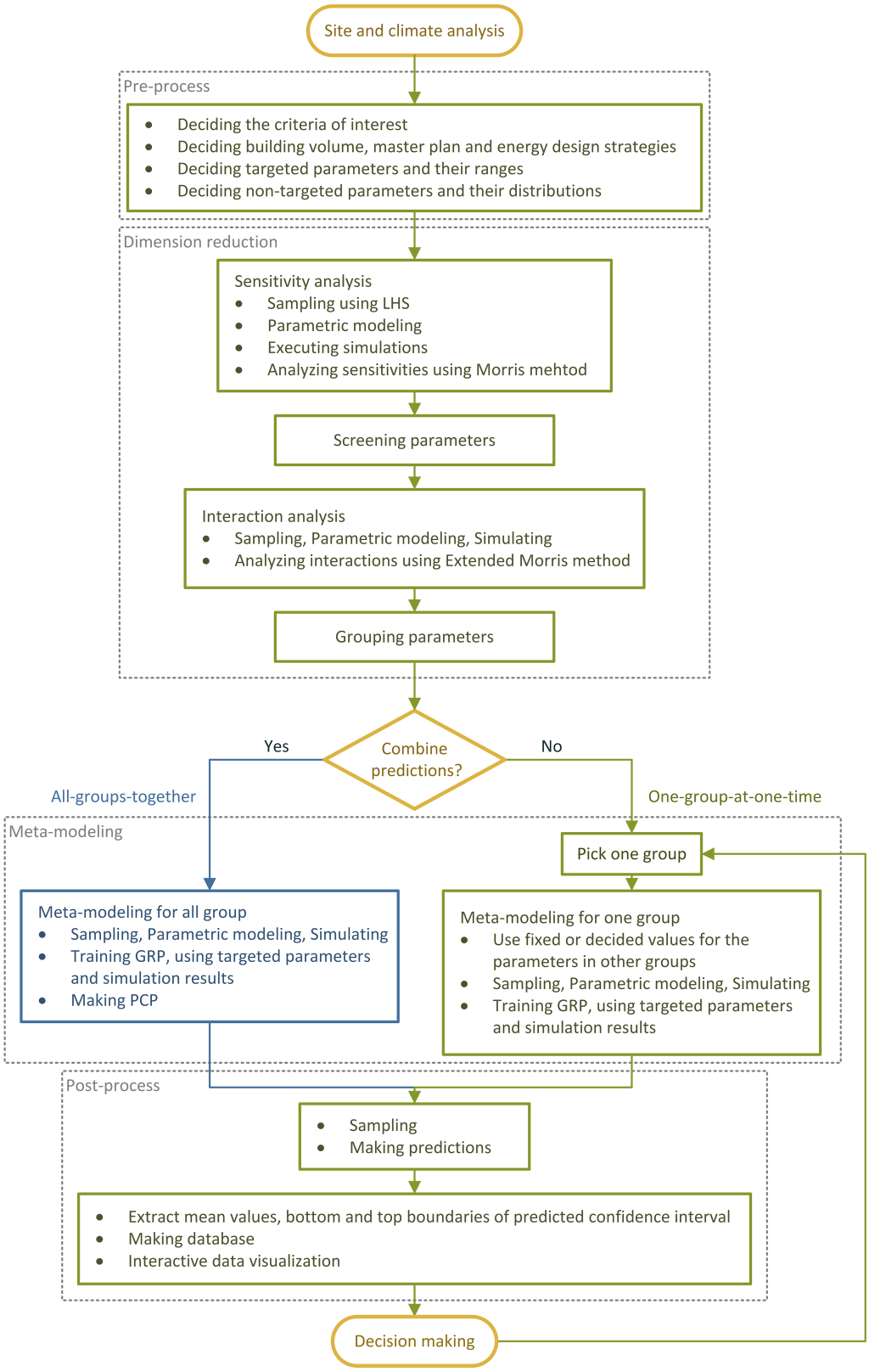


Figure 2-3: Framework for the design stage

(e.g. WWRs, overhangs), as well as their ranges, based on their environmental design strategies. The distribution of the TPs can simply be uniform distribution. The building environment specialists then list the NTPs, mainly occupant-related ones, and their distributions, based on their knowledge and experiences.

2.3.2 Dimension reduction

In the early stage of architecture environmental design, the number of parameters (dimension d), including both TPs and NTPs, are usually dozens. No matter what kind of sampling method (section 4.2) is used, the necessary sample size to perform full-set exploration is usually from 5^d to 20^d. 5^d is the minimum size suggested by many researchers. As the energy performance of a building does not change dramatically with the parameters, 10^d would be enough in environmental design. Additionally, in the case that uncertainties of the NTPs are considered, hundreds times of simulations are needed to indicate the distribution of energy performance of one sample, resulting in that the time cost rises by hundreds of times. It will take more than trillions of years[4] to run simulations, which is impracticable in design stage.

On the other hand, as the factors in architecture environmental design are so complicated, designers can hardly precisely judge the importance of each parameter. Meanwhile, there could be unexpected interactions between parameters, which would leads to troubles in latter stage.

In the dimension reduction phase, the sensitivities of all parameters, including both both TPs and NTPs, will be quantified, as well as their interactions. The purpose of dimension reduction phase is to:

- By analyzing the sensitivities of parameters,
  - Less important can be ignored, dimension can be reduced,
  - Designers is able to concentrate on important parameters,
  - Designers can be more free with less important parameters.
- By analyzing the interactions between parameters,
  - Parameters can be separated into groups, so that a high dimension problem can be decomposed into several low-dimension problems.
  - Designers is able to study interactive parameters together to avoid 'looped' works,
  - Designers can known which TPs are less influenced by NTPs, so that they can make relative aggressive decision on them.

With less important parameters ignored and parameters grouped, the number of parameters (in each group) is dramatically reduced. The volume of full-set, as well as the necessary sample size to perform a regression,

[4] 20 TPs, 10 levels for each, 100 samples of NTPs for each sample of TPs, 3s to run simulations once, 24 threads in parallel: $10^{20} \times 100 \times 3 \div 24 = 1.25 \times 10^{21}$ seconds $\approx 3.96 \times 10^{13}$ years

drops down by thousands of times (figure 2-4). The time cost of simulations is also correspondingly reduced. Meanwhile, informed with sensitivities and interactions, designers can better make decisions.

In this research, Morris method (section 4.3.2) is used to analyze the sensitivities of all parameters, while the interactions between the left parameters are quantified using Expanded Morris method (section 4.4.2). There are no essential differences between Morris method and traditional design of experiments using orthogonal array. Morris method is actually a kind of design of experiments using Monte Carlo method, which is very efficient in high-dimensional analysis.

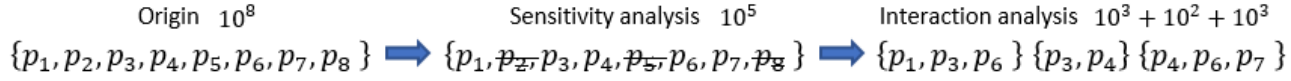


Figure 2-4: Dimension reduced after sensitivity and interaction analysis (schematic image)

2.3.3 Meta-modeling

Though in after the dimension-reduction phase, the number of parameters has been largely reduced, the time cost of simulations just drop from 'trillions of years' to 'several days'[5], which is still impracticable. A much more efficient method is necessary, to take place of simulations. Meta-modeling, or regression in another word, is one of the best choices.

The purpose of meta-modeling phase is to, by training meta-models equivalent to simulations,

- Replace simulations with trained meta-models and dramatically reduced time cost,
- Draw large amount of data from trained meta-models to realize full-set exploration.

From literature review, we have known that some researchers combined stratified sampling and meta-modeling to study uncertainty issues. However, it still takes several seconds to minutes to get the distribution of a criterion of a sample of TPs. The time cost is still a problem in full-set exploration. In this research, **Gaussian process regression (GPR)** (section 5.3.1), which is robust to deal with uncertain data, is used to make meta-models. A trained GPR model is able to predict the distribution of a criterion of a sample of TPs in no time. So that time cost of prediction is also reduced largely.

For each group, the parameters in that group are sampled randomly using Latin Hypercube Sampling (LHS, section 4.2.3), while fixed values are assigned to the other parameters. Simulations are then executed to get the values of criteria. TPs and calculated criteria are included in the training data. In cases that the Fluctuation in training data is big, STMN sampling method (section 5.6) should be adopted.

[5] 4 groups, 5 TPs in each group, 10 levels for each, 100 samples of NTPs for each sample of TPs, 3s to run simulations once, 24 threads in parallel: $4 \times 10^5 \times 100 \times 3 \div 24 = 5 \times 10^6$ seconds $\approx$ 60 days

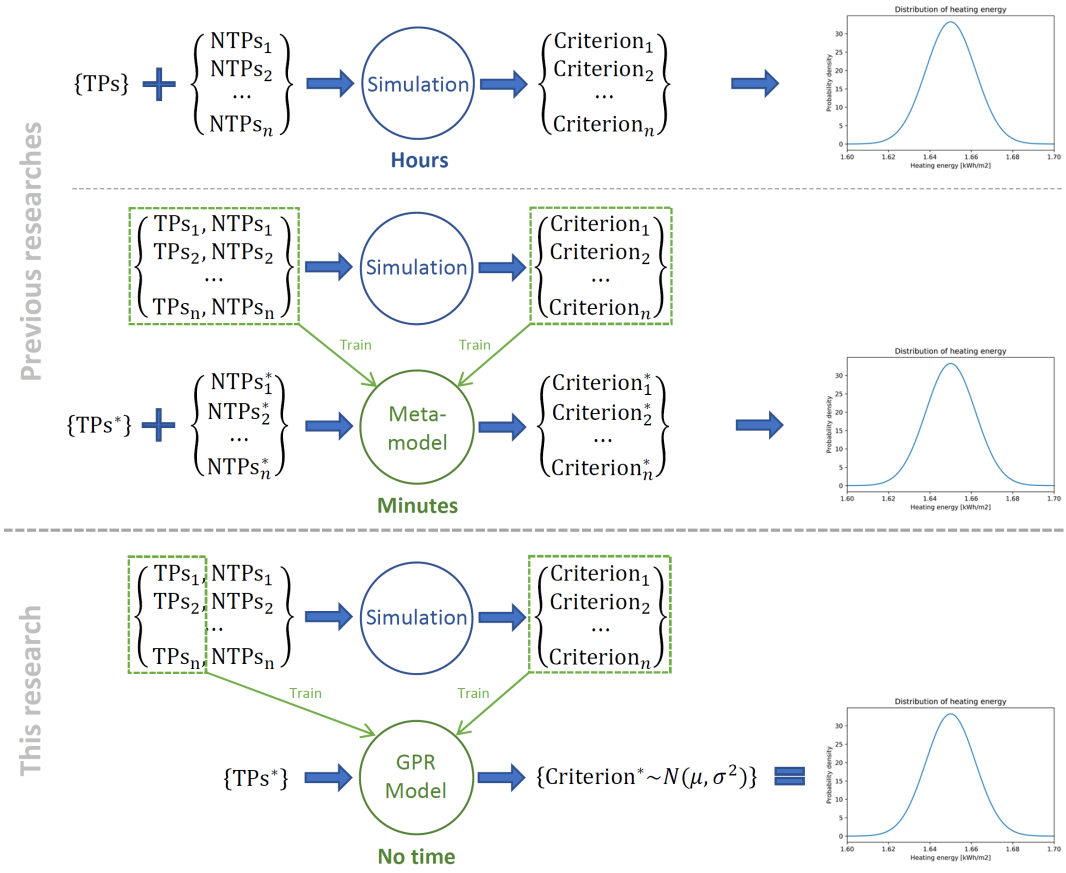


Figure 2-5: Comparison with previous researches (schematic image)

These is a question that should be pointed out. In most cases, the criteria are not in normal distributions, while the predictions made by GPR models are normal distributions. The purpose of considering uncertainty in this research is to inform designers the risk of underestimated energy consumption in the early stage. In this context, normal distributions would be enough to help designers make decisions.

Technically speaking, analyzing groups one by one separately dose not lead to any errors in the final results. However, there might be obstacles for designers, clients and those not familiar with interaction principle to understand the independent results of different groups. For that reason, an additional step is adopted in this research. A formula, which is called Prediction Combining Polynomial (PCP, section 4.6), is used to combine predictions of the same criterion from meta-models of different groups.

Therefore, we propose 2 sub-processes in the meta-modeling phase. If the members in the design team understand the interaction principle well, they can study the parameters group by group, which is called one-group-one-time sub-process in this research. The design team will firstly pick one group to train the meta-model, and then moves to the post-process phase. After decisions made, they move back and pick another group. The design team can use decided values for already-studied groups. On the other hand, in the cases that some members are not familiar with interaction principle, especially in the situation that faces the clients, another sub-process is proposed, which is called all-groups-together. The meta-models of all groups

should be made in the meta-modeling phase, as well as the PCP.

2.3.4 Post-process

Users can input the values of parameters that they want to test into the trained meta-models and get predicted results in a moment. However, even with millions of cases predicted, it is still impossibly difficult for users to read and understand the raw data then make decisions. Interactive data visualization is a key point in the post-process. The purpose of post-process is to, **by visualizing data interactively, help designers make decisions**. 2 sub-processes of data visualization, 'Real-time rendering' and 'Database retrieving', were proposed.

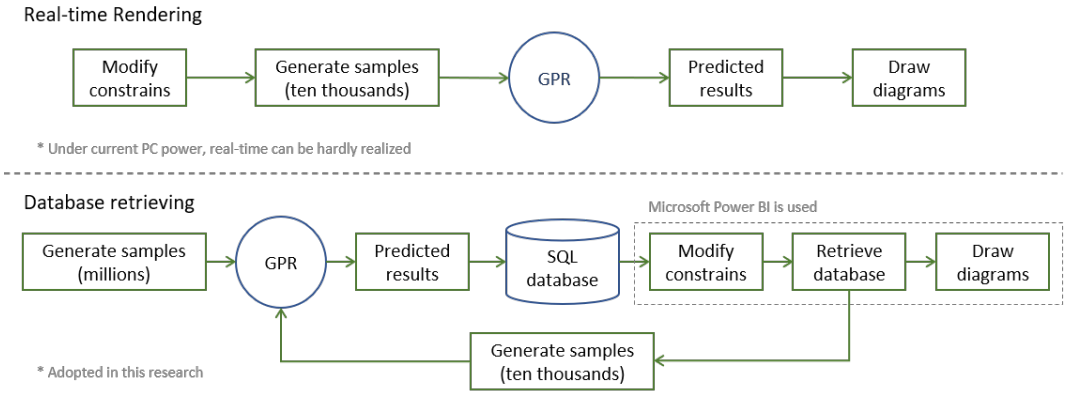


Figure 2-6: 'Real-time rendering and 'Database retrieving' sub-processes

In the 'Real-time rendering' sub-process, each time designers modify the constrains of TPs. Ten thousands of samples of TPs are generated using LHS, the corresponding values of criteria are predicted by GPR models. Then these predicted values are visualized. However, under current power of PCs and mobile terminals, it can hardly be real-time to go through from 'modify constrain' to 'draw diagrams'. This sub-process will be non-smooth. So we planned to try this sub-process in future works.

In this research, we adopted the 'Database retrieving' sub-process. Therefore, a database including a large mount prediction data is needed. The first step is to make millions of samples using LHS and input them into the trained meta-models. The predicted results, along with the samples, are stored in a SQL database. The next step is to make design dashboard. In this research, Microsoft Power BI is used. The design team can control the ranges of targeted parameters using sliders, modify the constrains. Power BI will retrieve the database correspondingly and display the predicted results in diagrams (e.g. histograms, pie chart) in real time. So that, using this dashboard, the design team can test different ideas and make decisions.

Additionally, there is another important reason that we chose dashboards as the main method in the post-process. There are quite a lot of factors that cannot be evaluated quantitatively in architecture design, like aesthetic problems. It's almost impossible to have computer make decisions decently on these problems for designers. Instead, the purpose of this research is to help designers make decisions. With dashboard, designers can adjust the parameters, watch the variety of criteria and evaluate those non-quantitable by themselves.

Chapter 3

Automation and acceleration of simulation

Data, without any doubt, is the base of any data analysis and any data-driven process. In this research, we mainly gather data from simulations. Therefore, how to enhance the efficiency of executing simulations is also a key point. The time cost of simulations consist of two parts, modeling and calculation. As millions of cases would be studied in this research, it is impossible to modify all the models manually. Automatic modeling methods are necessary. On the other hand, reduced the time cost of calculations is also a theme.

In this research, *EnergyPlus* was used as the simulation tool for heat balance and energy consumption analysis. There is a Python library called EPPY that can efficiently generated E+ models but is not easy to use. In this research, we made a improved version of EPPY named EPPiX, with the geometric functions reinforced. To execute multiple simulations simultaneously, a tool called Ultimate EP executor was developed using multi-process library in Python.

FlowDesigner was used to analysis the wind environment in this research. Advanced Knowledge Laboratory Inc., the mother company of FlowDesigner, provides a parametric modeling interface called *FD automation*. In this research, we converted the command-line styled *FD automation* into total object oriented programs and developed 2 tools, *GH2FD* for *Grasshopper 3D* and *PyFD* in pure Python environment.

3.1 Parametric modeling and parallel execution of *EnergyPlus*

In this research, *EnergyPlus*[65] (EP) has been used to execute the heat load simulations. EP is a dynamic whole building energy simulation program. The development of EP started in 1996, by the combination work of BLAST[1] and DOE-2 building energy simulation program[7], funded by the U.S. DOE[2]'s BTO[3]. NREL[4] manages the development of and in charge of the maintaining. The best capabilities of both software have been inherited by EP.[7] Due to its efficiency, robustness, expansibility, the free software EP is the very most widely used simulation software in researches, accoding to the survey carried out by Stevanović[24] in 2013, which covered all the previous research articles. Table 3-1 shows the results of the survey.

Table 3-1: Summary of a full review of passive solar design strategies by Stevanović 2013 (Yasin | 2017)

Usage frequency	Software	Scope of use
■■■■■■■■■■■■■■■■■■■■	EnergyPlus	Energy simulation software
■■■■■■■■■■	TRNSYS	Transient system simulation tool
■■■■■■■	DOE-2	Building energy use and cost analysis
■■■■	TAS	Thermal analysis simulation software
■■■	DEROB-LTH	Dynamic energy simulation
■■■	EPS-r	Integrated energy modeling tool
■■■	IDA ICE	Indoor climate and energy tool
■■	BEopt	Building energy optimization software
■■	GenOpt	Generic optimization program
■■	IES-VE	Building performance simulation
■■	Radiance	Synthetic imaging system
■	ADELINE	Daylight and artificial lighting simulation
■	EC501	Building thermal performance calculation
■	ENERGY	Building thermal performance prediction model
■	e-QUEST	Quick energy simulation tool
■	FLUENT	Computational fluid dynamic software
■	HTB2	Thermal simulation of building
■	LT	Method energy performance curves estimation
■	OPTI	Energy consumption, thermal comfort and daylighting
■	PHPP	Passive house planning tool
■	SUNCODE-PC	Building thermal analysis simulation
■	THERB	Thermal environment simulation software
■	VisualDOE	Building energy simulation software
	modeFRONTIER	Multi-objective optimization and design environment
	SIMBAD	Building energy consumption tool
	ZEBO	Decision support tool for zero energy design

[1] Building Loads Analysis and System Thermodynamics
[2] Department of Energy
[3] Building Technologies Office
[4] National Renewable Energy Laboratory

Since EP is an open-source software, developed in *C* language, researchers are able to access its core part through an interface or add customized functions or calculations into the simulations. The input and output files of EP, IDF and CSV files, are totally ASCII text file, which can be easily modified using programming language, or parsed to programming objects. Therefore, EP was also used in the optimization processes. Attia et al.[21] carried out a survey on the combination of simulation tools and optimization tools. As illustrated in the figure 3-1, EP was the most frequently used simulation tool, meanwhile EP + Matlab was the most frequently used combination.

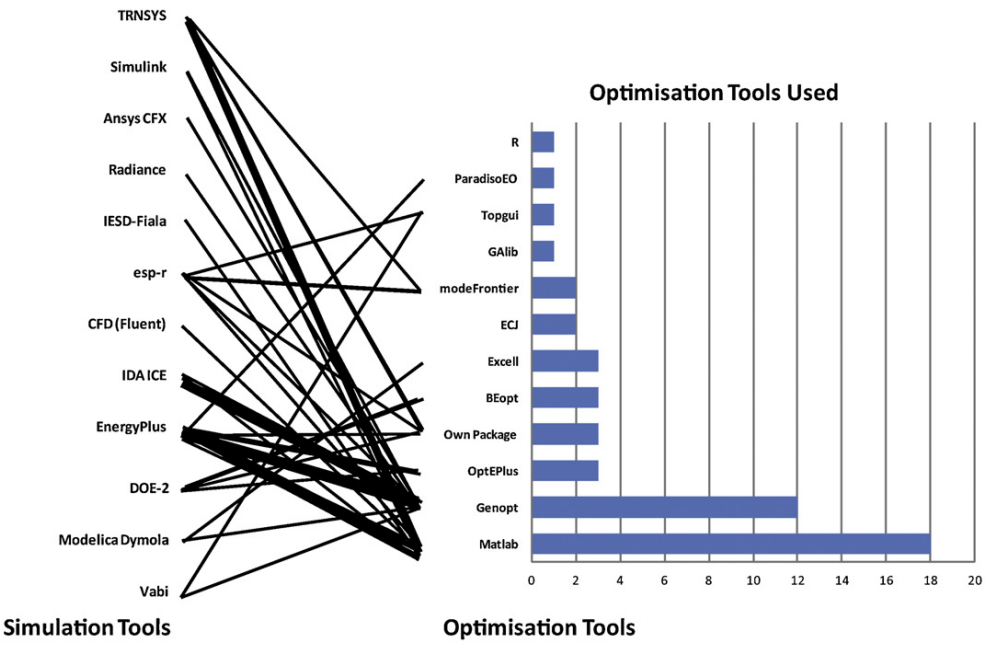


Figure 3-1: Optimization tools order by use (right) and simulation tools ordered by use (left). The line thickness is proportional to the frequency of the pairings (Attia et al. | 2013)

Beside its robustness and popularity, the reason that we chose to use EP in this research is that there are a lot of third-party tools developed for EP, such as OpenStudio[cite] and Honeybee[cite], which are based on the architects' favorite modeling tools and able to be smoothly incorporated into early stage of environmental design.

OpenStudio is SketchUp[cite] based modeling tool for EP. *OpenStudio* (OS) has the same mother as EP, NREL, so that it can be regarded as official. With the OS plug-in in SketchUp, we can finish geometric modeling, change the names of building components and match overlapping surfaces (figure 3-2). OS also has an independent GUI (figure 3-3), with which the non-geometric properties in a OS model can be edited. One of the big advantages of OS is that it carries a library including a lot of templates from ASHRAE Standard 90.1[61]. We can easily apply the schedules, materials and constructions, people definitions, etc. into our own models by dragging them from the library. Ideal AC system can also be easily configured in this GUI.

Besides, there is also a graphic HVAC modeling panel in OS. Regretfully, as we are architects who are lack of knowledge of AC system, we have not tried that part.

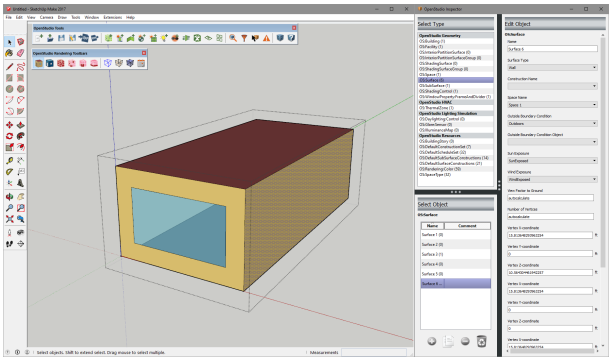


Figure 3-2: SketchUp with OpenStudio plug-in

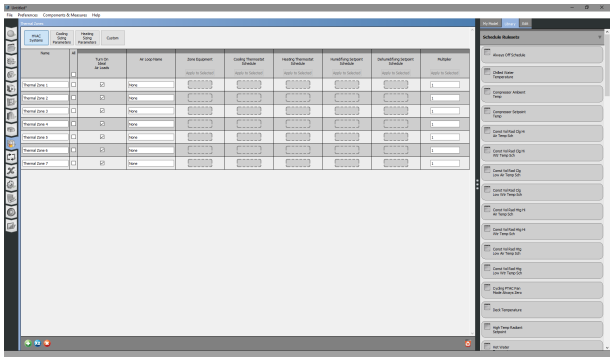


Figure 3-3: GUI of OpenStudio

The biggest disadvantage of OS is that its functions are limited. Many details in a EP model cannot be edited using the GUI of OS, such as Air Flow Network (AFN) or Phase Change Material (PCM). Therefore, we usually output the IDF files from OS and edit them in IDFEditor, a pre-process tool installed along with EP Windows version. Figure 3-5 shows the interface of IDFEditor, the information of official manual is also displayed to help users understand each property of each object in a model.

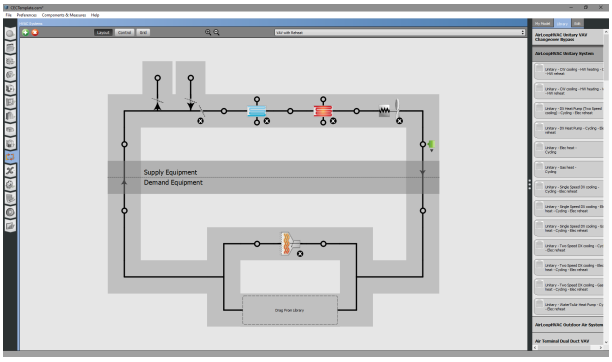


Figure 3-4: SketchUp with OpenStudio plug-in

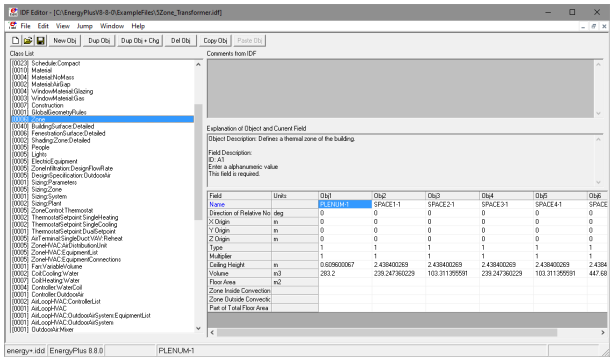


Figure 3-5: GUI of OpenStudio

By the way, there is another tool, named Parametric Analysis Tool, installed along with OS. However, as we could not find enough documentations for it, even not a operable official tutorial, we haven't use it.

3.1.1 Parametric IDF generation

Figure 3-6 shows the manual and parametric flows of using EP, which have been polished in the practices by the members of MaeLab and become quite sophisticated. The manual flow have already been introduced in the last section. In this section, we will concentrate on the parametric flow and explain the reason why we chose to improve EPPY into EPPiX.

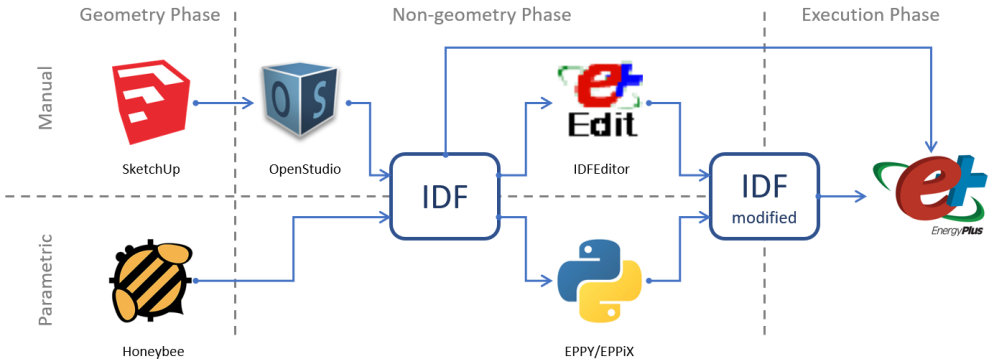


Figure 3-6: The manual and parametric flows of using *EnergyPlus*

Though, using OS and IDFEditor, we can control every detail in an EP model, this manual process is too low-efficient to make thousands of IDF files for statistic study. Therefore, we need a parametric process to automatically generate the models we need.

The most simple way is to use the parametric object in EP. As illustrated in figure 3-7, instead of inputting the exact value, the name of the corresponding parametric object is used as the value of an attribute. EP will firstly convert a IDF file into multiple files, corresponding to all rows in parametric object class. However, there are some disadvantages of this way. The first one is actually funny. A '=' mark is needed when referencing to the name of a parametric object. However, the '=' cannot be inputted in IDFEditor, which causes errors. orz. The second one is that the number of rows contained in a parametric object is limited up to 100. Thirdly, though EP will not run the converted multiple IDF files in parallel.

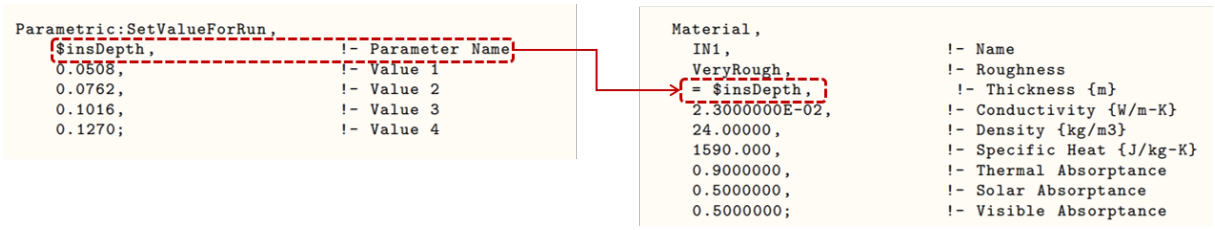


Figure 3-7: The usage of parametric object in EP (U.S. Department of Energy | 2017)

The most famous parametric modeling tool for EP should be Honeybee, a third-party plug-in in Grasshopper 3D (GH). Making full use of the curved surface modeling ability of *Rhino* and GH, Honeybee is very powerful at dealing with complex, especially curved, geometries. Though GH is my favorite modeling tool, Honeybee is not a proper choice for this research. First, rather than modify a model, Honeybee will build a new model from 0 when there is any modification in the input, which is quite low-efficient. GH is very good a process a bunch of objects with the same logic. But it very very troublesome to separate objects in a list and process each independently. Due to this characteristic of GH, it is also very annoying to do detailed configurations in Honeybee. Moreover, as there is no loop operation in GH, a GH master will make use of the DataTree, a individual data structure in GH, and control the Path of each branch carefully to perform an equivalent process of loop. However, the data structure control within most Honeybee components is in a mass. Therefore, Honeybee will flatten the input data by default, which total destroys the structure if data flow in GH and kills any hope of generating multiple models simultaneously. Finally, the things that can be configured in Honeybee are also limited.

A Python library called ***EPPY*** is used in this research. EPPY is able to parse the objects in a IDF file into Python objects. So that edit a EP model in an object oriented way (figure 3-8), instead of annoying string operations. EPPY is almost the ideal parametric modeling tool for this research.

```
Material,
    G01a 19mm gypsum board,    !- Name
    MediumSmooth,              !- Roughness
    0.019,                     !- Thickness
    0.16,                      !- Conductivity
    800,                       !- Density
    1090,                      !- Specific Heat
    0.9,                       !- Thermal Absorptance
    0.4,                       !- Solar Absorptance
    0.4;                       !- Visible Absorptance
```

```
material_0 = epm.newidfobject("material".upper())
material_0.Name = "G01a 19mm gypsum board"
material_0.Roughness = "MediumSmooth"
material_0.Thickness = 0.019
material_0.Conductivity = 0.16
material_0.Density = 800
material_0.Specific_Heat = 1090
material_0.Thermal_Absorptance = 0.9
material_0.Solar_Absorptance = 0.4
material_0.Visible_Absorptance = 0.4
```

Figure 3-8: Left: IDF object in text, Right: parsed object by EPPY

However, EPPY also has its disadvantages. Because EPPY faithfully parses the EP object into Python, the coordinates of vertices of a surface are not in a list, but stored in different attribute. So we need to manually type lot of codes to modify the geometric information of a surface, rather than simply use a loop.

```
surface_10 = epm.getobject("BuildingSurface:Detailed".upper(), "Surface 10")
surface_10.Vertex_1_Xcoordinate = 0
surface_10.Vertex_1_Ycoordinate = 9
surface_10.Vertex_1_Zcoordinate = 3.6
surface_10.Vertex_2_Xcoordinate = 0
surface_10.Vertex_2_Ycoordinate = 9
surface_10.Vertex_2_Zcoordinate = 0
surface_10.Vertex_3_Xcoordinate = 0
surface_10.Vertex_3_Ycoordinate = 0
surface_10.Vertex_3_Zcoordinate = 0
surface_10.Vertex_4_Xcoordinate = 0
surface_10.Vertex_4_Ycoordinate = 0
surface_10.Vertex_4_Zcoordinate = 3.6
```

Figure 3-9: Editing geometric information using EPPY

Therefore, we made some improvement on EPPY. We call this improved version EPPiX. By study the member dictionary, we find the structure of an EPPY is very special. Unlike a normal Python object, the names of all attributes are stored in a list called 'objls', and the values are stored in a list called 'obj'. Therefore, the __getattr__ and __setattr__ functions cannot be used, neither the __dict__ member. So we decided to directly edit the 'obj' list. Some simple geometric operations, such as move, scale and mirror, were also added into EPPiX. Besides, we also made a simple GH component that is able to convert the vertex coordinates of surfaces into a CSV file which can be directly utilized by EPPiX. The source code of EEPiX is attached as an appendix to the end of this thesis.

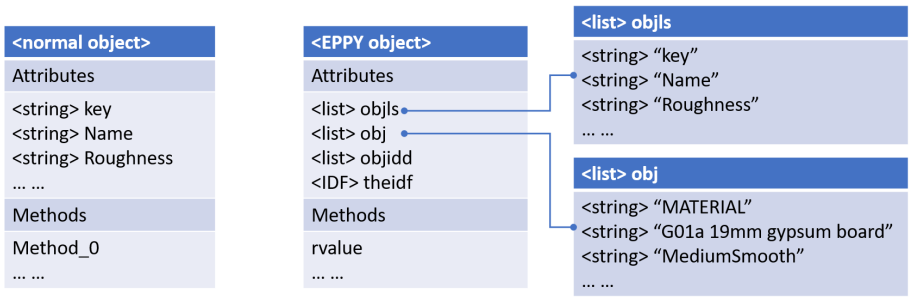


Figure 3-10: Left: The structure of a normal object, Right: The structure of a EPPY object

3.1.2 Multi-processing execution

Running multiple simulations in parallel simultaneously, making full use of the multi-core CPU power, is the key to reducing the time cost of simulation and accelerating the whole process.

In the official EP execution tool, EP-launch, that is installed along with the EP Windows version, there is a 'Group of Input Files' function which is able to execute multiple simulations in parallel. This function is very easy to use. A simulation group can be created from multiple IDF files as illustrated in figure 3-11 and the number of simultaneous processors to use can be configured as illustrated in figure 3-12.

However, in this research, this 'Group of Input Files' function is not perfect. We found EP-launch would become unstable in the situation that the number of IDF files exceeds 10,000. On the other hand, EP-launch remains all intermediate files so that the disk usage is very huge.

In this research, we developed a new tool called Ultimate EP Executor to run simulations in parallel. 2 versions have been developed respectively using Python and C#. The Python-ver. is cross-platform, which is able to run in Windows, Linux and MacOS, meanwhile the C#-ver. is Windows-only but has a GUI, which is more friendly for beginners.

By searching for the documentations, we found the way to trigger an EP simulation using command line.

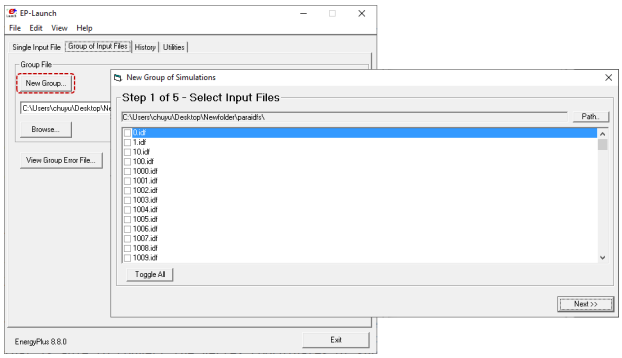


Figure 3-11: Creating a simulation group from multi IDFs

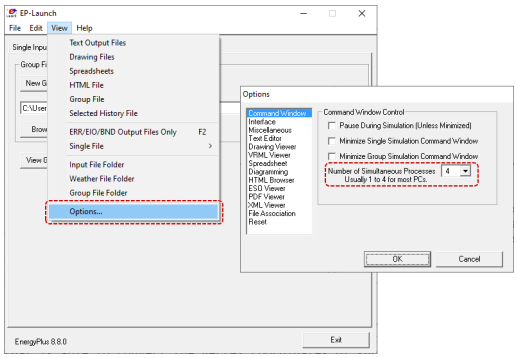


Figure 3-12: Configuring number of processors

Therefore, we can pass the commands to subprocess module in Python, or a Process object in C#, to launch an EP simulation in our program. If an IDF model contains any objects in 'HVAC Template' family, it needs to be expanded, converting the HVAC template objects to objects in 'Zone HVAC' series. EP will automatically expand an IDF file if we contains the '-x' optional in the commands. However, in practice, we found it is very unstable to ask EP automatically expand an IDF file. A better way is to expand the IDF file independently with ExpandObjects.exe in advance. Figure 3-13 illustrate the process of running one simulation. Only the error blog and the result CSV file remain, other intermediate files are deleted. Figure 3-13 shows the flow of run a single simulation in Ultimate EP Executor.

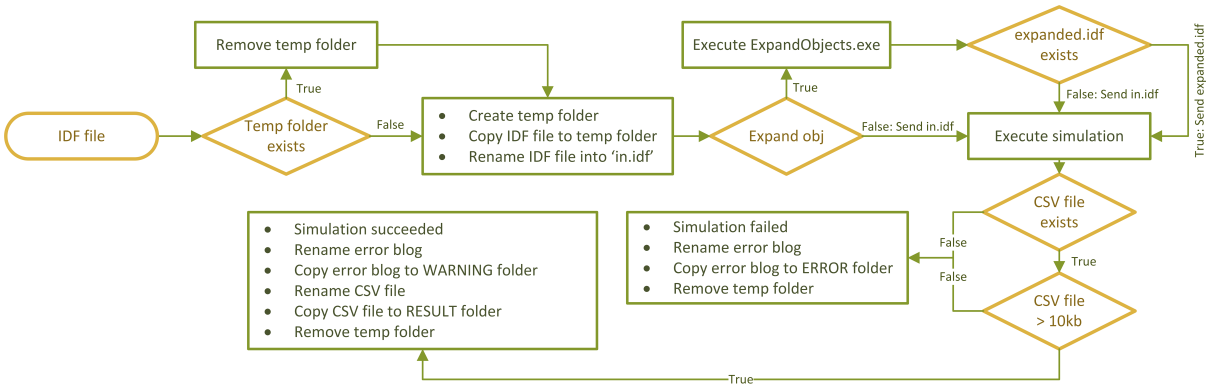


Figure 3-13: The flow of run a single simulation in Ultimate EP Executor

After succeeding in executing EP simulation in Python and C#, we used the multiprocessing module in Python, and the Parallel class under System.Threading.Tasks namespace in C#, to execute EP simulations in parallel.

3.2 Developments based on *FlowDesigner Automation*

FlowDesigner (FD)[48] is a CFD software. As indicated by its name, FD is designer oriented tool that is quite user friendly. The GUI of modeler of FD is very easy to use, including all kinds of geometric operations (figure 3-14). The result reviewer can also easily output different kinds of diagrams to visualize the simulation results vividly (figure 3-15).

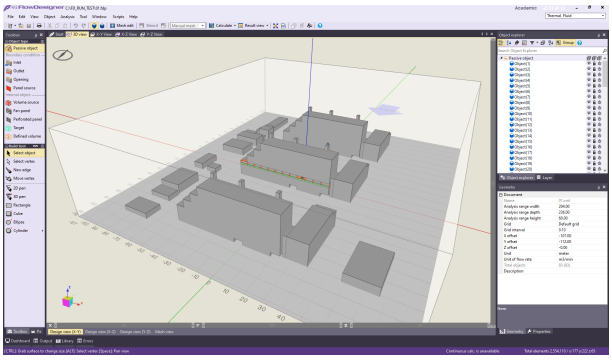


Figure 3-14: FD modeler GUI

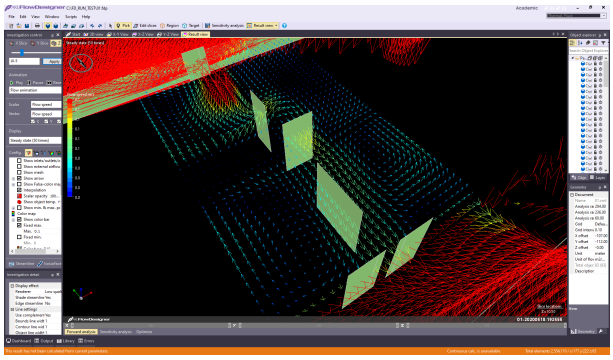


Figure 3-15: FD Viewer GUI

However, this modeler is total manual. In many cases, the user have to do duplicated works to build multiple similar geometries, such as the inlets and outlets on the ceiling in an office room. It is also troublesome to modify the geometries manually, for example, change the layout of inlets from 3×5 to 4×4. Therefore, in 2015, we proposed the development plan of a GH plug-in, that is able to convert the geometries in Rhinoceros and GH into FD objects and configure their properties, to AKL, the company that developed FD. We asked for a .NET API[5], as the developments of GH components or Revit plug-in are in .NET platform, using C# or Visual Basic .NET.

About 1 year later, in the late 2016, AKL released their cross-platform API for FD, named FlowDesigner Automation[47]. FD Automation can be utilized with .NET, VBA, WSH[6], Python and even PHP. The mechanism of FD Automation is to send text commands to FD through an AklModeler.CommandControl object. Therefore, we can say that FD Automation is a command-line based tool. Here are 3 examples. Figure 3-16 explains the command to create a geometry. Figure 3-17 explains the command to set the external wind speed.

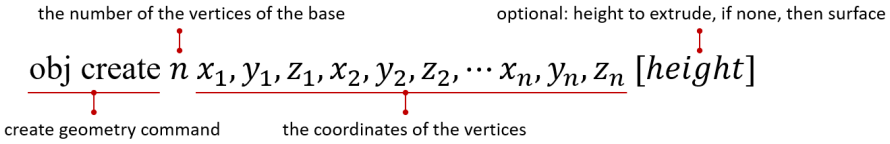


Figure 3-16: Command to create a geometry

[5] Application Programming Interface
[6] Windows Script Host

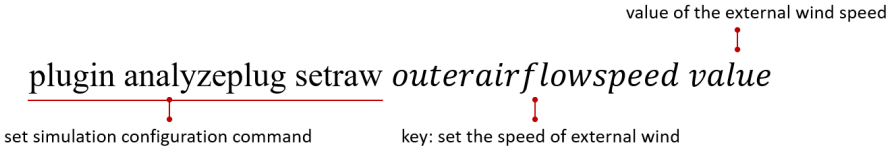


Figure 3-17: Command to set the external wind speed

3.2.1 GH2FD

Though, through FD Automation, we can fully control FD using commands, it is very inconvenient for most users, especially for designers. The learning curve is too long for a professional architect, as it is impossible for them to suspend their work and spend 1 or 2 weeks on learning how to use these commands.

The most popular parametric modeling tool used in architecture design industry is GH. Therefore, we developed a GH plug-in called GH2FD[70]. Rather than directly connect FD Automation with GH, we developed a intermediate .NET library called FD.NET, converting this command-line based API into a totally object oriented API. So that, with FD.NET, we can easily develop not only GH2FD, but also plug-in for Revit and other software with .NET API in the future.

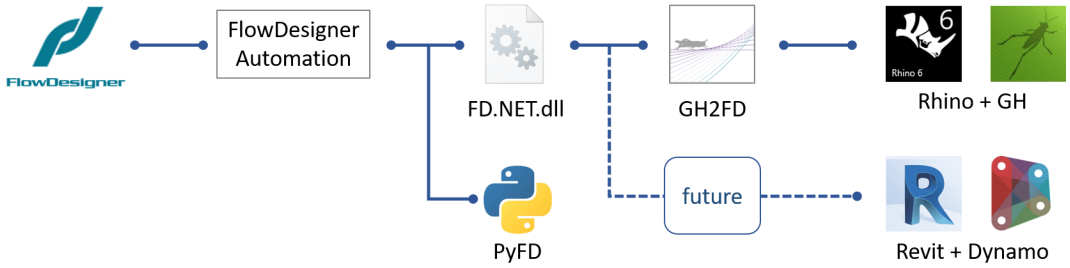


Figure 3-18: The framework of third-party development based on FD Automation

Figure 3-19, 3-20 and 3-21 show the composition of classes of FD.NET. There are four groups of classes: 1)FD commander, 2)Setting related, 3)Geometry related and 4)FD object related.

The FD commander is a static class that sends the commands to FD. It also includes some basic function, such as selecting object and deleting object.

FD_Setting is the parent of all other classes in Setting related group. In FD_Range, FD_External_Wind and FD_Mesh, each time the value of an attribute is modified, a corresponding command string will be added into the Update_strings list. Property, a kind of special variable in C#, is used to realize this operation. A

Property in C# is actually a combination 2 functions, get and set. The get function will be called when we get the value of a Property with '=', while the set function will be called when we set the value with '='. When the method Update() is called, all the command strings in Update_strings will be sent to FD. We did not choose to send command string directly to FD each time an attribute is modified. This is because the characteristic of GH. Each time one of the inputs of a GH components is modified, this component will be executed once. So that many commands would be sent to FD more than once, which reduces the efficiency significantly. Though FD.NET was not only for GH, GH was always the main factor to consider when we were designing FD.NET.

In the Geometry related group, FD_Object class is the root of FD_Cube, FD_Panel and FD_Mesh, while FD_Vertex is their aggregation. FD_Panel is a actually list of FD_Vertex, which includes the geometry information for panel-shaped objects in FD. Similarly, FD_Cube is a actually list of FD_Vertex, which defines the base, and the height of the geometry, as 3D objects in FD are generated by extrusion. FD_GeoMesh follows the definition of polygon mesh[84], which is used to creating complex geometries in FD.

A class in the FD object related group is actually 'a list of FD_Object' + 'non-geometric information', inherited from FD_Group. There are 12 classes corresponding to the 12 object types in FD. For illustration, we just pick up 2 of them, FD_Perf_Panel and FD_Defined_Volumn, which respectively stands for panel-shaped and 3D objects in FD. Each class includes a list of FD_Object, rather than a single FD_Object, because in many cases, different objects would share the same non-geometric properties. The Create_Set() function will first create all the geometries in the list, then select all these newly created objects and modify their non-geometric properties.

The classes in CSV related group is used to control the CSV output configurations in FD. Originally, we thought that it was more convenient to output CSV files within the FD modeler. But based on the feedback of users, as they want to fully control FD from GH, we added this part.

3.2.2 PyFD

Similar to Honeybee, GH2FD is also limited by the characteristics of GH. It is easy to created objects but a little difficult to modify an existing object. On the other hand, it is hard to realize loop in GH.

Therefore, we develop an equivalent of FD.NET in Python, named PyFD. The functions of PyFD are almost equivalent to those in FD.NET. However, there is a significant difference between PyFD and FD.NET. Each time the value of an attribute of an object in PyFD is modified, the corresponding command string will be directly sent to FD. We also added loop function into PyFD.

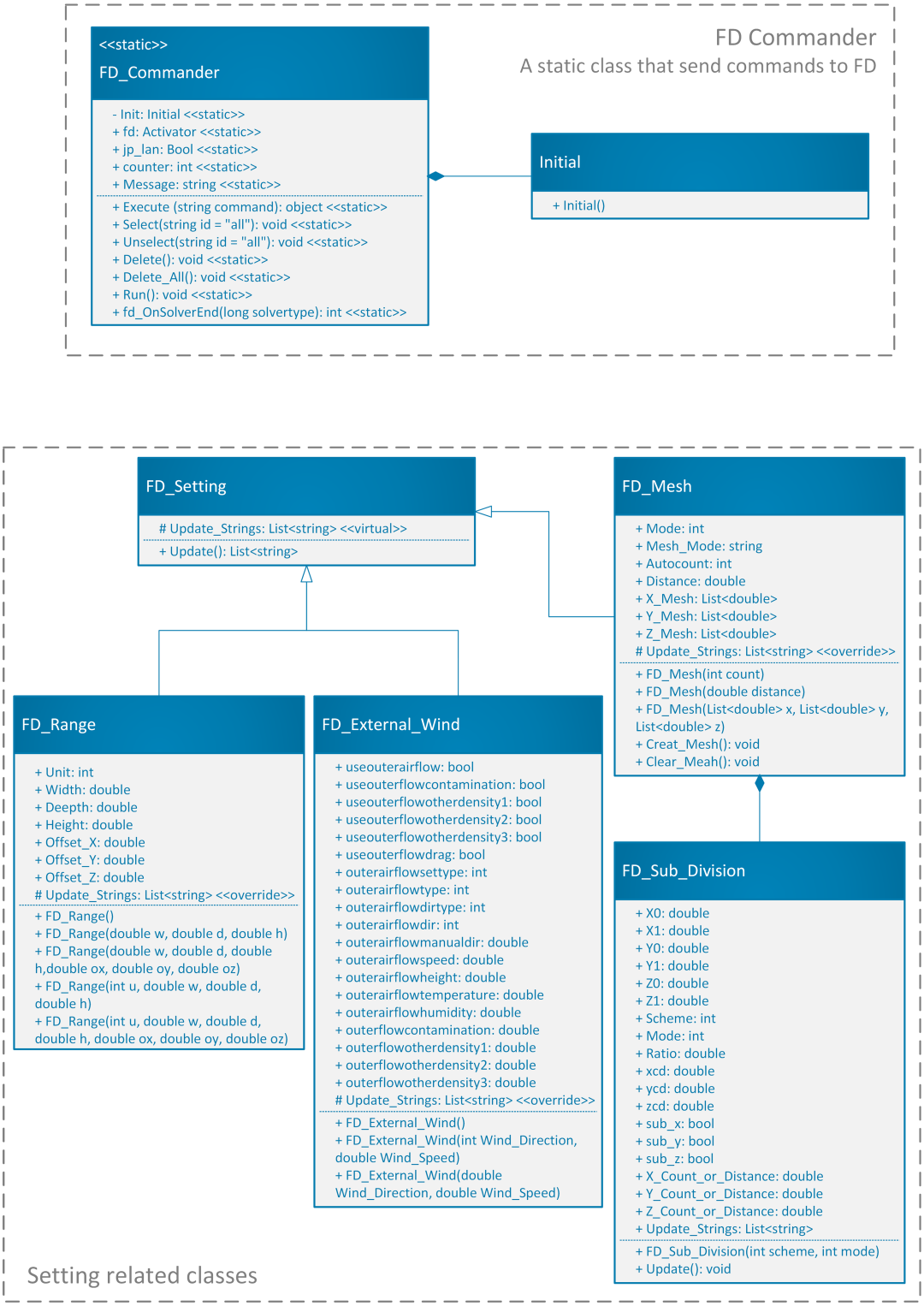


Figure 3-19: Class diagram of FD.NET (part 1)

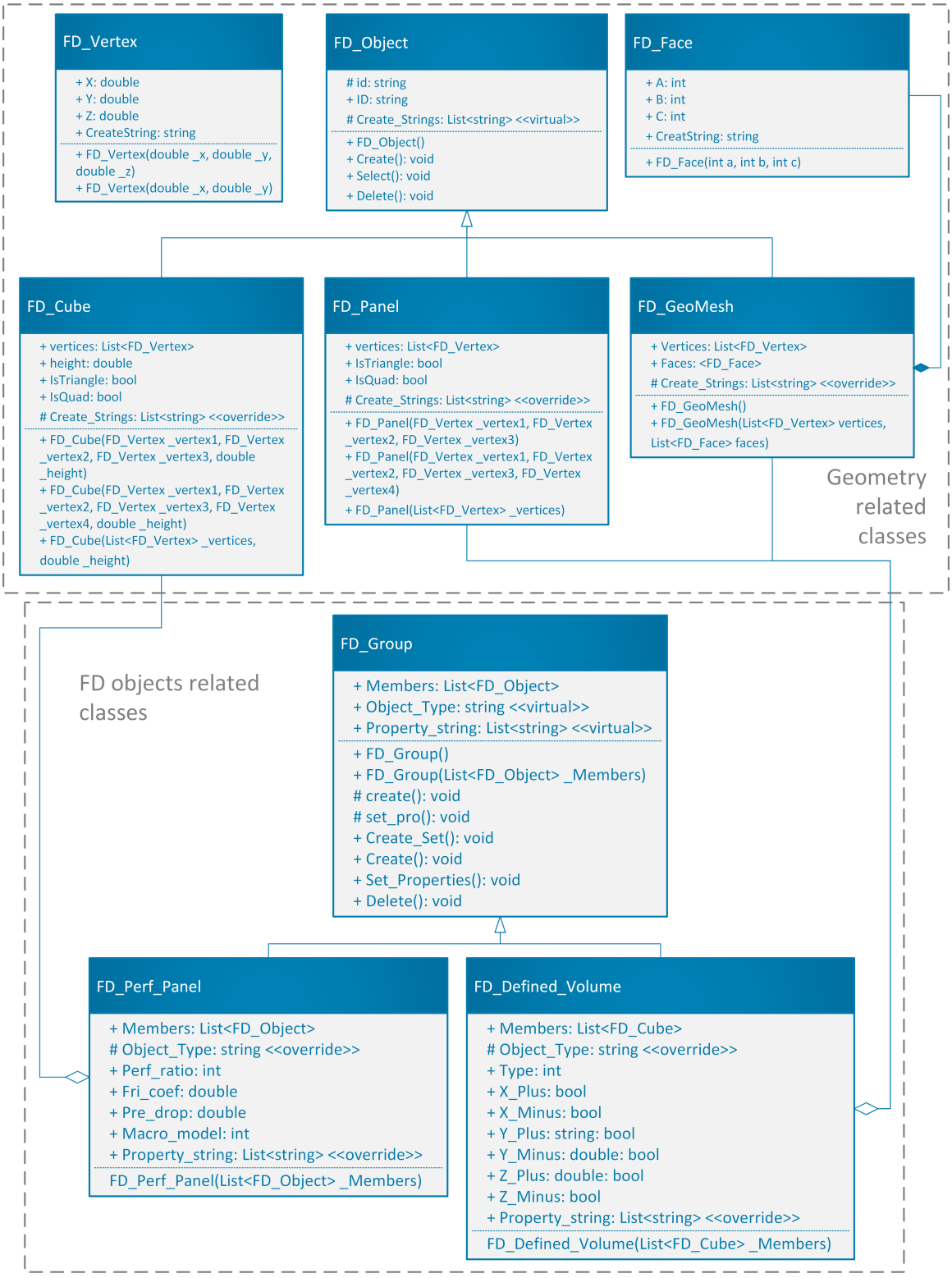


Figure 3-20: Class diagram of FD.NET (part 2)

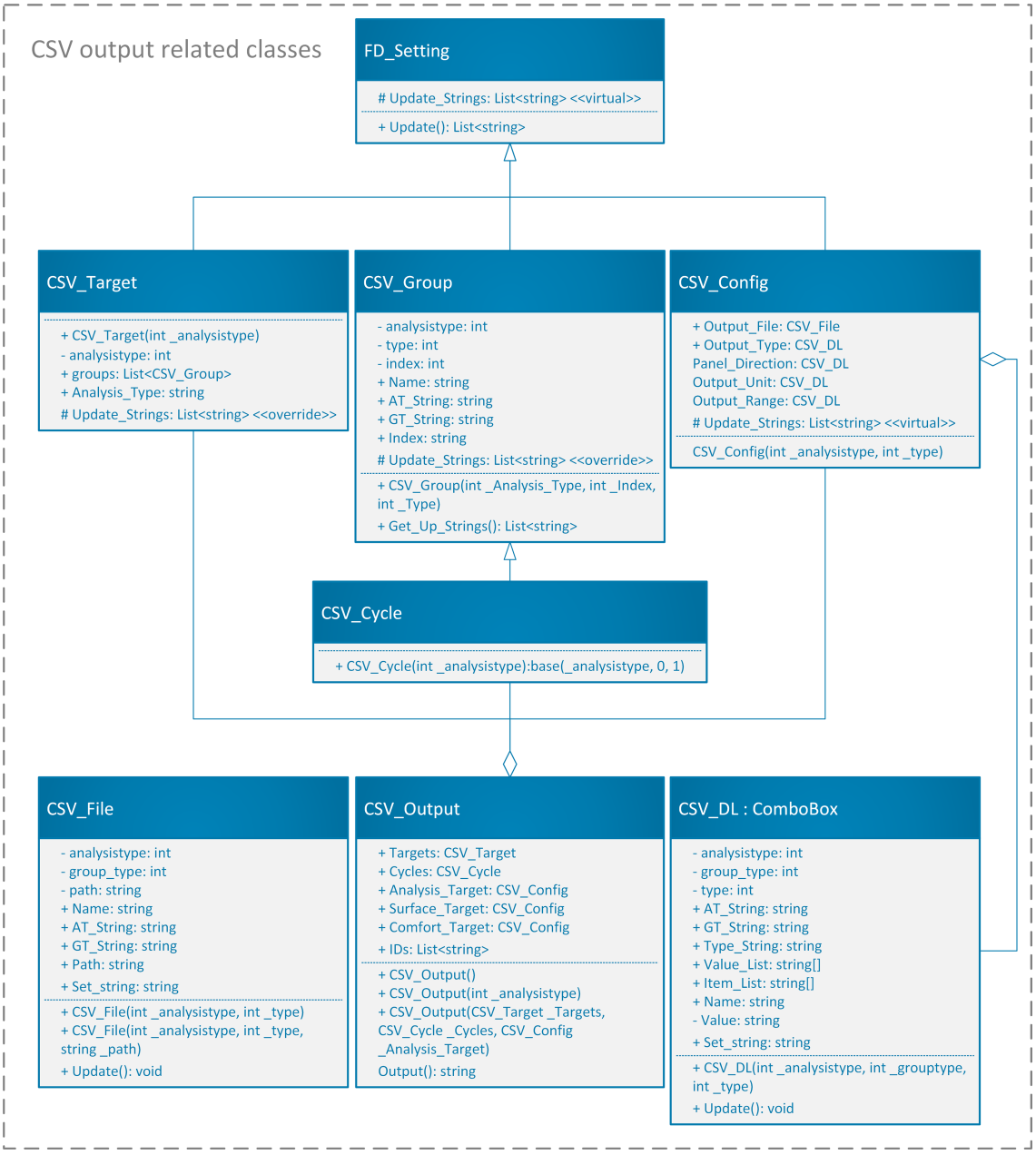


Figure 3-21: Class diagram of FD.NET (part 3)

Chapter 4

Methods of sensitivity, uncertainty and interaction analysis

In this research, we used sensitivity and interaction analysis to reduce the dimension of the problem. This chapter begins with introducing the concept of the curse of dimensionality and explaining the necessity to reduce the dimension, followed by introductions to the sampling methods used in this research.

The second part is the introduction the methods of sensitivity and interaction analysis, beginning with the definition of sensitivity and uncertainty. Sobol first-order indices and Morris method are disturbed. Including the algorithm, the advantage and disadvantage, as well as the improvements. The concept of interaction between parameters, along with the Sobol second-order indices and Expanded Morris method, were also introduced. This part ends with a demonstration of these method.

The last part id the original research of R^2 models. By introducing a new concept y^{true}, we improved the R^2 model to deal with uncertainties and proposed 2 new models, R^2_U and R^2_E.

4.1 The necessity to reduce the dimension

In the fields such as sampling, numerical analysis, meta-modeling and database, the curse of of dimensionality[79] is always a nightmare for researchers. Valid and effective sampling is the base of all researches in these fields. However, as the volume of the decision space expands exponentially with the arising of its dimension, the necessary sample size to perform valid sampling also increases exponentially. For example, in 1 dimensional space $[0, 1]$ (unit interval), 10 sample points is enough to sample this space with the average distance between adjacent points close to 10^{-1}. In a 2-D space (unit square), to perform a equivalent sampling, the necessary sample size is enlarged to $10^2 = 100$. Then sample space becomes 10 times larger. If we sampled $10^2 = 100$ points in an unit interval, the equivalent sampling in an unit square needs $(10^2)^2 = 10,000$ points, and $(10^2)^3 = 1,000,000$ in an unit cube. The expansion of the decision space is related to both the dimension and the sample size for each interval (equation 4.1).

$$E = \frac{s^d}{s} \tag{4.1}$$

$$= s^{d-1} \tag{4.2}$$

Where:

E = The times of expansion
d = Dimension of the decision spaces
s = Sample size for each unit interval

However, it is very tricky to decide the necessary sample size for each unit interval size to perform a valid sampling. Though many researchers follows a rule of thumb that samples 5 points for each dimension[79] and 5^d points in total, the necessary sample size for each dimension is strongly related to the smoothness of the landscape of the $f(x)$. Figure 4-1 shows the landscape of 2 functions, $f(x) = \sin(x)$ and $f(x) = 0.5x + 0.2\sin(20x) + 0.1$. 5 points are sampled within $x \in [0, 1]$ for each function.

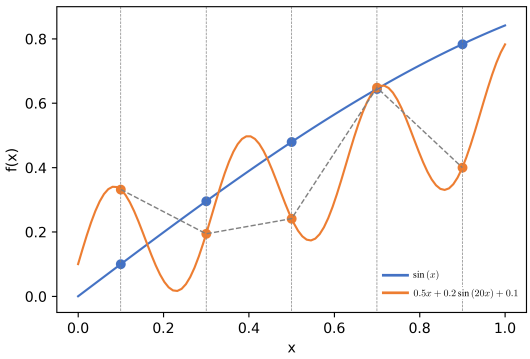


Figure 4-1: Smoothness of the $f(x)$ landscape

From this figure we can see that the landscape of $f(x) = \sin(x)$ is much more smooth than that of $f(x) = 0.5x + 0.2\sin(20x) + 0.1$. As a result, 5 points are enough to describe the curve of $\sin(x)$ but fail to reflect the shape of $0.5x + 0.2\sin(20x) + 0.1$.

It is very important that the sample points are able to reflect the landscape of the research objective. For example, in genetic algorithm, the population of the first generation is critical. Without enough population, the algorithm is likely to converge at a local optimal. Similarly, in the process of meta-modeling, if the sample points fail to reproduce the landscape of the original objective, usually the errors in the trained meta-model would be too big so that the model can hardly give out decent predictions.

It is a paradox that, if we want to estimate the necessary sample size, we need to know the landscape of $f(x)$. But we can hardly get an overview of the landscape until we have enough sample points. Theoretically, it is an unsolvable problem that how many sample points are necessary for each dimension. Most researcher tends to gradually add points into the original sample collection based on the analysis results, until they are satisfied with the results. To conclude, it is very difficult for us to reduce the total sample size by limiting the points on each unit interval.

Another factor related to the necessary sample size is the dimension of the decision space, which is also called sample space or input space in some contexts. The volume of the decision space is mainly determined by the dimension, or the number of parameters in another word. With the reduction of the dimension, the volume of the decision space shrink dramatically. Therefore, in a smaller decision space, we can better fill it with less samples. From the equation 4.1, we can also see that the relationship between expansion of the decision space E and the sample size of each dimension s_{1d} is cubic, while that between E and the dimension d is exponential. Therefore, it more effective to control the value of d, rather than s_{1d}, especially in high-dimensional problems.

In this research, we used both sensitivity and interaction analysis to reduce the dimension of decision space. Sensitivity analysis is used to reduced the number of parameters, while interaction analysis is used to divide a high-dimensional problem into several relative-low-dimensional problems (figure 4-2).

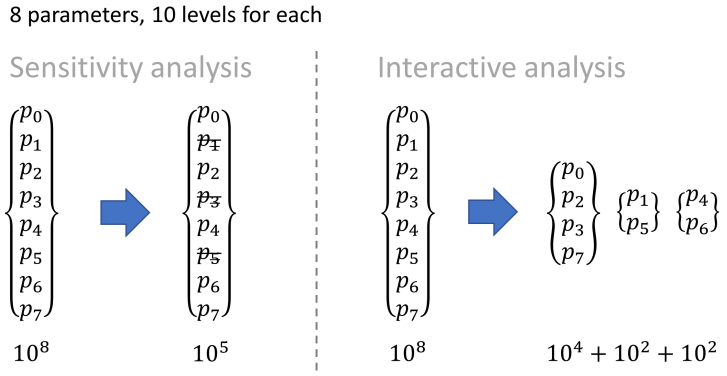


Figure 4-2: Reducing dimension by sensitivity and interaction analysis

4.2 Sampling methods

Valid sampling is the solid base of any statistical researches. Before talking about the sensitivity and interaction analysis methods, we would like to talk about the sampling methods in advance. Figure 4-3 illustrates the three most widely used sampling methods by researches.

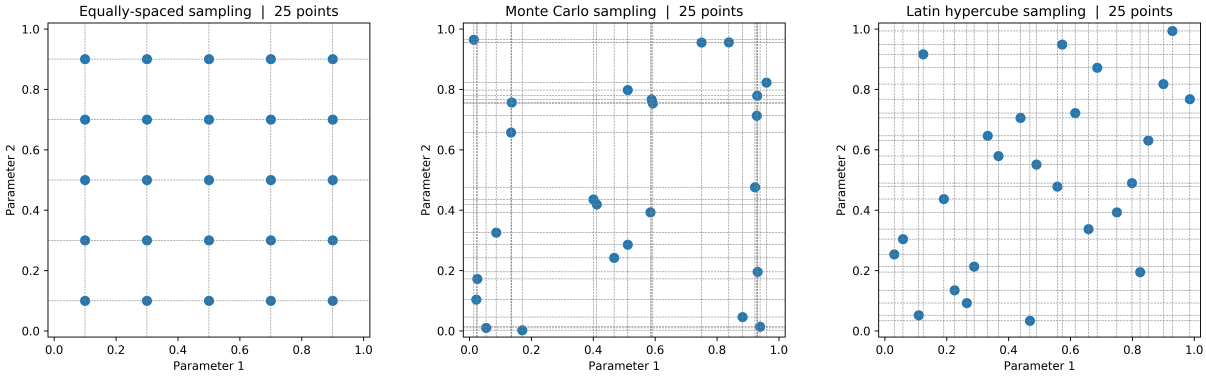


Figure 4-3: Comparison between sampling methods

4.2.1 Equally-spaced sampling

Equally-spaced sampling, which is also called evenly-spaced sampling, is a very traditional and the most widely used sampling method in all fields.

In the first step, several points will be sampled on each unit interval, the distance between 2 adjacent points are the same. Therefore, these point are equally/evenly spaced. There are normally 2 ways to do the equally-spaced sampling on an unit interval. The interval will be firstly evenly divided into n sub-intervals. Normally, The boundaries or mid-points of the sub-intervals are used as the sampling points, which generate respectively $n + 1$ or n points.

The second step is to cross reference the sampling points between dimensions. The sample collection can be regarded as a d-dimensional matrix. The sample size in total is s^d, where s denotes the number of points in each dimension.

Equally-spaced sampling is most widely used for that it is the most simple and intuit method. For there is no random process, this method is extremely practical in pre-computer age. The disadvantages of equally-spaced sampling are also very obvious. The points in each dimension is sparse. In each unit interval, the space between adjacent points is wide. If we despite the interactions between parameters, the exploring in each dimension is not enough. Another disadvantage is that the total sample size is unavoidably increases

exponentially with dimension. The curse of of dimensionality is more powerful on this method than on others.

There is a derivative of equally-spaced sampling (figure 4-4). The points in each unit interval is sampled randomly. Then, identically, points between dimensions are cross referenced. This method hasn't been 'officially' named, might because it hasn't been much adopted. We can call this method this method 'Cross-referenced Monte Carlo' method. This method is not much different from the original one. What's more, it also has the same disadvantages as MC method. The points might be too close or too faraway.

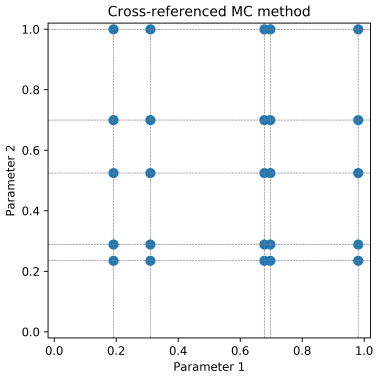


Figure 4-4: Cross-referenced Monte Carlo sampling

4.2.2 Monte Carlo sampling

Monte Carlo sampling, or simply called random sampling, is the most widely used sampling method by researchers. For each sample, the points in different unit interval are sampled in random independently, then combined into a sample. Different samples are also generated independently.

Compared to equally-spaced sampling, Monte Carlo sampling can fill the space of each unit interval much better. Theoretically, the points in each unit interval is equal to the total sample size. In each unit interval, the distance between adjacent points is much narrow. As all the samples are generated independently, it is not necessary to know the total sample size in advance. Special sampling scheme is neither needed. Researchers can also easily generated additional samples at any time. Another advantage of Monte Carlo sampling is that samples can be generated in respect to a specific distribution, such as normal distribution or Weibull distribution.

However, just because all points are sampled independently, there are chances that the points are too close to each other or too faraway. The extremely narrow or wide space between points would reduce the validness of the total sample, though this problem could be relieved by enhancing the sample size.

4.2.3 Latin hypercube sampling

Latin **H**ypercube **S**ampling (**LHS**)[cite] is a kind of derivative of Monte Carlo sampling, which is decried in as early as 1979 by Michael McKay of Los Alamos National Laboratory. LHS get it name from Latin square[cite], which is very similar to Sudoku[1] in Asia (figure 4-5). The basic conception of LHS is that, for example in a 2-D dicision space, there is only on sample point in each row and each column. With the dimension goes up, the square becomes a cube (3D) or hypercube (high-dimensional). LHS take the advantages from both equally-spaced sampling and Monte Carlo sampling.

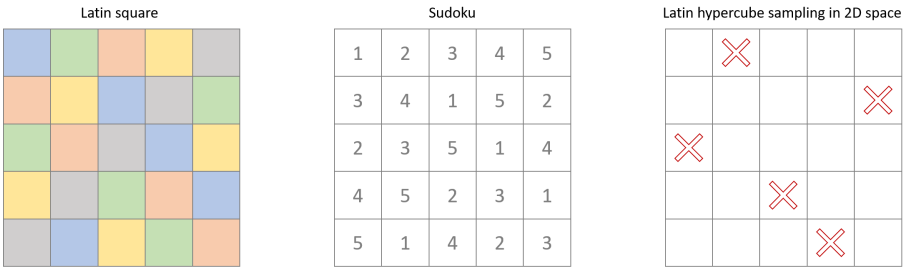


Figure 4-5: Comparion between Latin square, Sudoku and LHS

The first step of LHS is to decide the total sample size s_t. Then, each unit interval is divided evenly into s_t sub-intervals. In the second step, a random point will be generated within each sub-interval. Then, one point from each dimension will be picked up randomly to make a full sample. The total sample collection satisfies the requirements of Latin hypercube, as there is only one sample in one sub-interval in one dimension. Figure 4-6 illustrates the process of LHS.

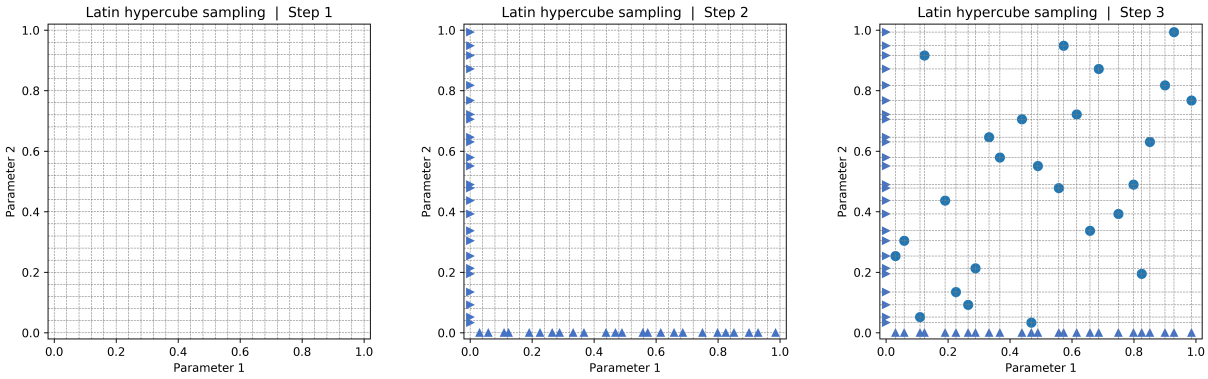


Figure 4-6: The process of Latin hypercube sampling

[1] 数独

Compared to equally-spaced sampling, for each dimension, LHS fills the unit interval much better. By evenly dividing the unit interval in advance, compared to the common Monte Carlo method, LHS avoid the problem that several points are too close to each other, or the distance between adjacent points is too wide. The sample collecting generated using LHS can better fill the sampling space than MC.

LHS begins with deciding the sample size in total, which is independent to the dimension. This independence means that LHS dose not ask a larger sample size for a higher dimension, which is one of the main advantage of adopting LHS. However, no matter what kind of sampling method is used, the volume of the decision space does not change. The necessary sample size that is able to reflect the landscape of the decision space is independent with the sampling method used. A large amount of samples are still needed in high dimensional decision space even LHS is used. The only thing guaranteed by LHS is that the necessary sample size is no larger than that of equally-spaced or MC sampling.

The theory of the LHS is based on the uniformed distribution, which can be find in that the unit intervals are evenly divided. As a result, it is difficult, though not impossible, to perform LHS in respect to a non-uniformed distribution (figure 4-7). The first thing is that many kinds of distributions, such as normal distribution, are boundaryless. The probable value of these distributions, though the probabilities are extremely low, extend to positive and negative infinite, while LHS needs to divide the interval in the first step. Surely, we can set constraints to these distribute, it is still complicated to divide the interval and sample points in respect to the distributions, as the percent point function (the inverse of cumulative distribution function) is needed.

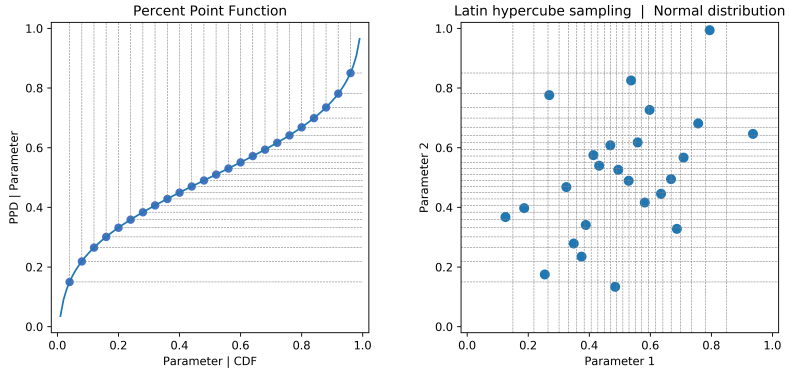


Figure 4-7: LHS in respect to normal distribution, mapped with PPD

Another disadvantage of LHS is that, as the total sample size should be decided in advance, add samples into the collection will break the Latin hypercube. This is not a big problem, as the validness of sample collection will never be worse than that of MC sampling.

In this research, we mainly used LHS with uniform distributions.

4.3 Sensitivity and uncertainty analysis

In this research, sensitivity analysis is used to assess the impact from the parameters on the criteria. The parameters has lower impact, or the less important parameters, were dropped in further study. In this way, the dimension the problem has been reduced.

Uncertainty is another topic in this research. Although the main purpose of this research is to find how to deal with uncertainty, not uncertainty itself, it is still very interesting to find that the concepts of sensitivity and uncertainty is so similar. What's more the methods to do sensitivity analysis and uncertainty analysis are exactly the same. Let's take a close look at the definitions of sensitivity and uncertainty analysis.

Sensitivity	:	The amplitude that a criterion varies with a parameter varying
Uncertainty	:	The uncontrollable or unpredictable variance in a parameter or criteria
Sensitivity analysis	:	Analyzing the relationship between the variety of a parameter and the corresponding variety in a criterion
Uncertainty analysis	:	Analyzing the relationship between the variance in a parameter and the variance caused by this parameter in a criterion

Mathematically, the variety of a variable is usually quantified by its variance or standard deviation. In this context, **there are no essential differences between sensitivity analysis and uncertainty analysis**, besides the purposes are diverse.

The methods of sensitivity analysis can be roughly divided into 2 categories, **OFAT** and **variance-based**.

The **O**ne-**F**actor-**A**t-**O**ne-**T**ime (**OFAT**)[cite] is a method of experiment designing. Rather than test multiple factors simultaneously, OFAT just changes the value of one factor, as indicated by its name, and keeps the values of other factors, then study the variety in the results. The control variables method we studied in the middle school physics class is actually OFTA.

One of the advantages of OFAT is that it is intuit, easy to understand and use. The experiment design is easy. Another advantage is that, if the precision is not so critical, OFAT can provide a nearly qualitative results with a few experiments/observations, though the method itself looks quantitative.

However, duo to the interactions between factors, the results given by OFAT are local results, not global, which are surely influenced by the values of other factors. To get more precise results or avoid missing the global optimal, OFAT requires more runs, which sometimes costs more time and calculation power than variance-based.

The variance-based method[cite] can be regarded as, to some extend, the inverse of OFAT. Using a variance-based method, the value of parameter of interests will be kept constant, or limited into a narrow range, while the values of other parameters are dynamic. The varieties in the variance of results.

Generally speaking, a variance-based method requires special experiment in advance. Compared to OFAT, it usually requires more data to perform a valid analysis using variance based method, so that it is more time-costing. The advantages of the variance-based method are precision and reliability. The analysis results from a variance-based method are usually global. As the values of multiple factors vary simultaneously, the interactions between the factors are considered.

In this research, we have tried both Sobol indices, a typical of variance-based method, and Morris method, a derivative of OFTA, as well as 2 kinds of interactions analysis method expanded from them. Qualitatively, the results from these 2 methods are similar, while Morris method is more efficient. As we aim at the early stage of design, qualitative analysis is sufficient, Morris method can satisfy all of our requirements.

4.3.1 Sobol first-order indices

Sobol indices[cite], named after the Russian mathematician Ilya M. Sobol, is the most typical variance-based sensitivity analysis method. It is so typical that the Wiki page of *Variance-based sensitivity analysis* says 'Variance-based sensitivity analysis (often referred to as the Sobol method or Sobol indices, after Ilya M. Sobol)'.[cite] Sobol indices is a method of global sensitivity analysis by decomposing the variance of results.

Sobol indices can be applied to any any model, especially black boxes, with d-dimensional input space, which is denoted with $y = f(x)$ in this section. x is a d-dimensional input vector (4.3) of this model, with consist of d parameters (4.3).

$$x \in \mathbb{R}^d \tag{4.3}$$

$$x = \{p_1, p_2, \cdots, p_d\} \tag{4.4}$$

Where:

x	:	The input vector of $f(x)$
d	:	The dimension of input space of $f(x)$, the number of parameters in x
p_j	:	The j^{th} parameter in x

The basic idea of Sobol indices can be described as, **if we reduce the variance of a parameter p_j, how much will be reduced in the variance of $f(x)$.**

Sobol indices analysis usually begins with equally-spaced sampling. The unit interval of each dimension is divided evenly into several sub-intervals. The mid point of each sub-interval is used. s denotes the number of points taken in each dimension (4.5). The points in different dimensions are cross referenced so that the total sample size n is s^d 4.7. X denotes the sample collection in total (4.6), while Y stands for the corresponding results (4.8).

$$p_j \in \{p_{j,1}, p_{j,2}, \cdots, p_{j,s}\} \tag{4.5}$$
$$X = \{x_1, x_2, \cdots, x_n\} \tag{4.6}$$
$$n = s^d \tag{4.7}$$
$$Y = \{f(x_1), f(x_2), \cdots, f(x_n)\} \tag{4.8}$$

Where:

$p_{j,i}$	:	The i^{th} sampled point/value of p_j
X	:	The sample collection
n	:	The total sample size
d	:	The number of sampled points in each dimension
Y	:	The collection of $f(x)$

When we want to study the sensitivity of p_j over $f(x)$, the first step is to divide the collection Y into sub-collections based on the value of p_j. The number of the sub-collections is surly s. $(Y|p_j = p_{j,i})$ denotes the sub-collections. The variance of a sub-collection will be calculated and compared with that of the whole collection Y. The reduction proportion of the variance can be regarded as the **local sensitivity** of p_j where its value is equal to $p_{j,i}$ (4.9). The average value of all the local sensitivities is the **global sensitivity** (4.11).

$$LS_{j,i} = \frac{Var(Y) - Var(Y|p_j = p_{j,i})}{Var(Y)} \tag{4.9}$$
$$= 1 - \frac{Var(Y|p_j = p_{j,i})}{Var(Y)} \tag{4.10}$$
$$GS_j = \sum_{i=1}^{s} \frac{LS_{j,i}}{s} \tag{4.11}$$
$$= 1 - \sum_{i=1}^{s} \frac{Var(Y|p_j = p_{j,i})}{s \times Var(Y)} \tag{4.12}$$

Where:

$Var(Y)$	=	The variance of the whole collection Y
$Y\|p_j = p_{j,i}$	=	The sub-collection of $f(x)$ where $p_j = p_{j,i}$
$LS_{j,i}$	:	The local sensitivity of p_j where $p_j = p_{j,i}$
GS_j	:	The global sensitivity of the j^{th} parameter p_j

There is a skill to reduce the calculation cost of GS_j. That is the mean value of difference of variance of the whole collection and sub-collections is equal to the variance of the mean values of each group (4.15). We hope that this sentence does not confuse you. Therefore, the formula of global sensitivity can be simplified into 4.16.

$$Y_{j,i} = (Y|p_j = p_{j,i}) \tag{4.13}$$

$$\{\overline{Y_{j,i}}\} = \{\overline{Y_{j,1}}, \overline{Y_{j,2}}, \cdots, \overline{Y_{j,s}}\} \tag{4.14}$$

$$\because Var(Y) - \overline{Var(Y_{j,i})} = Var(\{\overline{Y_{j,i}}\}) \tag{4.15}$$

$$\therefore GS_j = \frac{Var(\{\overline{Y_{j,1}}\})}{Var(Y)} \tag{4.16}$$

Where:

$\{\overline{Y_{j,i}}\}$: A array consisting of the mean values of sub-collections

$\overline{Var(Y_{j,i})}$: The mean value of the variances of sub-collections

Figure 4-8 illustrate the process of analyzing the sensitivity of p_i using Sobol indices. We can see that the addition of $\overline{Var(Y_{j,i})}$ and $Var(\{\overline{Y_{j,i}}\})$ is equal to the value of $Var(Y)$.

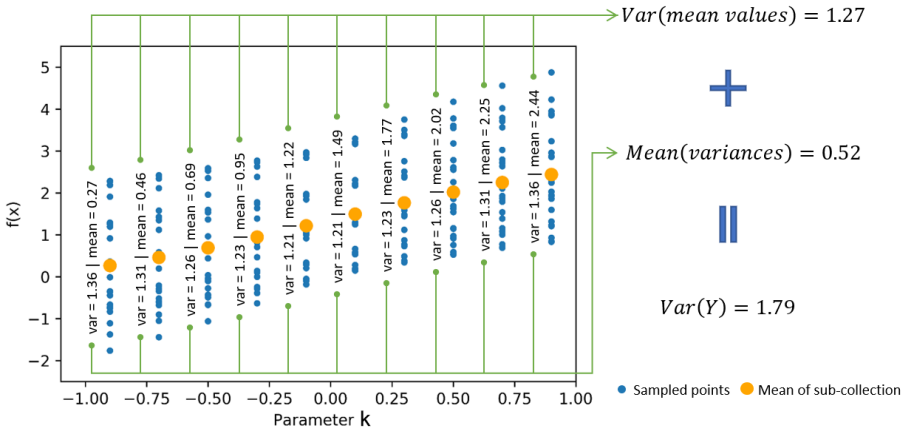


Figure 4-8: A demo of analyzing the sensitivity of p_i using Sobol indices

The calculated global sensitivity GS_j is also called **first-order index**. Actually, the correct sequence is that, because the first-order index indicate the overall sensitivity of p_j over $f(x)$, it is often used as the global sensitivity. The mathematical meaning of first-order index is the proportion decomposed part by p_j of variance of Y, which is more intuitively related to uncertainty analysis. It certainly **also indicated the uncertainty caused by p_j in Y**. Therefore, Sobol indices method is also widely used in uncertainty analysis.

There is a big disadvantage of first-order index. As the global index is a averaged value, it sometimes fail to reflect all the local sensitivities, especially there are both positive and negative values in the local sensitives. Figure 4-9 illustrates a failure sample of first-order index. The value of GS_j is 0, which indicate that p_j has no influence over $f(x)$. However, we can see that the variance of sub-collection changes with the value of p_j.

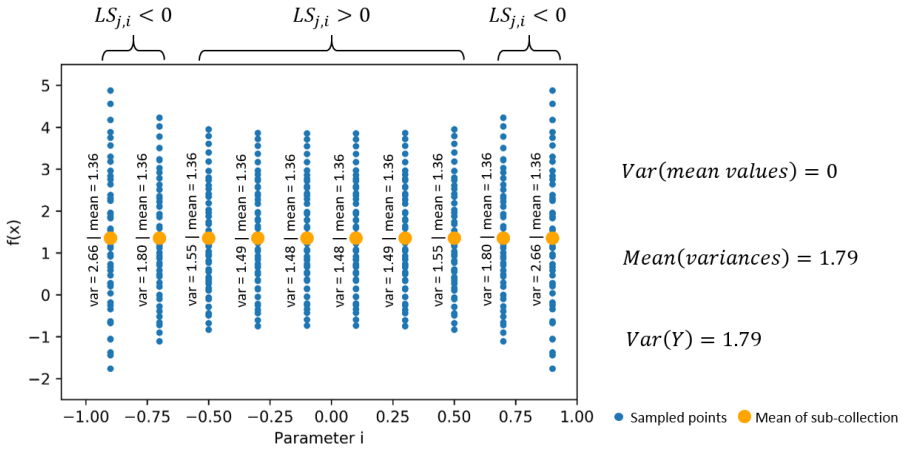


Figure 4-9: A failure sample of first-order index

Inspired by the Morris method that is introduced in the next section. We considered to use the absolute values of local sensitivities, rather than the raw values, to get the global sensitivity of p_j (4.17). In this case, the simplified formula of global sensitivity (4.11) cannot be used. But, as the computers nowadays are so advanced, the time cost difference between calculating variance of means and calculating mean of variances, with even a huge amount of data, is almost negligible.

$$GS_j^{ab} = \sum_{i=1}^{s} \left| \frac{Var(Y) - Var(Y|p_j = p_{j,i})}{s \times Var(Y)} \right| \tag{4.17}$$

Where:

GS_j^{ab} : The global sensitivity of p_j calculated with the absolute values of $LS_{j,i}$

Another disadvantage of the Sobol index method is the curse of of dimensionality. The sample size each sub-collection is equal to s^{d-1}. In high dimensional problems, as the value of d is high, this sample size s^{d-1} is much larger than necessary. A decent variance and a reliable mean value of each sub-collection can be get with a much smaller sample size. Instead of equally-spaced sampling, we considered a new sampling method for Sobol index.

Firstly, points are taken within the unit interval of the j^{th} dimension using LHS, the number of points is s (4.18). Then, for each point $p_{j,i}$, multiple samples of other parameters are generated randomly, the sample size is s_c. These samples are put into a collection (4.19) that is equivalent to concept of the sub-collection in last paragraphs. The following process is the same. The total sample size is $(s \times s_c) \times d$ (4.20).

$$X_j = \{X_{j,0}, X_{j,1}, \cdots, X_{j,s}\} \tag{4.18}$$

$$X_{j,i} = \begin{Bmatrix} x_{j,i,1} \\ \vdots \\ x_{j,i,s_c} \end{Bmatrix} = \begin{Bmatrix} p_{1,1} & \cdots & p_{j,i} & \cdots & p_{d,1} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ p_{1,s_c} & \cdots & p_{j,i} & \cdots & p_{d,s_c} \end{Bmatrix} \tag{4.19}$$

$$n = (s \times s_c) \times d \tag{4.20}$$

4.3.2 Morris method

Morris method, named after Max D. Morris, is an expansion of OFAT method for global sensitivity analysis. The basic idea of Morris method can be explained as, **if we change the value of a parameter p_j a little, how much variety will happen in $f(x)$**, as illustrated in equation 4.24

$$x = p_1, p_2, \cdots, p_j, \cdots, p_d \tag{4.21}$$

$$x + \Delta_j = p_1, p_2, \cdots, p_j + \Delta_j, \cdots, p_d \tag{4.22}$$

$$p_j \in [0, 1] \tag{4.23}$$

$$LS_j = \left| \frac{f(x + \Delta_j) - f(x)}{\Delta_j} \right| \tag{4.24}$$

Where:

x	:	A d-dimensional input vector of $f(x)$, normalized
Δ_j	:	The amount of change in the j^{th} parameter
LS_j	:	The local sensitivity of the j^{th} parameter

If the amount of change in the j^{th} parameter, Δ_j, approach 0, the local sensitivity can be regarded as the absolute of the partial derivative of p_j for the function $f(x)$ 4.25.

$$f'(p_j) = \lim_{\Delta_j \to 0} \frac{f(x + \Delta_j) - f(x)}{\Delta_j} \tag{4.25}$$

However, in the case that the errors in $f(x)$ are not negligible, it is not proper to limit the value of Δ_j to a very small range, otherwise the errors will be extremely enlarged. For example, in this research, we used simulations by *EnergyPlus* as the $f(x)$. *EnergyPlus*, though open sourced, is so complex that it is no different from a black box to us. Sometimes, even the inputs is changed very little, the varieties in outputs are larger than expected. Many researchers use 0.1 for Δ, in the premise that the input vector x is normalized into $[0, 1]^d$.

Another problem is that, similar to the characteristics of partial derivative, the value of local sensitivity LS_j is strongly influenced by the initial value of x. In order to get the global sensitivity of p_j over $f(x)$, researchers use different initial values of x, not only p_j, but also other parameters, to calculate the local sensitivity. The average value the local sensitivities is used as the global sensitivity (4.26).

$$GS_j = \sum_{i=1}^{s} \frac{LS_{j,i}}{s} \tag{4.26}$$

Where:

s	:	The size of initial sample collection
GS_j	:	The global sensitivity of the j^{th} parameter

That is how Morris method evaluate the sensitivity. But, the real skill of Morris method is the experiment design.

The first step of Morris method is to generate a initial sample collection X_0 randomly. If LHS is adapted, the size of this sample collection, s, should be decided in advance.

$$X_0 = \begin{Bmatrix} x_{0,0} \\ \vdots \\ x_{0,i} \\ \vdots \\ x_{0,s} \end{Bmatrix} = \begin{Bmatrix} p_{1,0} & \cdots & p_{j,0} & \cdots & p_{d,0} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ p_{1,i} & \cdots & p_{j,i} & \cdots & p_{d,i} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ p_{1,s} & \cdots & p_{j,s} & \cdots & p_{d,s} \end{Bmatrix} \tag{4.27}$$

Where:

$x_{0,i}$: A randomly generated input vector in this initial sample collection

X_0 : The initial sample collection

The second step is to expend a sample $x_{0,i}$ in the initial collection into $d+1$ samples, make a expansion collection X_i, into which the Δ is gradually added (4.28). In the expansion collection X_i, each sample equal to the addition of the last sample and Δ (4.30). So X_i can be presented as 4.31.

$$x_{j,i} = \{p_{1,i} + \Delta_1, \cdots, p_{j-1,i} + \Delta_{j-1}, p_{j,i} + \Delta_j, p_{j+1,i}, \cdots, p_{d,i}\} \tag{4.28}$$

$$= \{p_{1,i} + \Delta_1, \cdots, p_{j-1,i} + \Delta_{j-1}, p_{j,i}, p_{j+1,i}, \cdots, p_{d,i}\} + \Delta_j \tag{4.29}$$

$$= x_{j-1,i} + \Delta_j \tag{4.30}$$

$$X_i = \begin{Bmatrix} x_{0,i} \\ x_{1,i} \\ \vdots \\ x_{j-1,i} \\ x_{j,i} \\ \vdots \\ x_{d,i} \end{Bmatrix} = \begin{Bmatrix} p_{1,i} & \cdots & p_{j-1,i} & p_{j,i} & \cdots & p_{d,i} \\ p_{1,i} + \Delta_1 & \cdots & p_{j-1,i} & p_{j,i} & \cdots & p_{d,i} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ p_{1,i} + \Delta_1 & \cdots & p_{j-1,i} + \Delta_{j-1} & p_{j,i} & \cdots & p_{d,i} \\ p_{1,i} + \Delta_1 & \cdots & p_{j-1,i} + \Delta_{j-1} & p_{j,i} + \Delta_j & \cdots & p_{d,i} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ p_{1,i} + \Delta_1 & \cdots & p_{j-1,i} + \Delta_{j-1} & p_{j,i} + \Delta_j & \cdots & p_{d,i} + \Delta_d \end{Bmatrix} \tag{4.31}$$

If the all the parameters share the same value of Δ, Xi can also be presented as the addition of $x_{0,i}$ and the production of Δ and a transforming matrix that has $d+1$ rows and d columns (4.32).

$$X_i = x_{0,i} + \Delta \times \begin{bmatrix} 0 & \cdots & 0 & 0 & \cdots & 0 \\ 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 1 & \cdots & 1 & 0 & \cdots & 0 \\ 1 & \cdots & 1 & 1 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 1 & \cdots & 1 & 1 & \cdots & 1 \end{bmatrix} \tag{4.32}$$

Figure 4-10 illustrates the movement, in a 3-dimensional analysis, from $x_{j-1,i}$ to $x_{j,i}$ in X_i.

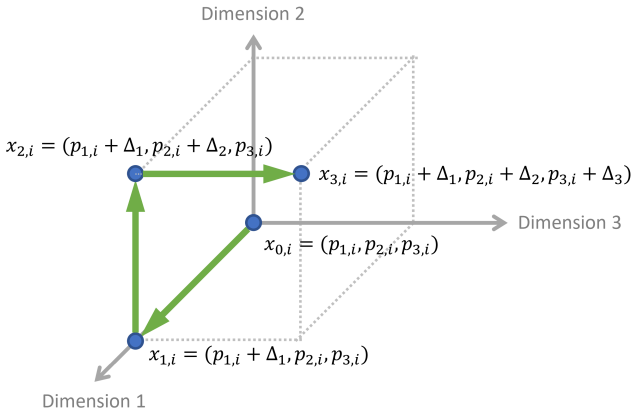


Figure 4-10: Movements of samples in Morris method

$x_{j-1,i}$ and $x_{j,i}$ are use to get the local sensitivity of p_j (4.33), as the difference between them is the Δ_j (4.34). The average value of all local sensitivities from different collections X_i are used as the global sensitivity (4.26).

$$LS_{j,i} = \left| \frac{f(x_{j,i}) - f(x_{j-1,i})}{\Delta_j} \right| \tag{4.33}$$

$$= \left| \frac{f(x_{j-1,i} + \Delta_j) - f(x_{j-1,i})}{\Delta_j} \right| \tag{4.34}$$

Figure 4-11 demonstrates the sensitivity analysis of a certain p_i. The blue points stands for the samples $\{x_{j-1,0}, \cdots, x_{j-1,i} \cdots, x_{j-1,s}\}$ from all the collections $\{X_1, \cdots, X_i \cdots, X_s\}$, which the orange points stands for the corresponding $\{x_{j,i}\}$. The points from the same collection X_i are linked using a green line. The absolute value of these lines is the local sensitivities.

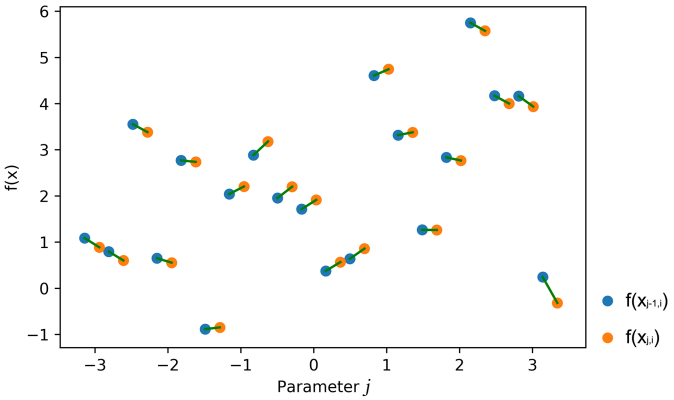


Figure 4-11: morris demo

One of the advantages of the Morris method is its high efficiency. The total calculation need by Morris method is $s \times (d+1)$, which is much more less than that needed by Sobol indices in most cases 4.35.

$$n = s \times (d+1) < (s \times s_c) \times d \ll s^d \tag{4.35}$$

Another advantage of this experiment design used by Morris method is that, the weight of each sample in X_i is similar. The first and last sample are used once, and other samples are used twice, in the calculations of local sensitivities. Let's make a comparison. As illustrated in figure 4-12, if we used $x_{0,i}$ and $x_{0,i} + \Delta$ to calculate the local sensitivity, the initial sample $x_{0,i}$ are used d times, while the other samples are used just once. The initial samples are over-weighted so that the results take too much influence from them.

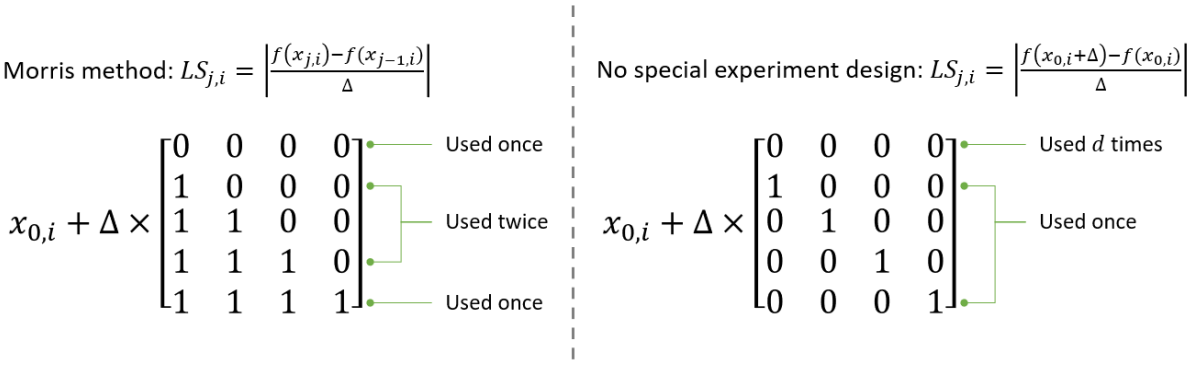


Figure 4-12: The advantage of the experiment design of Morris method

As the global sensitivity is an averaged values, similar to Sobol indices, this averaged values sometimes cannot faithfully reflect all the local sensitivities. In some cases, though the mean value of $\{LS_j\}$ is low, one or two values in this array might be high, so that the influence from p_j is not negligible. The variance of an array is sensitive to those extremely high or low members. A relative high variance of local sensitivities also indicates that the sensitivity of p_j is influenced by the values of other parameters, which also hints the interactions. Therefore, some researchers also take the variance of local sensitivities into consideration, reforming the formula of global sensitivity into equation 4.36.

$$GS_j^{mv} = \sqrt{\overline{LS_j}^2 + Var(LS_j)} \tag{4.36}$$

$$Var(LS_j) = \sum_{i=1}^{s} \frac{(LS_{j,i} - \overline{LS_j})^2}{s} \tag{4.37}$$

Where:

GS_j^{mv}	:	Global sensitivity of p_j considering both mean value and variance
$\overline{LS_j}$	:	The mean value of $\{LS_j\}$
$Var(LS_j)$	:	The variance of $\{LS_j\}$

4.4 Interaction analysis

If we say that the purpose of sensitivity analysis is to study the relationship between a parameter and a criterion, then it can be said that purpose of interaction analysis is to study the relationship between 2 parameters. The definition of the interaction between 2 parameters, $\boldsymbol{p_j}$ and $\boldsymbol{p_{j'}}$, can be defined as, **if the value of $\boldsymbol{p_{j'}}$ influences the relationship between $\boldsymbol{p_j}$ and $\boldsymbol{f(x)}$, then $\boldsymbol{p_j}$ and $\boldsymbol{p_{j'}}$ are interactive.**Generally speaking, the relationship between p_j and $p_{j'}$ is represented by the sensitivity of p_j.

Figure 4-13 is a simply illustration of the definition of interaction. In the first diagram, the function $f(x)$ is the addition of p_1 and p_2, the slope of the line $f'(p_1)$ do not varies with the value of p_2. So that there is no interaction between p_1 and p_2 in this case. In function $f(x)$ in the second diagram is the production of p_1 and p_2, the slope of the line $f'(p_1)$ is exactly equal to p_2. The sensitivity of p_1 over $f(x)$ is decided by the value of p_2. In this case, p_1 and p_2 is strongly interactive with each other. Surely, it more than true or false in the real world. In the third diagram, the gray dash line and the blue line ($f'(p_1|p_2 = 0.1)$) are in parallel. The angle between the gray line and orange line ($f'(p_1|p_2 = 0.4)$) is quite small. The slope of $f'(p_1)$ line is slightly influenced by the value of p_2. In this case, we can say that there is a weak interaction between p_1 and p_2.

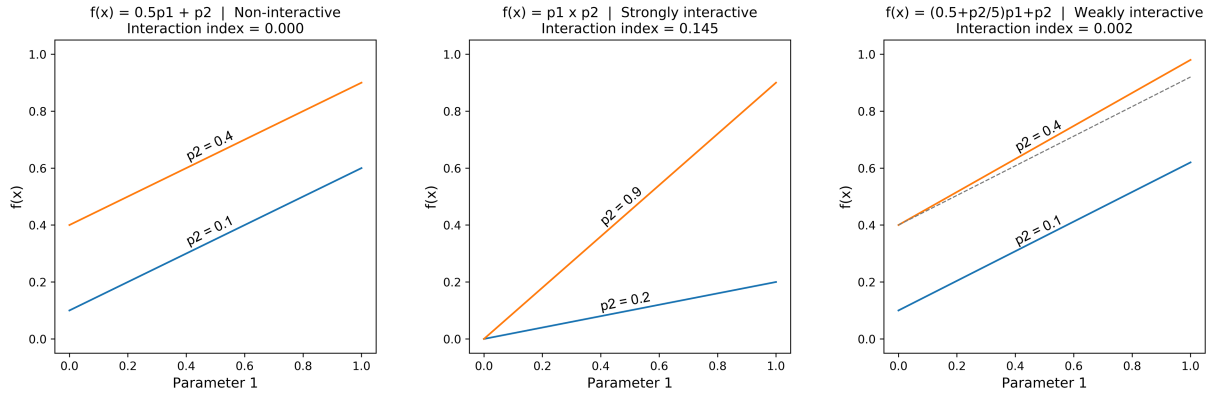


Figure 4-13: Interactions between p_1 and p_2

The methods of interaction analysis are the extensions of that of sensitivity analysis. Some researchers call the interaction analysis higher-order sensitivity analysis. There are also variance-based and OFAT methods for interaction analysis. In this research, we studied the Sobol second-order indices and Expanded Morris method.

4.4.1 Sobol second-order indices

Sobol second-order indices method is similar to Sobol first-order indices decried in section 4.3.1. Compared to the first-order indices that limit the range of a single parameter p_j, the basic idea seconded-order indices is to **limit the ranges of 2 parameters, p_j and $p_{j'}$, simultaneously, then observe the variety in the variance of $f(x)$.** By comparing the second-order index of $\{p_j, p_{j'}\}$ and their respective first-order indices (sensitivities), the interaction between p_j and $p_{j'}$ can be quantified.

The experiment design of Sobol second-order indices method is identical to that of the first-order indices (4.5 to 4.8). Then the whole collection Y is divided into sub-collections base on a cross referenced matrix of p_j and $p_{j'}$, denoted by $(Y|p_j = p_{j,i}, p_{j'} = p_{j',i'})$ or $Y_{(j,i),(j',i')}$ (4-8). The number of the sub-collections is s^2, equal to the square of the number of points taken in each dimension. Each sub-collection includes $s^(d-2)$ samples.

$$\begin{Bmatrix} Y_{(j,0),(j',0)} & \cdots & Y_{(j,i),(j',0)} & \cdots & Y_{(j,s),(j',0)} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ Y_{(j,0),(j',i')} & \cdots & Y_{(j,i),(j',i')} & \cdots & Y_{(j,s),(j',i')} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ Y_{(j,0),(j',s)} & \cdots & Y_{(j,i),(j',s)} & \cdots & Y_{(j,s),(j',s)} \end{Bmatrix} \tag{4.38}$$

Likewise, the difference between the variance of a sub-collection $Y_{(j,i),(j',i')}$ and the whole collection Y is used to represent the local second-order index of p_j and $p_{j'}$ (4.39). The average value of the local second-order indices is the global second-order index of p_j and $p_{j'}$ (4.41).

$$LSoi_{(j,i),(j',i')} = \frac{Var(Y) - Var(Y_{(j,i),(j',i')})}{Var(Y)} \tag{4.39}$$

$$= 1 - \frac{Var(Y_{(j,i),(j',i')})}{Var(Y)} \tag{4.40}$$

$$GSoi_{j,j'} = \sum_{i=1,i'=1}^{s,s} \frac{LSoi_{(j,i),(j',i')}}{s^2} \tag{4.41}$$

Where:

$LSoi_{(j,i),(j',i')}$: The local second-order index of p_j and $p_{j'}$ where $p_j = p_{j,i}, p_{j'} = p_{j',i'}$

$GSoi_{jj'}$: The global second-order index of p_j and $p_{j'}$

Identical the first-order indices, there is also a simplified formula of the global second-order indices 4.43.

$$\because Var(Y) - \overline{Var(Y_{(j,i),(j',i')})} = Var(\{\overline{Y_{(j,i),(j',i')}}\}) \tag{4.42}$$

$$\therefore GSoi_{jj'} = \frac{Var(\{\overline{Y_{(j,i),(j',i')}}\})}{Var(Y)} \tag{4.43}$$

The next the step is to get the Sobol first indices of both p_j and $p_{j'}$, using the method described in section 4.3.1, then compare their addition with second-order index to get the **interaction index** (4.4.1).

$$It_{jj'} = GSoi_{jj'} - (GS_j + GS_{j'}) \tag{4.44}$$

Where:

$GS_j, GS_{j'}$	:	The global sensitivities of p_j and $p_{j'}$
$It_{jj'}$	:	The global interaction index of p_j and $p_{j'}$

A positive interaction index indicates the interaction between 2 parameters. The higher the $It_{jj'}$ index is, the stronger the interaction is. Normally, if the $It_{jj'}$ is larger than 0.1, the p_j and $p_{j'}$ can be regarded as strongly interactive.

Similar to Sobol first-order index, the sample size of each sub-collection, $s^(d-2)$, is much larger than necessary. We also considered a new sampling method for Sobol second-order indices. Firstly, paired samples of p_j and $p_{j'}$ are generated using LHS, the number of samples is s' (4.45). Then, for each pair $\{p_{j,i}, p_{j',i}\}$, multiple samples of other parameters are generated randomly, the sample size is s_c. These samples are put into a collection (4.46) that is equivalent to concept of the sub-collection. As there are $d(d-1)$ pairs of parameters, the total sample size to get the second-order indices of all pairs of parameters is $(s' \times s_c) \times d(d-1)$. Besides, the first-order indices of all the parameters are also needed, which require $(s \times s_c) \times d$ samples. So the necessary sample size in total is $(s' \times s_c) \times d(d-1) + (s \times s_c) \times d$ (4.47).

$$X_{jj'} = \{X_{jj',0}, X_{jj',1}, \cdots, X_{jj',s'}\} \tag{4.45}$$

$$X_{jj',i} = \begin{Bmatrix} x_{jj',i,1} \\ \vdots \\ x_{jj',i,s_c} \end{Bmatrix} = \begin{Bmatrix} p_{1,1} & \cdots & p_{j,i} & \cdots & p_{j',i} & \cdots & p_{d,1} \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ p_{1,s_c} & \cdots & p_{j,i} & \cdots & p_{j',i} & \cdots & p_{d,s_c} \end{Bmatrix} \tag{4.46}$$

$$n = (s' \times s_c) \times d(d-1) + (s \times s_c) \times d \tag{4.47}$$

$$= (s_c \times d)[s'(d-1) + s \times d] \tag{4.48}$$

Where:

$X_{jj'}$	:	The whole sample collection to analysis the global second-order index of p_j and $p_{j'}$
$X_{jj',i}$	:	The sub-collection to analysis the local second-order index of p_j and $p_{j'}$, where $p_j = p_{j,i}, p_{j'} = p_{j',i}$
n	:	The number of samples needed to get the interaction index of p_j and $p_{j'}$

4.4.2 Expanded Morris method

Expanded Morris method, as its name indicates, is an interaction analysis method based on Morris method. The basic idea of Expanded Morris method can be explained as, if we change the value of $p_{j'}$, whether the sensitivity of p_j over $f(x)$ varies or not. If the value of $p_{j'}$ influence the sensitivity of p_j, then p_j and $p_{j'}$ are interactive.

Equation 4.49 shows the formula of local interaction between p_j and $p_{j'}$. Like Morris method, this local interaction index is influenced by the initial value of the input vector x. So that researchers usually generate several initial values for x, then use the average value of $LIt_{jj',i}$ as the global interaction index (4.50).

$$LIt_{jj',i} = \frac{|f(x_i + \Delta_j + \Delta_{j'}) - f(x_i + \Delta_j) - f(x_i + \Delta_{j'}) + f(x_i)|}{\Delta^2} \tag{4.49}$$

$$It_{jj'} = \sum_{i=1}^{s} \frac{Lit_{jj',i}}{s} \tag{4.50}$$

Where:

$LIt_{jj',i}$: The local interaction index of p_j and $p_{j'}$

$It_{jj'}$: The global interaction index of p_j and $p_{j'}$

If we adjust the polynomial a bit, changing the sequence of the items in numerator, we will find that this interaction index $LIt_{jj',i}$ indicates the variety in the local sensitivity of p_j when the value of $p_{j'}$ changes from $p_{j',i}$ to $p_{j',i} + \Delta$ (4.53). Identically, it also indicate the variety in $LS_{j',i}$ (4.55).

$$LIt_{jj',i} = \frac{|f(x_i + \Delta_j + \Delta_{j'}) - f(x_i + \Delta_j) - f(x_i + \Delta_{j'}) + f(x_i)|}{\Delta^2} \tag{4.51}$$

$$= \frac{f(x_i + \Delta_j + \Delta_{j'}) - f(x_i + \Delta_j)}{\Delta^2} - \frac{f(x_i + \Delta_{j'}) - f(x_i)}{\Delta^2} \tag{4.52}$$

$$= \frac{(LS_{j,i}|p_{j'} = p_{j',i} + \Delta) - (LS_{j,i}|p_{j'} = p_{j',i})}{\Delta} \tag{4.53}$$

$$= \frac{f(x_i + \Delta_j + \Delta_{j'}) - f(x_i + \Delta_{j'})}{\Delta^2} - \frac{f(x_i + \Delta_j) - f(x_i)}{\Delta^2} \tag{4.54}$$

$$= \frac{(LS_{j',i}|p_j = p_{j,i} + \Delta) - (LS_{j',i}|p_j = p_{j,i})}{\Delta} \tag{4.55}$$

Where:

$LS_{j,i}$: The local sensitivity of p_j

$LS_{j',i}$: The local sensitivity of $p_{j'}$

Figure 4-14 illustrate the relationship of sensitivity analysis using Morris method and interaction analysis using Expanded Morris method. We can see that the interaction is actually the difference in the sensitivities of one parameter caused by the variety of the other parameter.

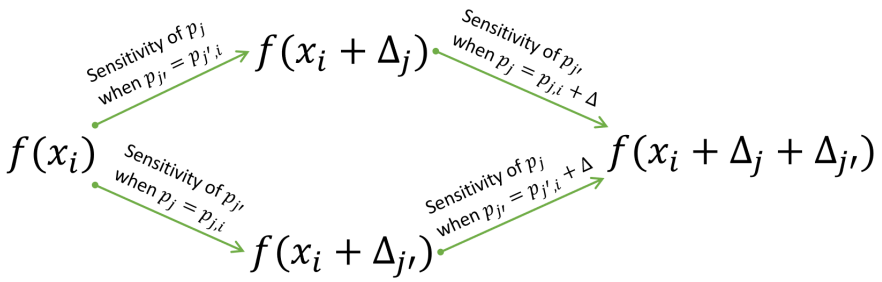


Figure 4-14: Movements of samples in Expanded Morris method

The experiment design of Expanded Morris method is also similar to that of Morris method but much more complex, which is a little difficult to explain. In a d-dimensional decision space, there are $d(d-1)/2$ pairs of parameters. For each pair, 4 samples are needed to get the interaction index. However, the total sample size needed for all pairs is not $4d(d-1)/2$, but $d^2/2 + d/2 + 1$. Let's see how it comes.

The first step in identical to that in Morris method, generating a collection including s initial samples using LHS (4.56).

$$X_0 = \begin{Bmatrix} x_{0,0} \\ \vdots \\ x_{0,i} \\ \vdots \\ x_{0,s} \end{Bmatrix} = \begin{Bmatrix} p_{1,0} & \cdots & p_{j,0} & \cdots & p_{j',0} & \cdots & p_{d,0} \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ p_{1,i} & \cdots & p_{j,i} & \cdots & p_{j',i} & \cdots & p_{d,i} \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ p_{1,s} & \cdots & p_{j,s} & \cdots & p_{j',s} & \cdots & p_{d,s} \end{Bmatrix} \tag{4.56}$$

Where:

$x_{0,i}$: A randomly generated input vector

X_0 : The initial sample collection

For each pair, the initial sample is expanded into 4 samples with the transform matrix in equation 4.58. The items in columns from 1^{st} to $j-1^{th}$, and from $j+1^{th}$ to $j'-1^{th}$, are all 1, while those in columns since $j'+1^{th}$ are all 0. The j^{th} column is $[0,1,0,1]^T$ and the j^{th} column is $[0,0,1,1]^T$.

$$X_i^{jj'} = \begin{Bmatrix} p_{1,i}+\Delta, & \cdots, & p_{j,i}, & p_{j+1,i}+\Delta, & \cdots, & p_{j',i}, & \cdots, & p_{d,i} \\ p_{1,i}+\Delta, & \cdots, & p_{j,i}+\Delta, & p_{j+1,i}+\Delta, & \cdots, & p_{j',i}, & \cdots, & p_{d,i} \\ p_{1,i}+\Delta, & \cdots, & p_{j,i}, & p_{j+1,i}+\Delta, & \cdots, & p_{j',i}+\Delta, & \cdots, & p_{d,i} \\ p_{1,i}+\Delta, & \cdots, & p_{j,i}+\Delta, & p_{j+1,i}+\Delta, & \cdots, & p_{j',i}+\Delta, & \cdots, & p_{d,i} \end{Bmatrix} \tag{4.57}$$

$$= x_{0,i} + \Delta \times \begin{bmatrix} \boldsymbol{1^{st}} & \cdots & \boldsymbol{j^{th}} & \boldsymbol{j+1^{th}} & \cdots & \boldsymbol{j'-1^{th}} & \boldsymbol{j'^{th}} & \cdots & \boldsymbol{d^{th}} \\ 1 & \cdots & 0 & 1 & \cdots & 1 & 0 & \cdots & 0 \\ 1 & \cdots & 1 & 1 & \cdots & 1 & 0 & \cdots & 0 \\ 1 & \cdots & 0 & 1 & \cdots & 1 & 1 & \cdots & 0 \\ 1 & \cdots & 1 & 1 & \cdots & 1 & 1 & \cdots & 0 \end{bmatrix} \tag{4.58}$$

In this way, we can find that, for the first pair $(1,2)$, 4 samples are needed. For pairs from $(1,3)$ to $(1,d)$, 2 additional sample are needed for each, totally $2(d-2)$. Then, for the rest pair, just one additional sample is needed for each, totally $d(d-1)/2-(d-1)$. As there are s samples in the initial collection, after rearranging, the total samples needed for all pairs is $s[d^2/2+d/2+1]$ (4.61).

$$n = s \times [4 + 2(d-2) + \frac{d(d-1)}{2} - (d-1)] \tag{4.59}$$

$$= s \times [2 + 2(d-1) + \frac{d(d-1)}{2} - (d-1)] \tag{4.60}$$

$$= s \times (\frac{d^2}{2} + \frac{d}{2} + 1) \tag{4.61}$$

Figure 4-15 illustrates how samples are reused in Expanded Morris method, with a 5-dimensional example.

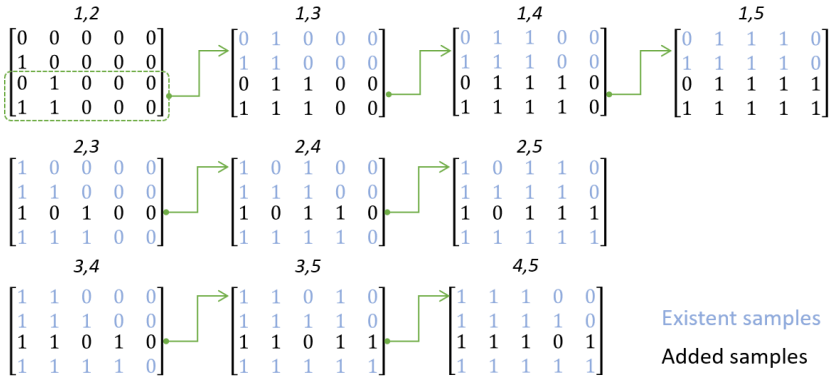


Figure 4-15: Reusage of samples in

Figure 4-16 is a 3-dimensional illustration of the movements of samples. 7 samples are enough.

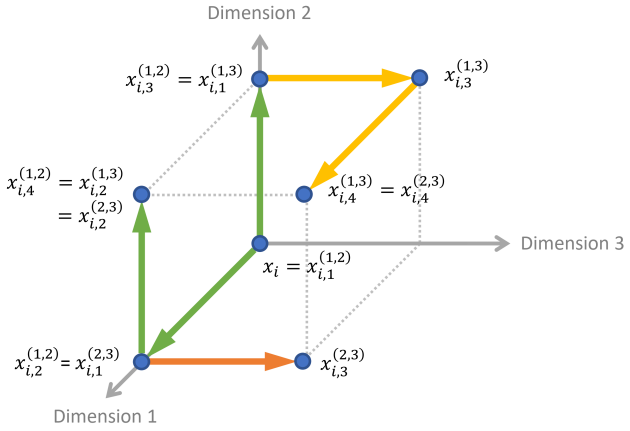


Figure 4-16: Movemrnts of samples in Expanded Morris method

4.5 Comparison between Sobol indices method and Morris method

In this section, Ishigami function (4.63) is used to test Sobol indices and Morris methods and demonstrate sensitivity and interaction analysis. Ishigami function, firstly proposed by T. Ishigami[2] and T. Homma[3] from JAEA[4] in 1990[5], is a 3-dimensional (4.62) function along with 3 coefficients. Ishigami faction is widely used by researchers in testing sensitivity interaction analysis methods.

$$x_i = \{p_{1i}, p_{2i}, p_{3i}\} \in \mathbb{R}^3 \tag{4.62}$$

$$f(x_i) = \sin(p_{1i}) + a\sin^2(p_{2i}) + bp_{3i}^4 \sin(p_{1i}) + c \tag{4.63}$$

4.5.1 Testing sensitivity analysis methods

Firstly, we tested Sobol first-order indices with $a = 7, b = 0.1, c = 0$ (4.64), the values of p_1, p_2, p_3 are within $[-\pi, \pi]$ (4.65). Equally-spaces sampling method is used to generate the sample collection. In the first step, 10 points were taken from each unit interval. The function $f(x)$ has been calculated 1000 times in total.

$$f(x) = \sin(p_1) + 7\sin^2(p_2) + 0.1p_3^4 \sin(p_1) \tag{4.64}$$

$$p_1, p_2, p_3 \in [-\pi, \pi] \tag{4.65}$$

Table 4-1 shows the sensitivity analysis results using Sobol indices method. The GS_3 is equal to 0, which is not reasonable. The range of p_3 is $[-\pi, \pi]$, which is wider than that of $\sin(p_1)$ which is $[-1, 1]$, while the power of p_3 is 4. The value of p_3 surely influences the result of $f(x)$. We also tried GS_3^{ab}, the mean value of the absolute values of local sensitivities, which showed a reasonable result.

Table 4-1: First trail of first-order indices analysis

	GS	GS^{ab}
p_1	0.316	0.316
p_2	0.465	0.465
p_3	0.000	0.615

Then we adjusted s (4.66), the number of points taken from each dimension, to see the influence from the sample size to the analysis results.

$$s_i = 2^i \tag{4.66}$$

$$i \in \{1, 2, 3, 4, 5, 6\} \tag{4.67}$$

[2]石神努

[3]本間俊充

[4]Japan Atomic Energy Agency, 日本原子力研究開発機構

Table 4-2 shows the relationship between the results of sensitivity analysis and s_i. Commonsensically, the results gradually get stable with s_i arising. In this case, the varieties in the results become quite small after s_i passing 8.

Table 4-2: Sample size and first-order indices

	2	4	8	16	32	64
p_1	1.0	0.739	0.316	0.315	0.314	0.314
p_2	0.0	0.000	0.477	0.451	0.445	0.443
p_3	0.0	0.879	0.616	0.648	0.656	0.658

The second step was to test Morris method. The same function ($a = 7, 6 = 0.1, c = 0$) was used. The size of the initial sample collection s is 20. Table 4-3 shows the results of sensitivity analysis. Both GS, the global sensitivity considering only mean value of local sensitivities, and GS^{mv}, the global sensitivity considering both mean and variance of local sensitivities, were calculated. Compared to the results by Sobol indices, the sensitivity of p_3 might be underestimated by GS. The value of GS^{mv} seems more reasonable.

Table 4-3: First trail of Morris method

	GS	GS^{mv}
p_1	10.095	14.763
p_2	25.705	28.689
p_3	16.849	29.744

The influence from the size of initial collection s was also studied. The s was adjusted from 2 to 256. As Morris method is a Monte-Carlo-based, even with the same configurations, the results vary from time to time. It is difficult to quantify the errors in the results in a single run. Therefore, for each s_i, we run the sensitivity analysis 10 times and calculated the variance.

Table 4-4 shows the relationship between the size of initial sample collection and the variance of the global sensitivities of the same s_i calculated using Morris method. With the sample size increasing, the variances reduced, which indicated that the results gradually converged.

Table 4-4: Initial sample size and variances of Morris method results

	2	4	8	16	32	64	128	256
p_1	111.08	56.51	10.51	5.34	1.39	0.71	0.60	0.30
p_2	87.16	24.34	1.88	0.88	0.26	0.02	0.00	0.00
p_3	179.31	125.07	41.58	14.41	10.09	6.14	0.45	1.16

Figure 4-17 illustrates that the results converges with the sample size increasing. We can see that, after the sample size passing 32, the variance gets enough small for comparing the sensitivities between parameters.

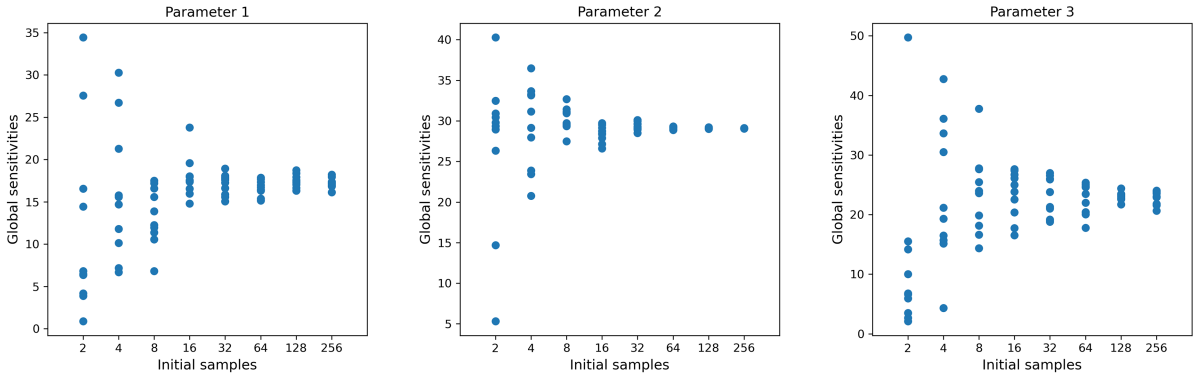


Figure 4-17: Initial sample size and Morris method results

Finally, we compared the results from Sobol indices and Morris method. GS^{ab} and GS^{mv} were used. For Sobol indices, the number of points taken from each unit interval is 10. For Morris method, the size of initial sample collection is 20. The ranges of parameters and the values of a and b were also adjusted. Table 4-5 shows the results. The values of the sensitivities are unit-less and hard to read. We normalized the results that the addition of each row is equal to 1 (table 4-6).

From the normalized results, we can see that the sensitivities, of the same parameter in the same circumstance, from different method, are close to each other.Therefore, it can be said that both Sobol indices and Morris method are effective methods. Morris method can get reliable results with a smaller sample size.

Table 4-5: Raw results of sensitivity analysis

		a=7, b=0.1		**a=5, b=0.2**			**a=7, b=0.1**		**a=5, b=0.2**	
		Sobol	Morris	Sobol	Morris		Sobol	Morris	Sobol	Morris
p_1	$[-\pi, \pi]$	0.078	4.387	0.147	4.507	$[-2\pi, 2\pi]$	0.121	12.409	0.276	13.342
p_2	$[-\pi, \pi]$	0.922	27.723	0.853	21.686	$[-\pi, \pi]$	0.868	30.184	0.658	20.492
p_3	$[-1, 1]$	0.003	0.243	0.012	0.485	$[-2, 2]$	0.067	3.971	0.268	8.328

Table 4-6: Normalized results of sensitivity analysis

		a=7, b=0.1		**a=5, b=0.2**			**a=7, b=0.1**		**a=5, b=0.2**	
		Sobol	Morris	Sobol	Morris		Sobol	Morris	Sobol	Morris
p_1	$[-\pi, \pi]$	0.078	0.136	0.145	0.169	$[-2\pi, 2\pi]$	0.115	0.266	0.230	0.316
p_2	$[-\pi, \pi]$	0.919	0.857	0.843	0.813	$[-\pi, \pi]$	0.822	0.648	0.547	0.486
p_3	$[-1, 1]$	0.003	0.008	0.012	0.018	$[-2, 2]$	0.063	0.085	0.223	0.198

4.5.2 Testing interaction analysis methods

We test the interactions between parameters in Ishigami function using both Sobol second-order indices and Expanded Morris method (4-7). As in Ishigami function, only p_1 and p_3 are weakly interactive, this is not a good demonstration for interaction analysis.

Table 4-7: Interactions between parameters in Ishigami function

	$\{p_1, p_2\}$	$\{p_1, p_3\}$	$\{p_2, p_3\}$
Sobol	0.00	4.59×10^{-5}	0.00
Morris	0.00	7.37×10^{-1}	0.00

Therefore, we defined a new function (4.68) to test the interaction analysis methods.

$$f(x) = p_1 \times p_2 + 2 \times p_1 \times p_4^2 + \frac{0.1 * \sin(p_3) * p_5}{p_2} \tag{4.68}$$

Table 4-8 shows the results of interaction analysis. Both Sobol second-order indices and Expanded Morris method were able to correctly reflect the interactions between parameters, though the analyzed interaction strengths for each pair were different. Expanded Morris method used much less samples to get the results, compared to Sobol second-order indices.

Table 4-8: Results of interaction analysis

	$\{p_1, p_2\}$	$\{p_1, p_3\}$	$\{p_1, p_4\}$	$\{p_1, p_5\}$
Sobol	2.58×10^{-2}	0.00	1.09×10^{-1}	0.00
Morris	1.00	0.00	2.20	0.00
	$\{p_2, p_3\}$	$\{p_2, p_4\}$	$\{p_2, p_5\}$	
Sobol	1.73×10^{-2}	0.00	1.99×10^{-2}	
Morris	1.34	0.00	1.10	
	$\{p_3, p_4\}$	$\{p_3, p_5\}$		
Sobol	0.00	3.46×10^{-3}		
Morris	0	1.96×10^{-1}		
	$\{p_4, p_5\}$			
Sobol	0.00			
Morris	0.00			

4.6 Prediction combination polynomial

Theoretically, the parameters in different groups can be studied independently. We can trained independent GPR models for different groups (equation 4.69).

$$Pd_i(g_i) \approx f(g_i) \tag{4.69}$$

Where:

$Pd_i(g_i)$ = The prediction made by a trained GPR model for the i^{th} group

$f(g_i)$ = $f(X)$ that the parameters not belonging to the i^{th} group are constants

Figure 4-18 illustrates the flow of training GPR for a single group. Firstly, we generate samples of the parameters in the group that we want to study. The other parameters are set to constant values, which are 0.5 in this illustration. Then calculations/simulations are executed. The sampled parameters in this group and the calculation results are used to train a GPR model. We can get predictions when input new values of the parameters in this group into the trained GPR model.

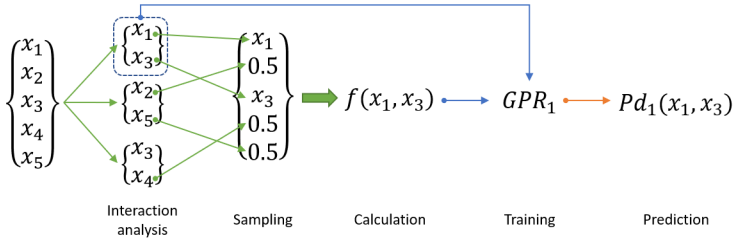


Figure 4-18: Training GPR for a single group

However, in some cases, we have to study parameters from different groups together, especially when cooperate with someone not familiar with interaction principles. Of course, we can put those parameters from different group together and train a GPR model. But, at the same time, we are facing the risks of high time cost and poor training. What's more, if we put them together, why should we do the interaction analysis.

To solve this problem, **P**rediction **C**ombination **P**olynomial (**PCP**) was introduced into this research. Here is the mathematical deduction.

Theoretically, the function $f(X)$ can be split into several sub-fuctions (equation 4.70). The inputs of each sub-faction f_i is the parameters in the i^{th} group. So that $f(g_i)$ is equal to the addition of $f_i(g_i)$ and a constant (4.71).

$$f(X) = f_1(group_1) + f_2(group_2) + \cdots + f_n(g_n) \tag{4.70}$$
$$f(g_i) = f_1(c_1) + \cdots + f_i(g_i) + \cdots + f_n(g_n) \tag{4.71}$$

However, it is impossible for ue to know the exact value of the sub-fuctions $f_i(g_i)$ or $f_j(g_j)$. What we can get is just $f(g_i)$ and $f(g_j)$. We need a few mathematical skills to convert $f(g_i)$ and $f(g_j)$ into $f(g_i, g_j)$, by adding a correction item $-f(c_{all})$ between them. Equation 4.72 shows the deduction process.

$$\begin{aligned} f(g_i, g_j) =& f_1(c_1) + \cdots + f_i(g_i) + \cdots + f_j(g_j) + \cdots + f_n(g_n) \\ =& [f_1(c_1) + \cdots + f_i(g_i) + \cdots + f_j(c_j) + \cdots + f_n(g_n)] \\ & - [f_1(c_1) + \cdots + f_i(c_i) + \cdots + f_j(c_j) + \cdots + f_n(g_n)] \\ & + [f_1(c_1) + \cdots + f_i(c_i) + \cdots + f_j(g_j) + \cdots + f_n(g_n)] \\ =& f(g_i) - f(c_{all}) + f(g_j) \end{aligned} \tag{4.72}$$

Where:
$f(c_{all})$ = $f(X)$ that all parameters are constants

In the case that there are same common parameters shared by the i^{th} and j^{th} (4.73), the correction item should be $-f(g_i \cap g_j)$ (4.77). Actually, if the $g_i \cap g_j$ is empty, then $f(g_i \cap g_j)$ is identical to $f(c_{all})$.

$$g_i \cap g_j \neq \varnothing \tag{4.73}$$
$$f(g_i) = \cdots + f_i(g_i) + \cdots + f_j(g_i \cap g_j, c_j^*) + \cdots \tag{4.74}$$
$$f(g_j) = \cdots + f_i(g_i \cap g_j, c_i^*) + \cdots + f_j(g_j) + \cdots \tag{4.75}$$
$$f(g_i \cap g_j) = \cdots + f_i(g_i \cap g_j, c_i^*) + \cdots + f_j(g_i \cap g_j, c_j^*) + \cdots \tag{4.76}$$
$$f(g_i, g_j) = f(g_i) - f(g_i \cap g_j) + f(g_j) \tag{4.77}$$

Based on this combination polynomial, we can combine the predictions from 2 GPR models trained independently for 2 groups (4.81). Because there are always some error in the GPR models, more all less, $Pd_i(g_i \cap g_j)$ is not exactly equal to $Pd_j(g_i \cap g_j)$. We use their average to stand for $f(g_i \cap g_j)$ (4.80).

$$f(g_i) \approx Pd_i(g_i) \tag{4.78}$$
$$f(g_j) \approx Pd_j(g_j) \tag{4.79}$$
$$f(g_i \cap g_j) \approx \frac{Pd_i(g_i \cap g_j) + Pd_j(g_i \cap g_j)}{2} \tag{4.80}$$
$$f(g_i, g_j) \approx Pd_i(g_i) + Pd_j(g_j) - \frac{Pd_i(g_i \cap g_j) + Pd_j(g_i \cap g_j)}{2} \tag{4.81}$$

If we want to involve a third group, we just need to add a correction item (4.82). The deduction process is no different from above.

$$f(g_i, g_j, g_k) = f(g_i) - f(g_i \cap g_j) + f(g_j) - f(g_j \cap g_k) + f(g_k) \tag{4.82}$$

Equation 4.83 is the general formula to combine n groups. Likewise, the general formula to combine the predictions of n groups (4.84) is in the similar form.

$$f(X*) = \sum_{i=1}^{n-1} (f(g_i) - f(g_i \cap g_{i+1})) + f(g_n) \tag{4.83}$$

$$Pd(X*) = \sum_{i=1}^{n-1} \left(Pd_i(g_i) - \frac{Pd_i(g_i \cap g_j) + Pd_{i+1}(g_i \cap g_{i+1})}{2} \right) + Pd_n(g_n) \tag{4.84}$$

Figure 4-19 is an illustration of the flow adapting PCP. The parameters are separated into 3 groups based on their interactions. The 1^{st} and 3^{rd} groups are picked up to train GPR models. Finally, the predictions from the trained GPR models are combined using PCP method.

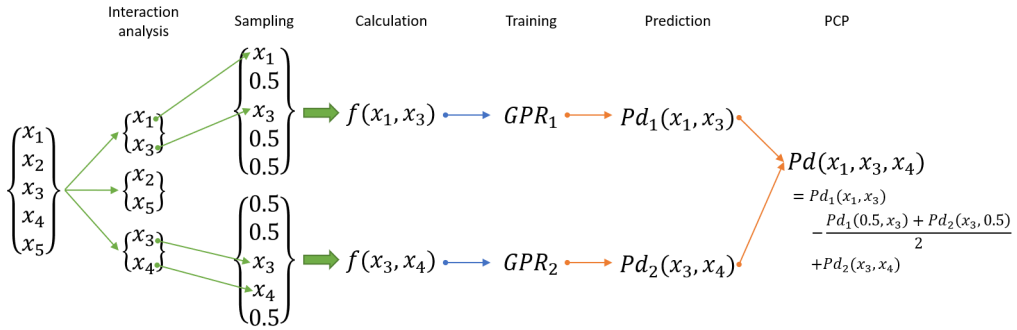


Figure 4-19: Training GPR for a single group

Certainly, this PCP method is also suitable to combine the calculation/simulation results of different groups, or predictions using other regression methods. The demonstration of PCP method is in section 6.5

Chapter 5

Regression and prediction using Gaussian process

In this chapter, the concept and characteristics of **GPR**[1], the meta-modeling method used in this research, are introduced. The first section is an overview of meta-modeling methods. Then the theory of Bayesian inference that is the base of GP is described, followed by the introduction to the principles of GP and kernel functions. The robustness of GPR when dealing with noisy data is demonstrated.

To solve the problem that GPR fails with high fluctuation in the training data, an original sampling and training method called **STMN**[2] method is considered in this research. Section 5.6 includes the introduction to the STMN method and the demonstration with Ishigami function

The last section introduces another original method, called **PCP**[3], that combines the prediction from regression models trained for different groups of parameters. PCP method is validated with mathematical deduction.

[1] **G**aussian **P**rocess **R**egression
[2] **S**ingle sample of **T**argeted parameters paired with **M**ultiple samples of **N**on-targeted parameters
[3] **P**rediction **C**ombination **P**olynomial

5.1 Derivatives of R square models

R^2 is commonly used to evaluate the errors in predicted values and validate a regression model (equation 5.1).

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i^{pd} - y_i^{ob})^2}{\sum_{i=1}^{n}(y_i^{ob} - \overline{y^{ob}})^2} \tag{5.1}$$

Where:

y^{ob} = The observations

y^{pd} = The predictions by a mathematical model

The numerator, $\sum(y_i^p - y_i^o)^2$, indicates the distance from prediction data to observation data. The denominator, $\sum(y_i^o - \overline{y^o})$, indicates the distance of observed data from its mean value. As the observations and predictions share the same sample size n, we can rewrite the formula in this form (equation 5.2),

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\frac{(y_i^{pd} - y_i^{ob})^2}{n}}{\sum_{i=1}^{n}\frac{(y_i^{ob} - \overline{y^{ob}})^2}{n}} \tag{5.2}$$

$$MSE = \sum_{i=1}^{n}\frac{(y_i^{pd} - y_i^{ob})^2}{n} \tag{5.3}$$

$$Var(y^o) = \sum_{i=1}^{n}\frac{(y_i^{ob} - \overline{y^{ob}})^2}{n} \tag{5.4}$$

Where:

MSE = **M**ean **S**quared **E**rror of the prediction data

$Var(y^o)$ = The variance of the observation data

The numerator is actually the mean squared error of the predicted data (equation 5.3), which is also often used to evaluate the errors in prediction data. However, as the value of MSE is deeply relied on the range and unit of data, there is not a objective standard for MSE. For example, we cannot say 0.4 $(km)^2$ is better than 40 $(kg)^2$. By dividing the MSE with the variance of observations (equation 5.4), this shortage has been overcame. Normally, the R^2 of predictions from a correctly regressed model should be larger than 0. A R^2 that is larger than 0.8 reflects the good performance of the regression model.

5.1.1 Variety, Fluctuation and Relative uncertainty R_U^2

By introducing the concept of targeted and non-targeted parameters, we found an interesting utilization the R^2 that is to separate uncertainties in observed data from target and non-targeted parameters (equation 5.5). If we know the 'true' value of y, then

$$R_U^2 = 1 - \frac{\sum_{i=1}^{n}(y_i^{ob} - y_i^{true})^2}{\sum_{i=1}^{n}(y_i^{true} - \overline{y^{true}})^2} \tag{5.5}$$

$$fluctuation_i = y_i^{ob} - y_i^{true} \tag{5.6}$$

$$variety_i = y_i^{true} - \overline{y^{true}} \tag{5.7}$$

$$SI^{tar} = 1 - \frac{\sum_{i=1}^{n}\frac{(y_i^{ob} - y_i^{true})^2}{n}}{Var(y^{ob})} \tag{5.8}$$

$$SI^{nt} = 1 - \frac{\sum_{i=1}^{n}\frac{(y_i^{true} - \overline{y^{true}})^2}{n}}{Var(y^{ob})} \tag{5.9}$$

Where:

y^{ob}	=	The observations (with fluctuation)
y^{true}	=	The 'true' value (without fluctuation)
SI^{tar}	=	The first-order Sobol indices of targeted parameters
SI^{nt}	=	The first-order Sobol indices of non-targeted parameters

y^{ob} means the observed value of y, which includes the uncertainties caused by the non-targeted parameters. y^t means the 'true' values of y, which is an ideal value without the uncertainties caused by non-targeted parameters. In another word, y^t is not something in the real world but just an ideal value that the influence from non-targeted parameters are excluded.

The numerator, $\sum \left(y_i^o - y_i^{true}\right)^2$, can be regarded as, with the targeted parameters decided, the offset in y (figure 5-1) caused by non-targeted parameters. Here, we introduced a new concept, ***Fluctuation*** (equation 5.6), into this research to describe this value. Actually, as shown in equation 5.8, the variance of this value value can easily converted into the first-order Sobol indices of targeted parameters (not the non-target).

Likewise, the denominator, $\sum \left(y_i^{true} - \overline{y^{true}}\right)^2$, cam be regarded as, with non-targeted parameters decided, the variance (figure 5-2) caused by targeted parameters, which can also be easily converted into the Sobol indices of non-targeted parameters (equation 5.9). We defined another concept, ***Variety*** (equation 5.7), to describe the variety of y caused by targeted parameters.

In this way, we can use R^2 to evaluate the proportions between uncertainties caused by targeted and non-

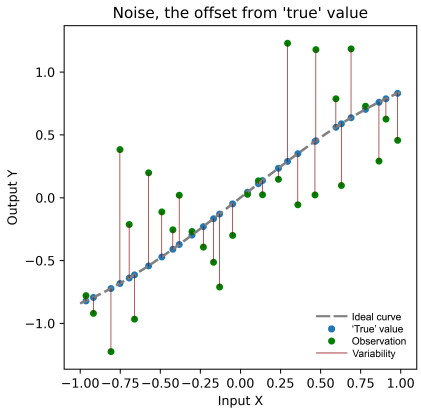


Figure 5-1: The concept of *Fluctuation*

Figure 5-2: The concept of *Variety*

targeted parameters, which is called ***relative uncertainty*** R_U^2 (equation 5.5). in this research.

If we regarded the observation as the 'true' values added with *fluctuation* (5.10), we can rewrite R_U^2 as formula 5.15. Normally, the mean value of the fluctuation is close to 0 (5.11).

$$y_i^{ob} = y_i^{ture} + fluctuation_i^{ob} \tag{5.10}$$

$$\overline{fluctuation_i^{ob}} \to 0 \tag{5.11}$$

$$R_U^2 = 1 - \frac{\sum_{i=1}^{n}(y_i^{ob} - y_i^{true})^2}{\sum_{i=1}^{n}(y_i^{true} - \overline{y^{true}})^2} \tag{5.12}$$

$$= 1 - \frac{\sum_{i=1}^{n}(fluctuation_i^{ob})^2}{\sum_{i=1}^{n}(y_i^{true} - \overline{y^{true}})^2} \tag{5.13}$$

$$= 1 - \frac{\sum_{i=1}^{n}(fluctuation_i^{ob} - \overline{fluctuation_i^{ob}})^2}{\sum_{i=1}^{n}(y_i^{true} - \overline{y^{true}})^2} \tag{5.14}$$

$$= 1 - \frac{\sigma_{nob}^2}{\sigma_t^2} \tag{5.15}$$

Where:

σ_{nob} = The variance of fluctuation in observation

σ_t^2 = The variance of 'true' values

For demonstration, we still use the Ishigami function (equation 4.63). The values of $f(x)$ when $a = 7, b =$

$0.1, c = 0$ are temperately used as the 'true' values (equation 5.16).

$$y_i^{true} = \sin(x_{1i}) + 7\sin^2(x_{2i}) + 0.1x_{3i}^4 \sin(x_{1i}) \tag{5.16}$$

Table 5-1 shows the configurations and results. Figure 5-3 illustrates the observed data.

Table 5-1: Configurations and results of the relative uncertainty test

	x_1	x_2	x_3	a	b	c	R_U^2
case 1	$[-\pi, \pi]$	$[-1, 1]$	$[-\pi, \pi]$	[6,8]	[0.05,0.15]	[-0.5,0.5]	0.96
case 2				[4,10]	[0.0,0.2]	[-1。 0,1.0]	0.78
case 3				[2,12]	[-0.1,0.3]	[-2.0,2.0]	0.35
case 4				[0,14]	[-0.2,0.4]	[-3.0,3.0]	-0.38

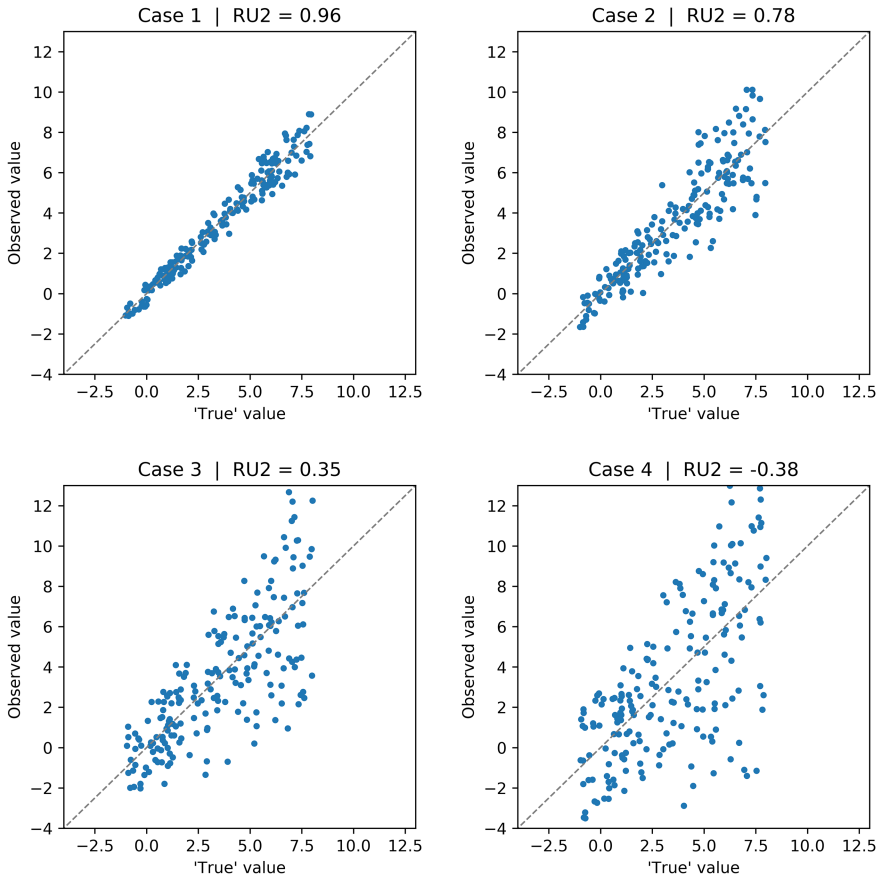


Figure 5-3: Demonstration of relative uncertainty

5.1.2 'True' R-square error R_E^2

Normally speaking, if the predictions are closer to the observation ($R^2 \rightarrow 1$), the regression is better. However, considering the fluctuation in the observation and overfitting issues, R^2 cannot correctly assess the validation of a regression, because its value is influenced by both the ***fluctuation*** in the observation and ***errors*** in the regression (equation 5.17).

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i^{pd} - y_i^{ob})^2}{\sum_{i=1}^{n}(y_i^{ob} - \overline{y^{ob}})^2} = 1 - \frac{\sum_{i=1}^{n}[(y_i^{pd} - y_i^{true}) - (y_i^{ob} - y_i^{true})]^2}{\sum_{i=1}^{n}(y_i^{ob} - \overline{y^{ob}})^2} \tag{5.17}$$

To solve this problem, we introduced 2 new concepts, ***'True' R-square error*** R_E^2 and ***error***, into this research. The word ***Error*** represents a value that meters the offset from predictions to the 'True' value (equation 5.18). ***'True' R-square error*** R_E^2 shows the proportion between ***error*** and ***variety***. Using R_E^2 can better assess a regression than the common R^2.

$$error_i = y_i^{pd} - y_i^{true} \tag{5.18}$$

$$R_E^2 = 1 - \frac{\sum_{i=1}^{n}(y_i^{pd} - y_i^{true})^2}{\sum_{i=1}^{n}(y_i^{true} - \overline{y^{true}})^2} \tag{5.19}$$

To theoretically prove this argument, let's suppose the true values of y obey a normal distribution (equation 5.20), while the offset from the true values to the observations and predictions also respectively obey normal distributions (equation 5.22 and 5.24).

$$\because y_i^{true} \sim N(\mu, \sigma_t) \tag{5.20}$$

$$y_i^{ob} = y_i^{true} + fluctuation_i \tag{5.21}$$

$$fluctuation_i \sim N(0, \sigma_{nob}) \tag{5.22}$$

$$y_i^{pd} = y_i^{true} + error_i \tag{5.23}$$

$$error_i \sim N(0, \sigma_{epd}) \tag{5.24}$$

$$\therefore y_i^{ob} \sim N\left(\mu, \sqrt{\sigma_t^2 + \sigma_{nob}^2}\right) \tag{5.25}$$

$$y_i^{pd} \sim N\left(\mu, \sqrt{\sigma_t^2 + \sigma_{epd}^2}\right) \tag{5.26}$$

$$\tag{5.27}$$

Formula 5.30 is the common R^2. From 5.25 and 5.26, we can get the distribution of $y^{pd} - y^{ob}$ (5.28). Because the mean value of $y^{pd} - y^{ob}$ is 0 (5.29), so we can rewrite the numerator into $Var(y^{pd} - y^{ob})$ (5.32). Finally, we represent the R^2 with the variances of true values of y, the fluctuation in observation and the errors in prediction (5.33). From 5.33 we can see the value of R^2 is influenced by σ^2_{nob}. And there is a $2\sigma^2_t$ in the numerator which is not desirable. So the tendency of the R^2 is not clear.

$$\because y_i^{pd} - y_i^{ob} \sim N\left(0, \sqrt{2\sigma_t^2 + \sigma^2_{epd} + \sigma^2_{nob}}\right) \tag{5.28}$$

$$\therefore \overline{y^{pd} - y^{ob}} = 0 \tag{5.29}$$

$$\therefore R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i^{pd} - y_i^{ob})^2}{\sum_{i=1}^{n}(y_i^{ob} - \overline{y^{ob}})^2} \tag{5.30}$$

$$= 1 - \frac{\sum[(y_i^{pd} - y_i^{ob}) - \overline{y^{pd} - y_i^{ob}}]^2}{\sum_{i=1}^{n}(y_i^{ob} - \overline{y^{ob}})^2} \tag{5.31}$$

$$= 1 - \frac{Var(y^{pd} - y^{ob})}{Var(y^{ob})} \tag{5.32}$$

$$= 1 - \frac{2\sigma_t^2 + \sigma^2_{epd} + \sigma^2_{nob}}{\sigma_t^2 + \sigma^2_{nob}} \tag{5.33}$$

On the other hand, R^2_E only includes the variance of the 'true' values of y and the prediction errors (5.33), which is much more ideal.

$$\because y_i^{pd} - y_i^{true} = error_i \sim N(0, \sigma_{epd}) \tag{5.34}$$

$$\therefore \overline{y^{pd} - y_i^{true}} = 0 \tag{5.35}$$

$$\therefore R^2_E = 1 - \frac{\sum_{i=1}^{n}(y_i^{pd} - y_i^{true})^2}{\sum_{i=1}^{n}(y_i^{true} - \overline{y^{true}})^2} \tag{5.36}$$

$$= 1 - \frac{\sum[(y_i^{pd} - y_i^{true}) - \overline{y^{pd} - y^{true}}]^2}{\sum_{i=1}^{n}(y_i^{true} - \overline{y^{true}})^2} \tag{5.37}$$

$$= 1 - \frac{Var(error)}{Var(y^{true})} \tag{5.38}$$

$$= 1 - \frac{\sigma^2_{epd}}{\sigma_t^2} \tag{5.39}$$

To illustrate this problem more vividly, we simply use a $y = \sin(x)$ curve as the 'true' values. The observations are simply the 'true' values added with ***fluctuation*** that obey a normal distribution. Similarly, the observations are simply the 'true' values added with ***error*** that proportional to ***fluctuation***.

We generated 2 collections of data, each collection includes 50 samples. For the 1^{st} collection, we intentionally set relatively higher fluctuation and lower error, to represent a good regression that find the 'true' values even with very noisy observations. For the 2^{nd} collection, we set relatively lower fluctuation and higher error, to represent a overfitted regression that approaches the observations. (table 5-2)

Table 5-2: Configurations and results of the illustration

	x	y^{true}	y^{ob}	y^{pd}	R^2	R_E^2
Collection 1	$U(-1,1)$	$y = \sin(x)$	$y^{true} + fluctuation \sim N(0, 0.5)$	$y^{true} + fluctuation * 0.5$	0.77	0,97
Collection 2			$y^{true} + fluctuation \sim N(0, 0.3)$	$y^{true} + fluctuation * 0.8$	0.99	0.80

From figure 5-4, we can see the common R^2 says that the overfitted predictions in collection 2 is better, which goes against our purpose. The judgments from R_E^2 is more suitable in this research.

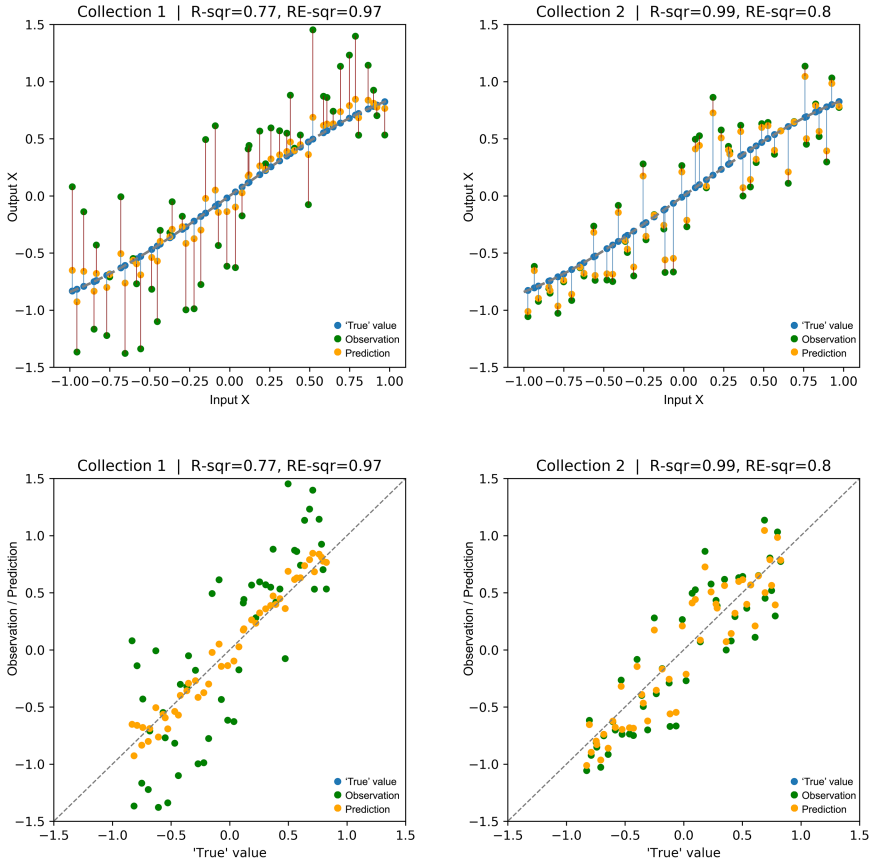


Figure 5-4: Validation of R_E^2

5.1.3 Section conclusion

In this section, in order to make things more clear and easy to understand. We defined some new concepts and introduced 2 derivatives of the common R^2 model. These new definitions are concluded in table 5-3 and 5-4. Though these concepts do not directly relate to the proposed framework for the early design stage in this research, they can help us better analyze and understand the data.

Table 5-3: Concepts defined in this research

Concept	Statistical meaning
'True' value	An ideal value that the influence from non-targeted parameters is excluded. The base of all others.
Variety	The variance in y caused by targeted parameters
Fluctuation	The variance in y caused by non-targeted parameters, the offset from observations to 'true' value
Error	The offset from predictions to 'true' value

Table 5-4: R^2 derivatives used in this research

Model	Symbol	Raw formular	Converted formular
R-square error	R^2	$1-\sum(y^{pd}-y^{ob})^2/\sum(y^{ob}-\overline{y^{ob}})^2$	
Relative uncertainty	R_U^2	$1-\sum(y^{ob}-y^{true})^2/\sum(y_i^{true}-\overline{y^{true}})^2$	$1-\sigma_{nob}^2/\sigma_t^2$
'True' R-square error	R_E^2	$1-\sum(y^{pd}-y^{true})^2/\sum(y_i^{true}-\overline{y^{true}})^2$	$1-\sigma_{epd}^2/\sigma_t^2$
Model	**Statistical meaning**		
R-square error	Metering the offset from predictions to observations		
Relative uncertainty	Metering the proportion between *fluctuation* and *variety*		
'True' R-square error	Metering the proportion between *error* and *variety*		

The concept of ***'true' value*** is the base of all other concepts, as well as R_U^2 and R_E^2. However, in many cases, maybe most cases, the ***'true' values*** of y are not available. In this research, we introduced a new sampling and training method called **STMN** to solve this problem, which is described with details and illustrations in section 5.6.2. How to make use of the R_U^2 and R_E^2 is also introduced in that section.

5.2 An overview of meta-modeling methods

In this research, meta-modeling is used to solve the problem that it costs too much time to explore the design possibilities widely using simulations in the early design stage. Gaussian process is used to regress the simulation data and make predictions. Before talking about Gaussian process, we would like to give an overview of meta-modeling methods, and explain the reason that Gaussian process, rather than others, is used in this research.

Meta-modeling is a mathematical method to analyze the data from survey or experiments, including computational experiments. With meta-models made, one can quickly get predicted new results without doing additional survey or experiments. Meta-modeling is not something fashionable but actually classical, since Newton. For example, equation $h = \frac{1}{2}gt^2$ is the meta-model of free fall.

Besides some classical methods, such as linear regression[82], the most meta-modeling methods nowadays are regraded as machine-learning, while some machine-learning methods are declared to be the so called Artificial Intelligence. However, the words of 'machine-learning' and 'AI' is now abused in commercial context. As researchers, in this research, we just use the word 'meta-modeling'

Based on the purpose, the meta-modeling can be separated into regression and classification. The classification meta-modeling can further separated into supervised and unsupervised. k-means clustering[81] and SVM[85] are powerful and easy to use unsupervised methods. As the main purpose to use meta-modeling is regression, the classification methods will not be introduced in this section. Figure 5-5 illustrate the inclusion relation in the field of meta-modeling.

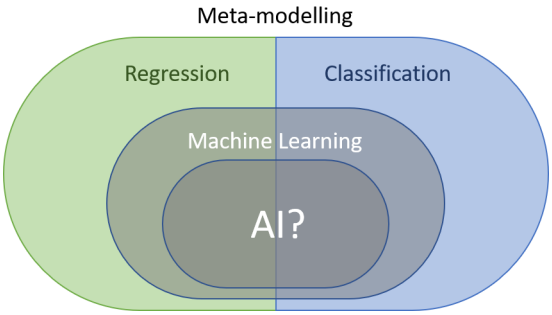


Figure 5-5: Overview of meta-modeling

In this research, we have tried **MLR**, the most popular **ANN**, an unpopular method called **PCE**, as well as **GPR** that is finally adopted in this research.

Multiple **L**inear **R**egression is a modern update of the classical linear regression, which is now also regarded as a kind of machine-learning method. **(WHAT?)** The expression of MLR is very simple (5.41). However, in this research, the relationship between parameters and criteria is not always linear. Besides, the interactions between parameters cannot be reflected in a MLR model. Therefore, MLR model is not adopted.

$$x = p_1, p_2, \cdots, p_d \tag{5.40}$$

$$y_i = b_0 + \sum_{j=1}^{d} b_j p_{ji} + e_i \tag{5.41}$$

Where:

b_0 : The regression intercept

b_j : The j-th parameter regression slope

$e_i \sim N(0, \sigma^2)$, a Gaussian error term

Artificial **N**eural **N**etwork, the most popular machine learning method since the victories of Alpha-Go, is inspired by the neural network in brains. ANN is very powerful in classification and vague estimations issues. For example, CNN[cite] is widely used in pattern recognizing and RNN[cite] is widely used in language processing. Many ANN models ends with softmax function[cite], while the error is estimated using cross entropy[cite], which is specialized for classification issues. However, we have tried ANN to regress the simulation data. ANN models failed to give decent predictions.The reason might be our misuse but, anyway, we abandoned ANN.

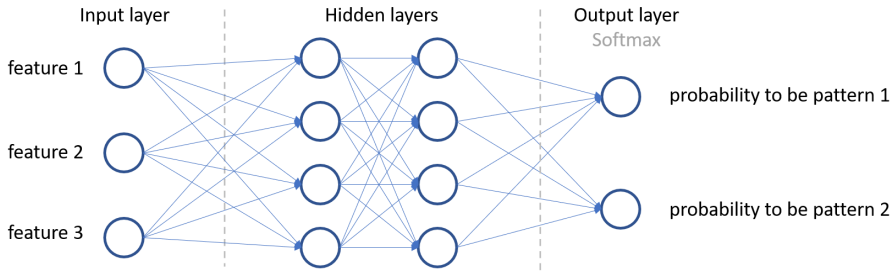


Figure 5-6: A typical construction of an ANN model

Polynomial **C**haos **E**xpansion (**PCE**) can be regarded as an extension of Volterra's theory of nonlinear functionals for dynamic systems. Equation 5.42 is the expression of PCE of 2^{nd} order stochastic processes [cite wiener 1938].

$$y(x, t, \theta) = \sum_{k=0}^{\infty} \beta_k(x, t) \Psi_k(\xi(\theta)) \tag{5.42}$$

PCE is usually used to analysis the sensitivity and uncertainty. Data-driven PCE can also be utilized in regression. We have tried to regress the simulation data using PCE with a Python library called OpenTurns[cite], but failed to get good results. Therefore, we also gave up PCE.

5.3 Introduction to Gaussian Process

5.3.1 Gaussian process

The very basic presumption of Gaussian process is that the distribution of observations is a normal/Gaussian distribution. The reasons that normal distribution is used are that, firstly many phenomenons in the real world is normal distributed, secondly the mathematical transforming result of a normal distribution is still normal distributed.

For any function $f(x)$, x is a d dimensional real number input (5.43), we assume that the $f(x)$ is normal distributed with mean μ and variance σ^2 (5.44).

$$x \in \mathbb{R}^d \tag{5.43}$$

$$f(x) \sim N(\mu, \sigma) \tag{5.44}$$

Now, we have observed a set of points X that includes n input vectors, $\{x_1, x_2, x_3, \cdots, x_n\}^T$. Then the joint distribution of $\{f(x_1), f(x_2), f(x_3), \cdots, f(x_n)\}^T$ is still normal distributed (5.47).

$$X = \begin{Bmatrix} x_1 \\ \vdots \\ x_n \end{Bmatrix} = \{x_1, x_2, x_3, \cdots, x_n\}^T \tag{5.45}$$

$$Y = \begin{Bmatrix} f(x_1) \\ \vdots \\ f(x_n) \end{Bmatrix} = \{f(x_1), f(x_2), f(x_3), \cdots, f(x_n)\}^T \tag{5.46}$$

$$P(Y|X) = N(\mu_Y, \Sigma_Y) \tag{5.47}$$

$$\mu_Y = \{\mu_1, \mu_2, \cdots, \mu_n\}^T \tag{5.47a}$$

$$\Sigma_Y = \begin{bmatrix} cov_{1,1} & \cdots & cov_{1,n} \\ \vdots & \ddots & \vdots \\ cov_{n,1} & \cdots & cov_{n,n} \end{bmatrix} \tag{5.47b}$$

$$cov_{i,i} = 1 \tag{5.47c}$$

Where:

$P(Y\|X)$	:	The conditional joint distribution of Y
μ_i	:	The mean value of the distribution of $f(x_i)$
$cov_{1,1}$	:	The covariance of $f(x_i)$ and $f(x_j)$
Σ_Y	:	The covariance matrix of Y and Y

The covariance matrix of this distribution(5.47b) is an positive-definite matrix, whose diagonal consists of values of 1 (5.47c).

Then, if we have another set of testing points, $X*$, that includes m input vectors, the conditional joint distribution of the corresponding $Y*$ is certainly normal distribution (5.50).

$$X* = \{x*_1, x*_2, x*_3, \cdots, x*_m\}^T \tag{5.48}$$

$$Y* = \{y*_1, y*_2, y*_3, \cdots, y*_m\}^T \tag{5.49}$$

$$P(Y*|X*) = N(\mu_{Y*}, \Sigma_{Y*}) \tag{5.50}$$

$$\mu_{Y*} = \{\mu*_1, \mu*_2, \cdots, \mu*_n\}^T \tag{5.50a}$$

$$\Sigma_{Y*} = \begin{bmatrix} Cov_{1,1} & \cdots & Cov_{1,m} \\ \vdots & \ddots & \vdots \\ Cov_{m,1} & \cdots & Cov_{m,m} \end{bmatrix} \tag{5.50b}$$

Based on these assumptions, we can get the joint distribution of Y and $Y*$ can be inferred (5.51). The covariance matrices, Σ_{YY*} and Σ_{Y*Y} are positive semi-definite.

$$P(Y, Y*|X, X*) = N\left(\begin{bmatrix} \mu_Y \\ \mu_{Y*} \end{bmatrix}, \begin{bmatrix} \Sigma_Y & \Sigma_{YY*} \\ \Sigma_{Y*Y} & \Sigma_{Y*} \end{bmatrix}\right) \tag{5.51}$$

$$\Sigma_{YY*} = \begin{bmatrix} Cov_{1,1} & \cdots & Cov_{1,m} \\ \vdots & \ddots & \vdots \\ Cov_{n,1} & \cdots & Cov_{n,m} \end{bmatrix} \tag{5.51a}$$

$$\Sigma_{Y*Y} = \Sigma_{YY*}^T \tag{5.51b}$$

Where:

$P(Y, Y*\|X, X*)$	:	The conditional joint distribution of Y and $Y*$
Σ_{YY*}	:	The covariance matrix of Y and $Y*$
Σ_{Y*Y}	:	The covariance matrix of $Y*$ and Y
Σ_{Y*}	:	The covariance matrix of $Y*$ and $Y*$

Base on the principles of multivariate normal distribution[83], the conditional distribution of $Y*$ can be expressed as formula 5.52.

$$P(Y*|X*, X, Y) = N(\mu*, \sigma*) \tag{5.52}$$

$$\mu* = \mu_{Y*} + \Sigma_{Y*Y}\Sigma_Y^{-1}(Y - \mu_Y) \tag{5.52a}$$

$$\sigma* = \Sigma_{Y*} - \Sigma_{Y*Y}\Sigma_Y^{-1}\Sigma_{YY*} \tag{5.52b}$$

However, as the distribution of $f(x)$ is not actually known, correspondingly, the covariance matrices are also unknown. In GP, a method called *kernel function* is used to generate 'fake' covariance matrices.

$$\begin{bmatrix} Y \\ Y* \end{bmatrix} = N\left(\begin{bmatrix} \mu(X) \\ \mu(X*) \end{bmatrix}, \begin{bmatrix} K & K_* \\ K_*^T & K_{**} \end{bmatrix} \right) \tag{5.53}$$

$$K = \begin{bmatrix} k(x_1, x_1) & \cdots & k(x_1, x_n) \\ \vdots & \ddots & \vdots \\ k(x_n, x_1) & \cdots & k(x_n, x_n) \end{bmatrix} \tag{5.53a}$$

$$K_* = \begin{bmatrix} k(x_1, x*_1) & \cdots & k(x_1, x*_m) \\ \vdots & \ddots & \vdots \\ k(x_n, x*_1) & \cdots & k(x_n, x*_m) \end{bmatrix} \tag{5.53b}$$

$$K_{**} = \begin{bmatrix} k(x*_1, x*_1) & \cdots & k(x*_1, x*_m) \\ \vdots & \ddots & \vdots \\ k(x*_m, x*_1) & \cdots & k(x*_m, x*_m) \end{bmatrix} \tag{5.53c}$$

Where:

X	:	The observation points
Y	:	The observed values of $f(x)$
$X*$	:	The points to test
$Y*$	:	The values that we want to predict using regression model
$\mu(x)$	:	The mean value function
$k(x_i, x_j)$	:	The kernel function that generate 'fake' covariance of x_i and x_j

So, the expression of the conditional distribution of $Y*$ is the formulas 5.54, 5.54a and 5.54b.

$$P(Y* | X*, X, Y) = N(\mu*, \sigma*) \tag{5.54}$$

$$\mu* = \mu(X*) + K_*^T K^{-1}(Y - \mu(X)) \tag{5.54a}$$

$$\sigma* = K_{**} - K_*^T K^{-1} K_* \tag{5.54b}$$

5.3.2 Kernel function

The problem now is that we have no information about the covariance matrices to inference the probability distribution of $Y*$. In GP, ***kernel functions*** are used to make a 'fake' covariance matrices. For that reason, the *kernel functions* are often called ***kernel tricks***, because they do not stand for anything in the real world but are a kind of mathematical skill.

The inputs of a kernel function are 2 input vectors, x_a and x_b, of the function $f(x)$ that we want to regress. Generally speaking, a kernel function is used to evaluate the distance between x_a and x_b then decide the corresponding covariance.

There are various kinds of kernel functions. For **most (not all)** of them, if x_a and x_b are close to each other (5.55), the 'fake' covariance of $f(x_a)$ and $f(x_b)$, $k(x_a, x_b)$, is close to it maximum value (5.56), which is an important principle for most kinds of kernel functions. So that the values of $f(x_a)$ and $f(x_b)$ are highly probably close (5.57).

$$\because x_a \to x_b \tag{5.55}$$
$$k(x_a, x_b) \to k_{max} \tag{5.56}$$
$$\therefore f(x_a) \to f(x_b) \tag{5.57}$$

In another word, the probability that $f(x_a)$ equal to itself is always greater than that it equal to others (5.58). Therefore, the 'fake' covariance between $f(x_a)$ and $f(x_b)$ should be less than the 'fake' variance of $f(x_a)$ (5.59).

$$\because P(f(x_a) = f(x_a)) \geq P(f(x_a) = f(x_b)) \tag{5.58}$$
$$\therefore k(x_a, x_b) \leq k(x_a, x_a) \tag{5.59}$$

Normally, the input vector x is normalized before regression, which means the range of x is within $[0, 1]$ (5.60). Therefore, the distance between x_a and x_b should less than or equal to d (5.61), as x is a d-dimensional vector. $x_a = 0$ and $x_b = 1$ should be the most the faraway pair in X. So that the value of $k(1, 0)$ or $k(0, 1)$ should be the smallest (5.61), which is close to 0 in many cases. It is another important principle for most kinds of kernels.

$$\because x \in [0, 1]^d \tag{5.60}$$
$$\therefore ||x_a - x_b||_2^2 \leq d^2 \tag{5.61}$$
$$\therefore k(x_a, x_b) \geq k(0, 1) \tag{5.62}$$
$$k(x_a, x_b) \geq k(1, 0) \tag{5.63}$$

So the main idea of most kinds of GP can be explained as, **if x_a and x_b are close to each other, then $f(x_a)$ and $f(x_b)$ are highly probably close to each other**, which is quite intuit.

Radial **B**asis **F**unction (**RBF**) kernel is the very most widely used kernel in GPR. Equation 5.64 in the expression of the RBF. Notice that x is a d-dimensional vector.

$$k_{rbf}(x_a, x_b) = \exp\left(-\frac{||x_a - x_b||_2^2}{2\sigma^2}\right) \tag{5.64}$$

$$x = p_1, p_2, \cdots, p_d \tag{5.65}$$

$$||x_a - x_b||_2^2 = \sum_{i=1}^{d}(p_{ai} - p_{bi})^2 \tag{5.66}$$

Where:

$k_{rbf}(x_a, x_b)$	:	The RBF kernel on x_a and x_b
$\|\|x_a - x_b\|\|_2^2$	:	The squared Euclidean distance between x_a and x_b
σ	:	A free parameter

For the reason that the distance between x_a and x_b is always greater than or equal to 0, the maximum value of RBF is 1, when $x_a = x_b$. The minimum value of RBF is 0 when the distance is ∞. The value of the RBF decreases with the distance increasing, and ranges within $(0, 1]$, the RBF can be regarded as a kind of similarity measurement.

$$\because 0 \leq ||x_a - x_b||_2^2 < \infty \tag{5.67}$$

$$\therefore k_{rbf} \leq \exp\left(-\frac{0}{2\sigma^2}\right) = 1 \tag{5.68}$$

$$k_{rbf} > \exp\left(-\frac{\infty}{2\sigma^2}\right) = 0 \tag{5.69}$$

$$k_{rbf} \in (0, 1] \tag{5.70}$$

Figure 5-7, 5-8 and 5-9 are illustrations of a RBF with 1-dimensional input space. The value of $2\sigma^2$ is 1. As x is 1-dimensional, the squared Euclidean distance $||x_a - x_b||_2^2$ is simply equal to $(x_a - x_b)^2$.

$$k_{rbf}(x_a, x_b) = \exp\left(-(x_a - x_b)^2\right) \tag{5.71}$$

$$x \in [0, 2] \tag{5.72}$$

Figure 5-7 shows the relationship curve between $x_a - x_b$ and $k_{rbf}(x_a, x_b)$. The value of the RBF increases with the value of $x_a - x_b$ and reaches the maximum value 1 when $x_a = x_b$, then RBF decreases.

Figure 5-8 illustrates the covariance matrix of $f(x_a)$ and $f(x_b)$. The thickness of blue color denotes the value of RBF, the thicker the larger. This symmetric diagram also hints the positive definiteness of the covariance matrix.

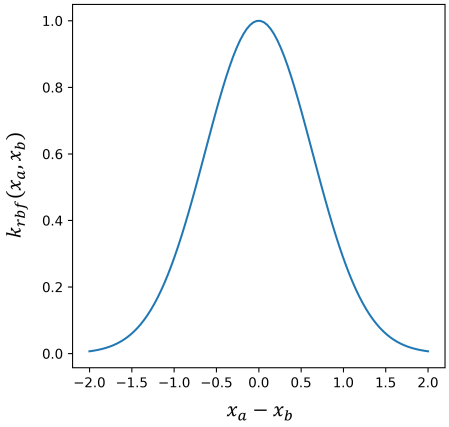


Figure 5-7: Observed points

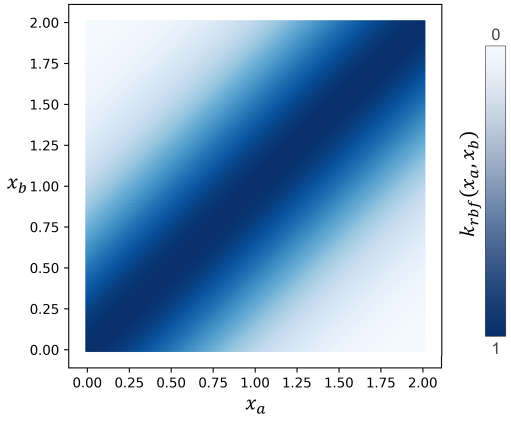


Figure 5-8: Regressed curve

If we display this matrix in 3-D space, the illustration will be more vivid and intuitive. The length and depth stand respectively for the value of $f(x_a)$ and $f(x_b)$. The height stands for the value of RBF.

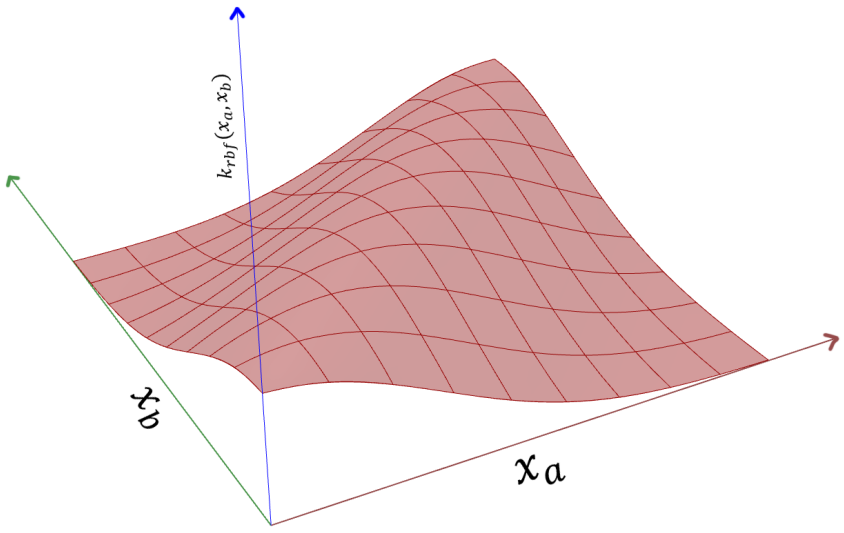


Figure 5-9: RFB illustration in 3-D

Besides RBF, we also used Matern 3/2 and 5/2 kernels in this research, which is denoted by **Mat32** and **Mat52** in most documents.

Both Mat32 and Mat52 kernels are based on the Matérn covariance function[cite] (equation 5.73), named after the Swedish forestry statistician Bertil Matérn.

$$C_\upsilon(d) = \sigma^2 \frac{2^{1-\upsilon}}{\Gamma(\upsilon)} \left(\sqrt{2\upsilon}\frac{d}{\rho}\right)^{\upsilon} K_\upsilon \left(\sqrt{2\upsilon}\frac{d}{\rho}\right) \tag{5.73}$$

$$d = \sqrt{||x_a - x_b||_2^2} \tag{5.73a}$$

$$= \sqrt{\sum_{i=1}^{d}(p_{ai} - p_{bi})^2}$$

Where:

d	:	The Euclidean distance between x_a and x_b
$\Gamma(\upsilon)$	:	The Gamma function
K_υ	:	The modified Bessel function of the second kind
υ, ρ	:	Positive parameters

In the case that $\upsilon = p + 1/2$ and p is an integer, the Matérn covariance function can be simplified into a product of an exponential and a polynomial of order p. When $\upsilon = 3/2$, it is the Mat32 kernel (5.74).

$$C_{3/2}(d) = \sigma^2 \left(1 + \frac{\sqrt{3}d}{\rho}\right) \exp\left(-\frac{\sqrt{3}d}{\rho}\right) \tag{5.74}$$

Likewise, Mat52 kernel (5.75) is the simplification of Matérn covariance function when $\upsilon = 5/2$.

$$C_{5/2}(d) = \sigma^2 \left(1 + \frac{\sqrt{5}d}{\rho} + \frac{\sqrt{5}d^2}{3\rho^2}\right) \exp\left(-\frac{\sqrt{5}d}{\rho}\right) \tag{5.75}$$

The expression of Mat32/Mat52 (5.74,5.75) kernel is similar to that of RBF (5.64). Actually, they also share the similar characteristics and curves of k and $x_a - x_b$. We did not find essential difference between RBF and Mat32/Mat52, just a little difference in the R_E^2 in the trained models. In this research, we used all RBF, Mat32 and Mat52 kernels when training model and chose the one with least R_E^2. In most cases, the differences were negligible.

There are also other kinds of kernels, such as Linear, Bias, Brownian. These kernels are not suitable in regression works. Therefore, they were not considered in this research.

5.4 Time cost to establish and train a GPR model

Time cost and memory usage to train a GPR model is very sensitive to the sample size of the training data. For each pair of x_a and x_b, the time cost to calculate the Euclidean distance is linear to the dimension of input space d, which is much less than that to execute the kernel function (5.78, 5.79). So, the time cost for each pair, or each element in the covariance matrix in another word, is nearly decided by kernel function. To establish a GPR models, the kernel functions is executed s^2 times, equal to the size of the matrix. Therefore, the time cost to establish a GPR model rises exponential with the sample size. The time cost to train a GPR model is strong influenced by the converging configurations, but the tendency is similar.

$$T_m = (c_{md} \times d + c_{mk}) \times s^2 + c_{m1} \tag{5.76}$$

$$T_t = r \times [(c_{td} \times d + c_{tk}) \times s^2 + c_{t1}] + c_{t2} \tag{5.77}$$

$$c_{md} \times d \ll c_{mk} \tag{5.78}$$

$$c_{td} \times d \ll c_{tk} \tag{5.79}$$

Where:

Tm	:	Time cost to establish an untrained GPR model
Tt	:	Time cost to train a GPR model
d	:	The dimension of the input space
s	:	The sample size of the training data
r	:	The number of iterations
c	:	Coefficients

In this research, the Python library we used to train GPR models is GPy. We tested the time cost with a very simple function (5.80).

$$y = \sum_{i=1}^{n} (d + 1 - i) p_i^i \tag{5.80}$$

Table 5-5 shows the test results. We can clear see that the time cost to establish a model T_M is linear to s^2, while influence from d is almost negligible. The relationship between T_t and s is not clear, but T_t surely increases with s.

Table 5-5: Time cost to establish and train a GPR model

d	1		2		3		4		5		6	
s	Tm	Tt	Tm	Tt	Tm	Tt	Tm	Tt	Tm	Tt	Tm	Tt
200	0.016	0.438	0.018	1.006	0.015	0.844	0.011	0.730	0.010	0.959	0.016	0.394
400	0.059	4.447	0.054	6.215	0.071	6.356	0.054	6.819	0.053	7.141	0.065	5.448
600	0.129	9.659	0.126	13.736	0.121	21.115	0.147	9.995	0.117	10.415	0.117	14.977
800	0.239	17.349	0.222	22.842	0.201	19.602	0.189	15.494	0.182	14.493	0.205	17.135
1000	0.312	26.743	0.331	31.111	0.291	46.919	0.322	30.945	0.342	18.466	0.320	38.091

5.5 Gaussian process and uncertainty

The reason that we chose Gaussian process regression is the robustness of GPR when dealing with noisy data.

In order to illustrate with 2-D diagrams, we simply used a case with only one TP and one NTP. Figure 5-10 shows the configurations of the case and simulation.

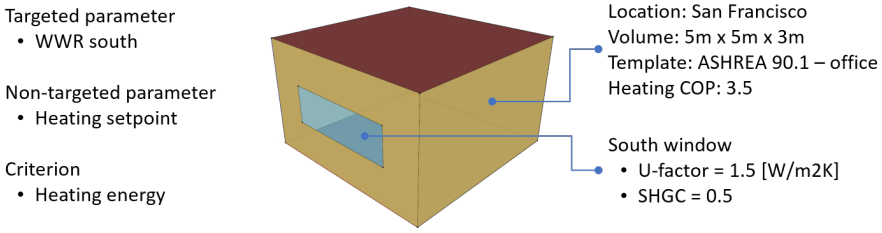


Figure 5-10: Predicted probability density

We set the range of WWR $[0.1, 0.9]$, then introduce a little uncertainty into Htsp, setting the range of Htsp $[17.95, 18.05]$. 100 samples were made using LHS and heating energy (HE) was calculated using EnergyPlus. At this stage, we regarded the $(HE|Htsp = 18°C$ as 'true' values HE^{true}. As illustrated in figure 5-11, the values of HE^{ob}s were noisy because of the existence of the uncertainty in Htsp. If we use a curve, such as sine or square, to regress these observed points in a traditional mathematical way, or use ANN[4] or PCR[5], we will get a curve (figure 5-12). The predicted HE^{pd} for a single WWR^* is a single value, which can hardly reproduced the Fluctuation, caused by the non-targeted parameters, in the observation data.

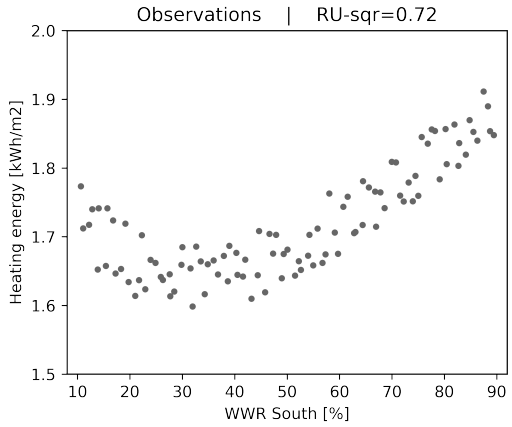


Figure 5-11: Observed points

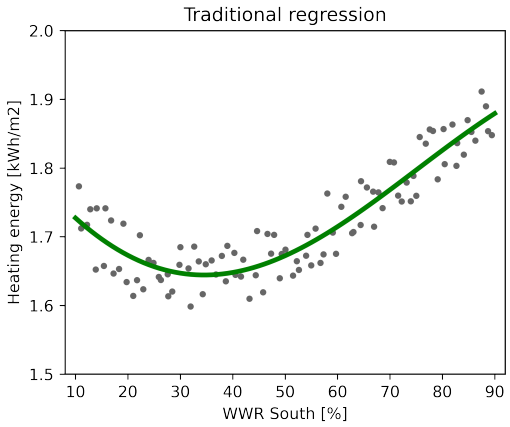


Figure 5-12: Regressed curve

[4] Artificial Nerual Network
[5] Polynomial Chaos Regression

On the other hand, for a single value of WWR^*, the prediction from a trained GPR model is, rather than a single value, the probability distribution of HE^{pd}, $P(HE^{pd}|WWR^*, (WWR, HE^{ob}))$, with respect to the prior $P(WWR, HE^{ob}))$. The mean value of this distribution is vary close to the 'true' value of HE. The R_E^2 of the mean values of predictions is 0.99. This distribution, or the variance of this distribution, can reflects the uncertainties in the observation data. To display this distribution in a 2-D diagram, confidence interval can be used, like figure 5-13. Another way is to use the color scale to represent the probability (figure 5-14).

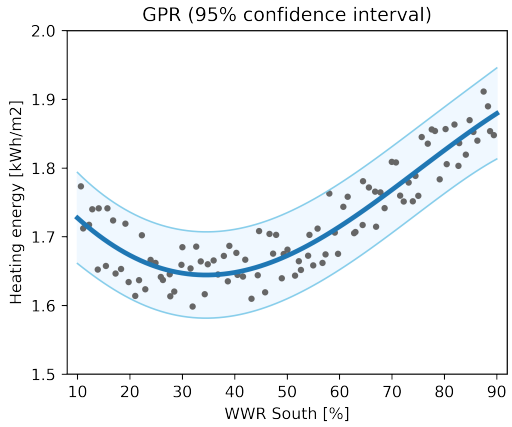


Figure 5-13: Predicted 95% confidence interval

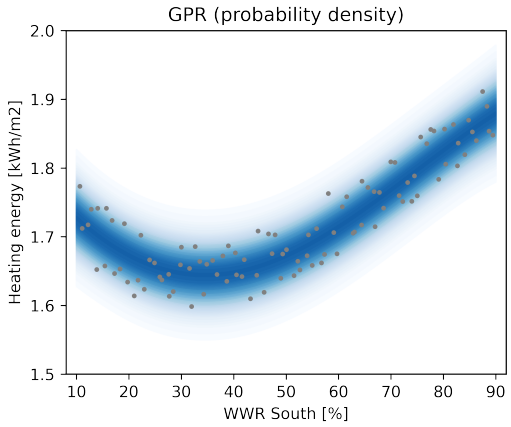


Figure 5-14: Predicted probability density

Actually, rather than a curve, a trained GPR model can be displayed in a 3D diagram, as illustrated in figure 5-15.

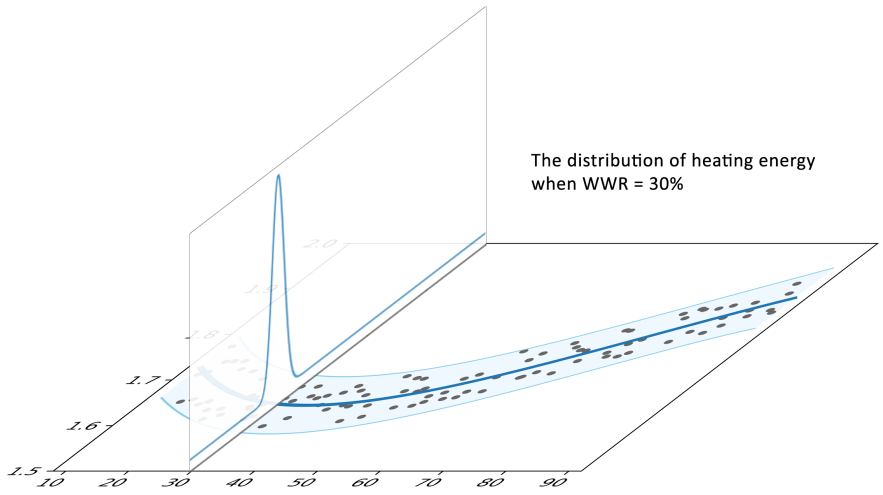


Figure 5-15: Predicted distribution of Heating energy displayed in 3D

5.6 STMN sampling and training method

5.6.1 Failure of GPR coming with a high relative uncertainty

Though GPR is robust at regressing fluctuation data, in this research, we still found that, with fluctuation in the observations getting high and sample size getting small, a trained GPR model will fail to give out valid predictions. However, as one of characteristics of building energy performance calculation, the sensitivities of NTPs are usually much larger than the sensitivities of TPs.

To illustrate this problem, we enlarged the range on Htsp to $[17.5, 18.5]$. We sampled 200 pairs of WWR and Htsp using LHS, calculated the corresponding HE^{ob} and HE^{true}. The *relative uncertainty* R_U^2 was analyzed. Then a GPR model was trained using WWR, HE^{ob}. To assess the GPR model, 200 more WWR^* were sampled and the corresponding HE^{true*} were calculated, as well as the HE^{pd} predicted. The *'true' R-square error* R_E^2 of the predictions was calculated.

The first diagram in figure 5-16 shows the observations and 'true' curve and confidence interval. The R_U^2 was as big as -5.70, which hints that the Fluctuation caused by NTP was much bigger than the Variety caused by TP. As illustrated in the second and third diagrams in figure 5-16, the training was strongly influenced by the Fluctuation. The R_E^2 was not good. One solution is to enlarge the sample size. As illustrated in figure

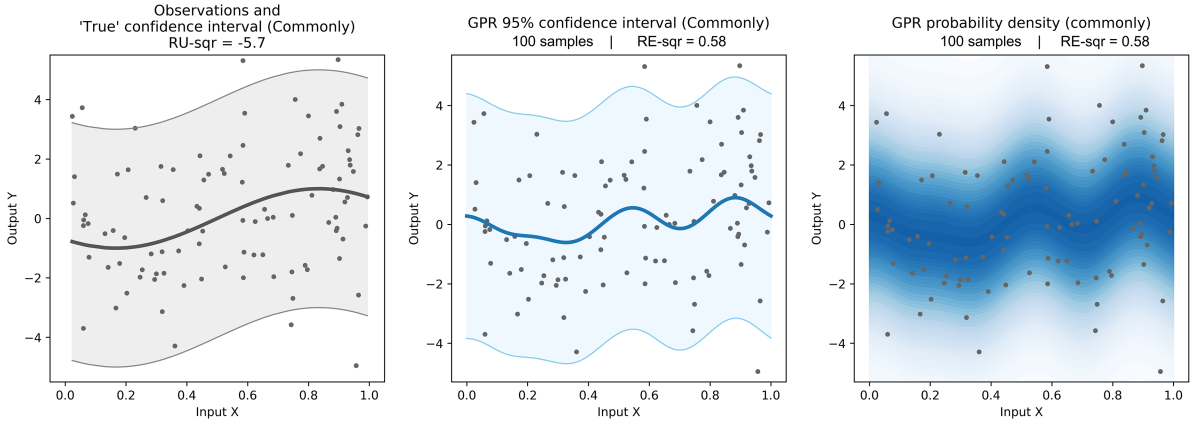


Figure 5-16: Failure of GPR

5-17, with the sample size increasing, the predictions gets closer to the *'true' value*. However, enhancing the sample size would lead to 2 unexpected results.

- Linear growth in time cost to do the calculation/simulation/survey to collect training data
- **Exponential growth** in time cost and memory usage to train a GPR model (section 5.4)

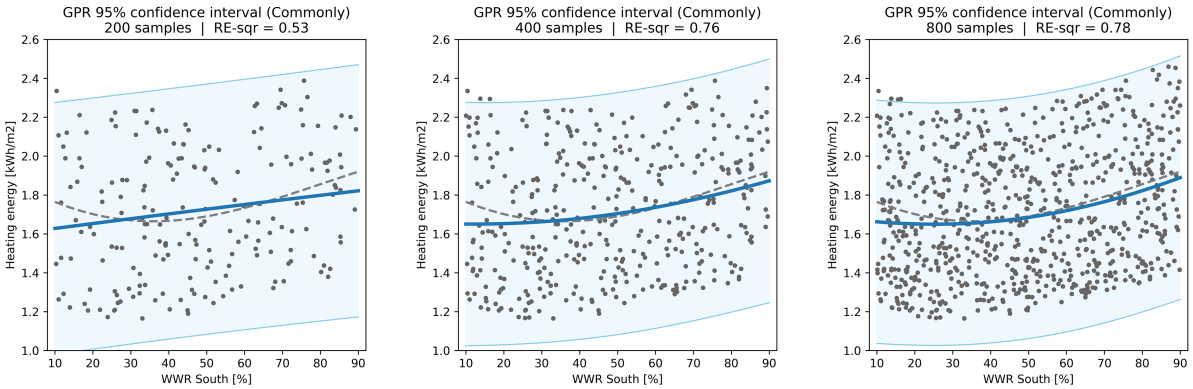


Figure 5-17: R_E^2 enhances along with the sample size

5.6.2 Single sample of targeted parameters paired with multiple samples of non-targeted parameters

In order to perform successful training with a relative small sample size when the *relative uncertainty* is large, we considered a kind of stratified sampling method named "**S**ingle '**T**rue' value paired with **M**ultiple **N**oise (**STMN***)" to collect training data.

Again, we used the single-input sine function. Firstly, we randomly generated 20 WWR and calculated the corresponding HE^{true}. Then, for each WWR, we randomly generate 10 $Htsp$. So, 200 cases were generated in total. Then a GPR was trained.

In the first diagram of figure 5-18, we can see that a single WWR corresponds with 10 HE^{ob}. The R_E^2 is 0.90, proving that we performed a relatively successful training with 20 sample sets, 200 cases in total.

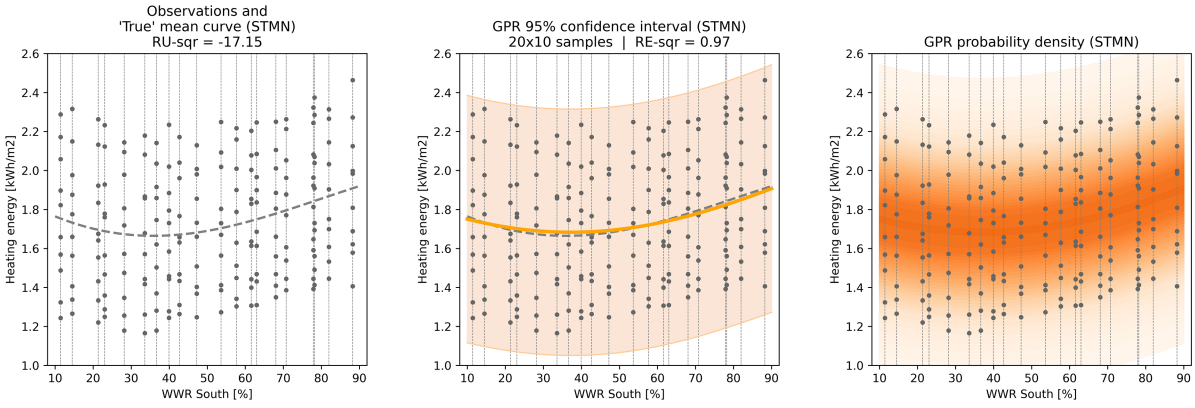


Figure 5-18: Failure of GPR

In most cases, there are several obstacles to adopt STMN* method,

1. It is very difficult to find a proper definition for y^{true}.

2. Due to the interactions between targeted and non-targeted parameters, the relationship between y^{true} and Fluctuation is not simply addition.

3. It is almost impossible to directly put our hands on the value of fluctuation.

To overcome these obstacles, we considered a '**S**ingle sample of **T**argeted parameters paired with **M**ultiple samples of **N**on-targeted parameters (**STMN**)' method. Figure 5-19 illustrate the flow of STMN method. A single sample of targeted parameters is paired with multiple samples of non-targeted parameters randomly generated and extended to a sample set including multiple cases. After the execution of calculations/simulations/survey, we get a set of observations y^{ob}. GPR models are trained with targeted parameters and y^{ob} values. The mean value of observations in each set, $\overline{y_i^{ob}}$, is used as y_i^{true} in R_U^2 and R_E^2 analysis.

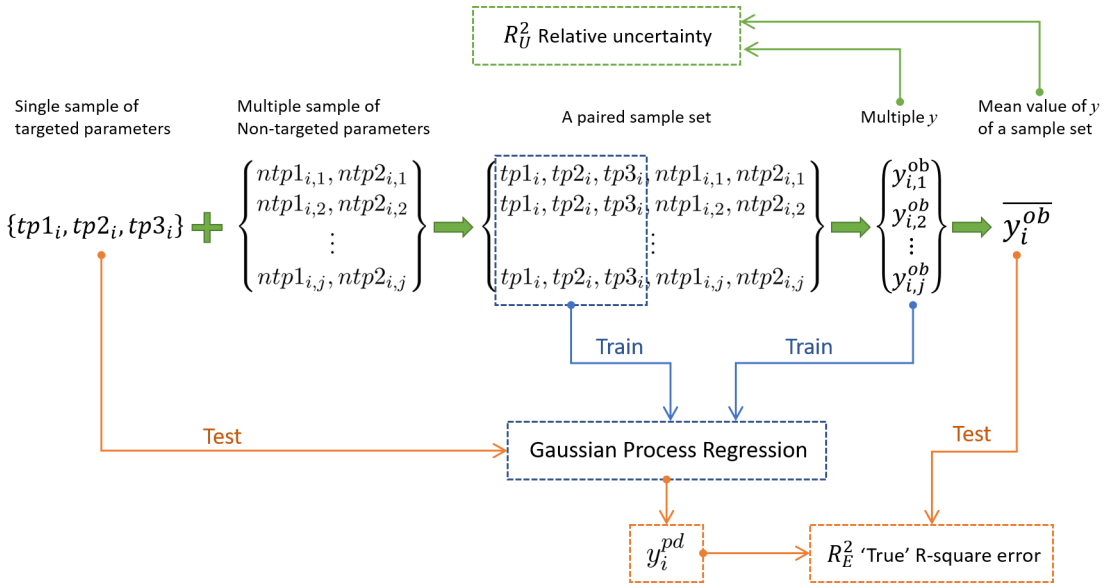


Figure 5-19: Predicted probability density curved surface

To define the 'true' value in this STMN method, we have considered 2 ways at the very beginning. (1) One is to use the mean values of the distributions of non-targeted parameters, which are constant values, in the calculation. Then use the results as y^true (equation 5.81).

$$y_i^{true} = f(tp1_i, tp2_i, \cdots, \overline{ntp1}, \overline{ntp2}, \cdots) \tag{5.81}$$

(2) Another way is the mean value of observations in each set as the 'true' value, which is finally adopt in this research (equation 5.82).

$$y_i^{true} = \sum_{i=1}^{n} \frac{f(tp1_i, tp2_i, \cdots, ntp1_{i,j}, ntp2_{i,j}, \cdots)}{n} \tag{5.82}$$

Equation 5.81 has been negated with this research going deeper. Due to the interactions between targeted and non-targeted parameters, we cannot say that using mean values of non-targeted parameters gives us ideal values of y. On the other hand, it is very common in statistic to enhance the sample size and use the mean value of multiple samples to cancel out the fluctuation. For that reason, statistically, it is more reasonable to use $\overline{y_i^{ob}}$ as y_i^{true}.

The formulas of R_U^2 and R_E^2 can be reformed as 5.83 and 5.84.

$$R_U^2 = 1 - \frac{\sum_{i=1}^{n} \left(\sum_{j=1}^{m} \frac{(y_{i,j}^{ob} - \overline{y_i})^2}{m} \right)}{\sum_{i=1}^{n} (\overline{y_i} - \overline{y})^2} \tag{5.83}$$

$$R_E^2 = 1 - \frac{\sum_{i=1}^{n} (y_i^{pd} - \overline{y_i})^2}{\sum_{i=1}^{n} (\overline{y_i} - \overline{y})^2} \tag{5.84}$$

5.6.3 Demonstration of STMN method with Ishigami function

To validate the proposed STMN method, we still used the Ishigami function. The configurations (table 5-6) were identical to that in section 5.1.1 where we did the demonstration of *relative uncertainty* R_U^2. In this demonstration, x_1, x_2, x_3 were regarded as the targeted parameters, while a, b, c were non-targeted parameters.

Table 5-6: Configurations of demonstration with Ishigami function

	x_1	x_2	x_3	a	b	c
case 1	$[-\pi, \pi]$	$[-\pi, \pi]$	$[-1, 1]$	[6,8]	[0.05,0.15]	[-0.5,0.5]
case 2				[4,10]	[0.0,0.2]	[-1.0,1.0]
case 3				[2,12]	[-0.1,0.3]	[-2.0,2.0]
case 4				[0,14]	[-0.2,0.4]	[-3.0,3.0]

Firstly, we generated 50 samples of x_1, x_2, x_3. For each sample, we randomly generate 10 samples of a, b, c and paired them with the sampled x_1, x_2, x_3. 500 cases were generated in total, the corresponding y^{ob}

values were calculated. The mean value of y^{ob} of each sample set was used as the y^{true}. Relative uncertainty R_U^2 was calculated. Then, a GPR model was trained with targeted parameters and y^{ob} of the 500 cases.

For comparison, other 500 cases, including both targeted and non-targeted parameters, were generated using common Monte Carlo method. Another GPR model was trained.

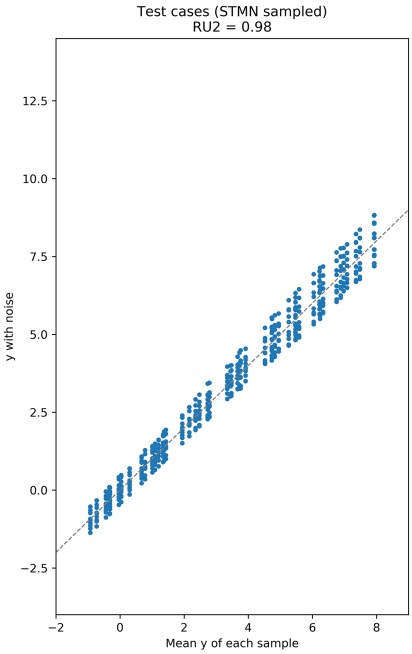
To test the trained GPR model, we generated other 50 sample sets, including $(50 \times 10 =)500$ cases in total, using STMN method. We fed the targeted parameters into the trained GPR models and get y^{pd*}. R_E^2 of $\overline{y^{ob*}}$ and y^{pd*} was analyzed, which is concluded in table 5-7.

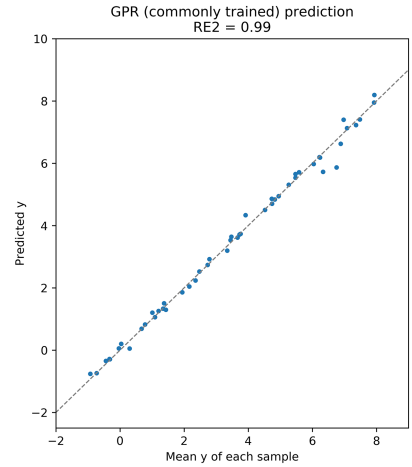
Table 5-7: Configurations of demonstration with Ishigami function

	STMN			**Commonly**	
	Training sample size	Testing sample size	R_E^2	Training sample size	R_E^2
case 1	$50 \times 10 = 500$	$50 \times 10 = 500$	1.00	500	0.99
case 2			1.00		0.95
case 3			0.99		0.93
case 4			0.99		0.82

From table 5-7 and figures 5-20-5-23, we can see that with the range of non-targeted parameters getting wider and fluctuation becoming larger, the STMN method shows better performance than a common way.

Based on our personal experience, if the value of *relative uncertainty* R_U^2 is lower than 0, STMN method should be adopt.





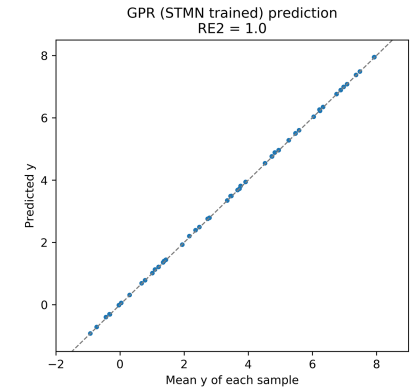
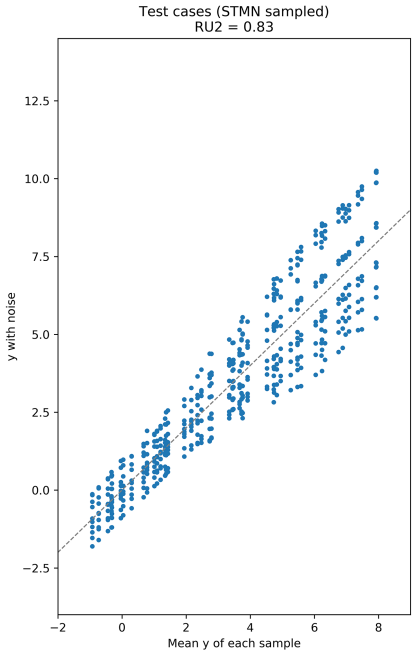
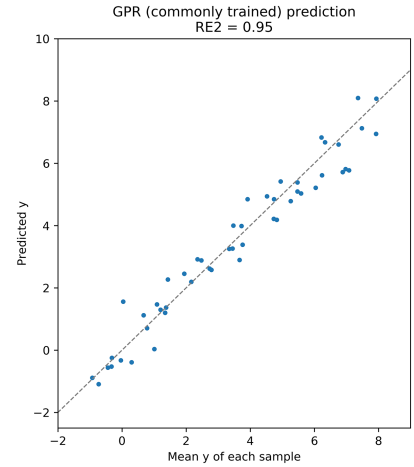


Figure 5-20: Test points and results of case 1





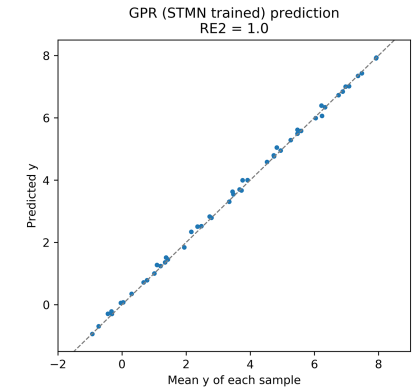
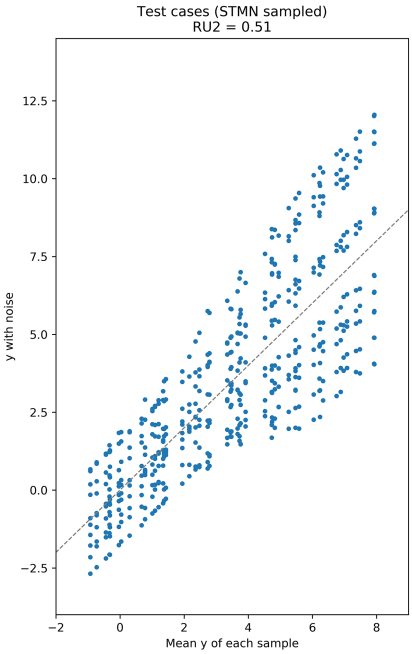
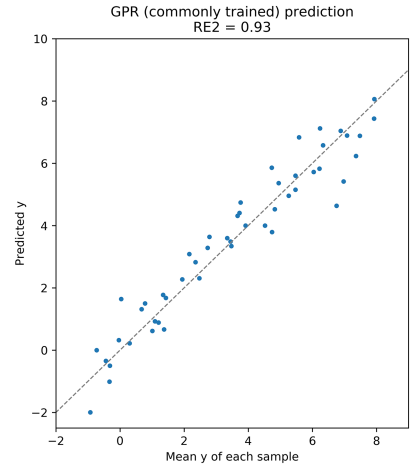


Figure 5-21: Test points and results of case 2





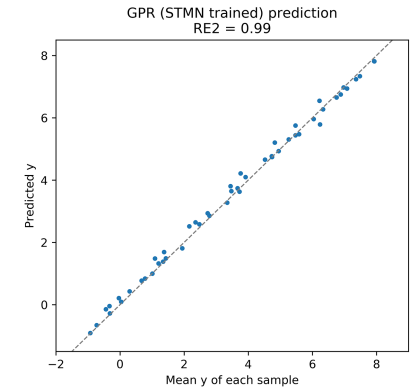
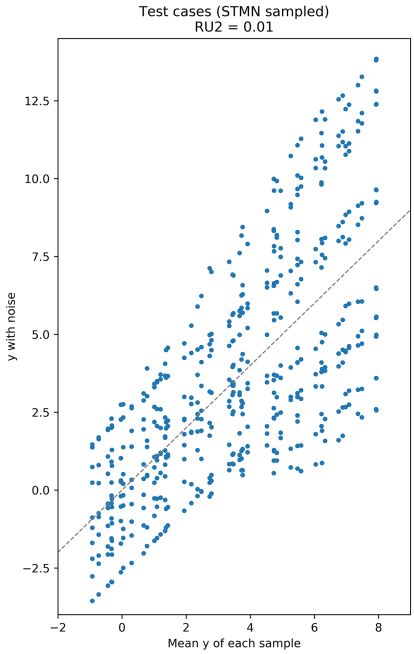
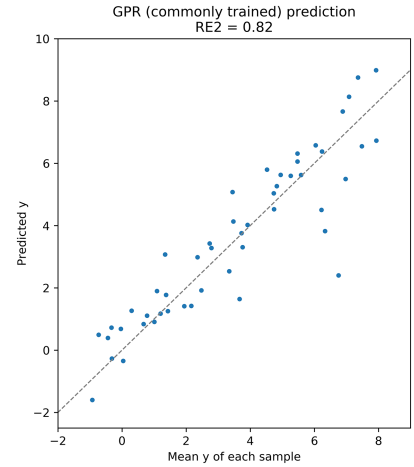


Figure 5-22: Test points and results of case 1





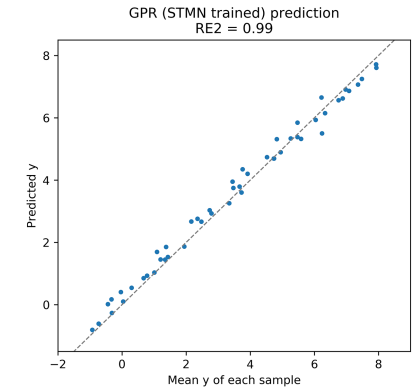


Figure 5-23: Test points and results of case 2

Chapter 6

Case study

In this chapter, the proposed decision-making framework for the early design stage is demonstrated with an imaginary office project located in Tokyo. 23 targeted parameters were studied, as well as 4 non-targeted ones. The sensitivity of each parameter was quantified using Morris method. The left 19 parameters were separated into 4 groups based on their interactions which is quantified using Extended Morris method. Using the meta-models made with GPR, a database including 20 thousand cases was made. Microsoft Power BI was used to visualize the data interactively and make dashboards. Additional predictions were made and added into the database to make the histograms smooth when the range of parameters getting small. Even in the case that some modifications were made in the base model, new dashboard could be made very efficiently as the whole process was highly automated.

6.1 Introduction

6.1.1 Description of the case study

In order to demonstrate the proposed decision-making framework for the design stage, we did a case study with an imaginary office project.

The site of this project was located in the suburban area of Tokyo. The dimension of this site was about 30mx50m, which was a tiny and typical size in Japan. We supposed that, to the south of the site, there was a small lake which provided beautiful scenery. The main entrance of this building was located on the east facade, facing a road. On the other side the road, there was a other mid-rise building, from which the distance to the east boundary of this site was about 15m. The west side was adjacent to another building, with a distance of 6m. The orientation of this site was 15 degree south east.

Though the orientation and aspect ratio have important influence on the energy performance, these two factors are often constrained by the tiny site. On the other hand, we believe that the designing considerations might be more important than energy performance when deciding the orientation and volume mass of a building. We adopted the office plane form *A proposal of standard plan for energy research (office building)*[3], a floor plan consisting of two office rooms in south and north and one core, which was quite widely used in energy performance study in Japan. Figure 6-1 illustrates the master plan and the standard floor plan of this project.

We imagined that the designers decided to make big openings on the south facade to make full use of the nice view from south, but the engineers stood on opposite site. Based on the climate in Tokyo, making big opening on south, even with overhangs, would result in high cooling energy consumption. Though in residential buildings, openings on the south are good for collecting solar energy and reducing heating load in winter, as the internal heat gain density in office buildings is much higher, the cooling load is dominant in in warm area. Attaching overhangs to the south facade is a effective solution that shades the openings from solar radiance and reduces the cooling load, but blocks some part of the view. How to balance the view and energy performance would be a key point in this design. In this case, a design dashboard could help when the designers negotiates with the engineers. On the east facade, as there was a distance to the other building, shadings were necessary. Fins were attached to east facade, to shade the building from the rising sun, which were also important visual elements. How to decide to size of the fins, balance the appearance and performance was another question. Figure 6-2 illustrate the appearance image of this project.

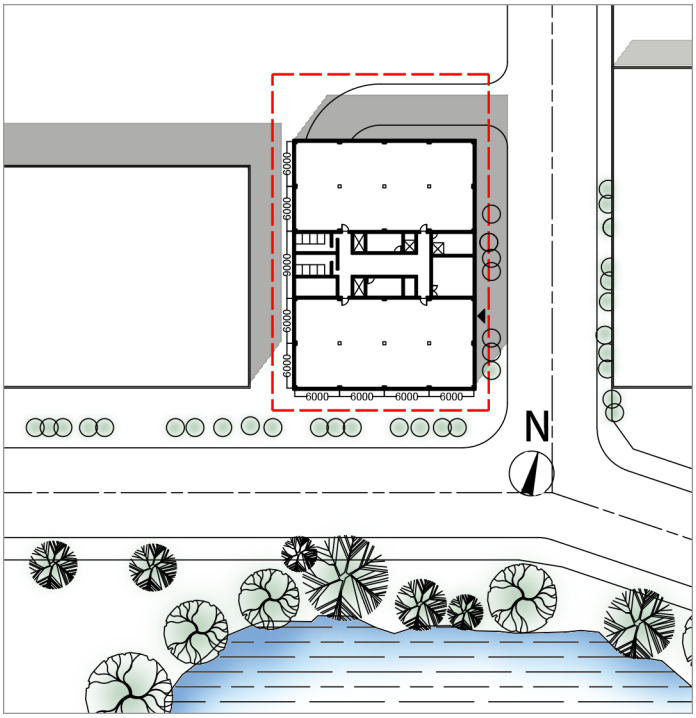


Figure 6-1: Master plan and standard floor plan

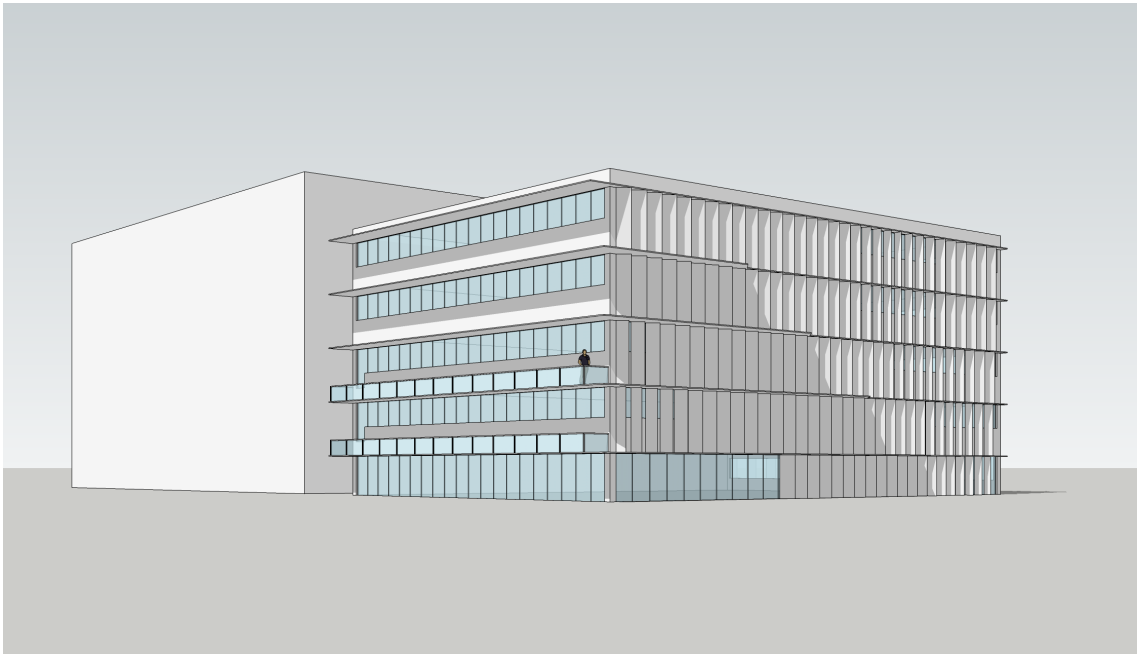


Figure 6-2: The appearance of east and south facades

6.1.2 Parameters and criteria

In this study, we mainly focused on the envelop performance.

The sizes of the openings on all 4 facades were represented by *Window to wall area ratio (WWR)*. As this research aims at the early stage, we simply represented the performance of the window with *Solar heat gain coefficient (SHGC)* and *U-factor* of the glass. Theoretically, the *SHGC* of a window also includes the influence from the window frames and the shading devices. In this case study, *SHGC* was simply used as a transparency property of window glass.

Figure 6-3 illustrates the section of a overhang. The ratio between overhang depth and window height was used to represent the sizes of shading devices on the south facade, which is called *overhang scale* in this case study. The overhangs are attached to the upper edges of the openings so that the occupant can widest view from the lake. The depth and width of the fins on the east facade were also studied. As shown in figure 6-4, the word depth meant the size of a fin on south-north axis, while width meant that on east-west axis. The distance between 2 adjacent fins was 1m.

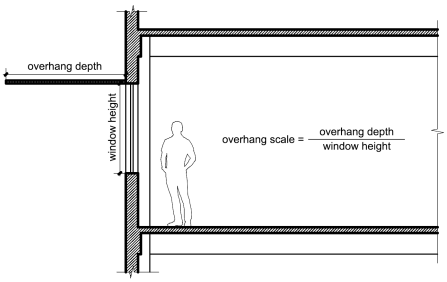


Figure 6-3: Overhang section

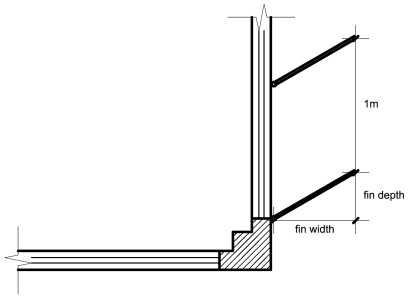


Figure 6-4: Fin plan

The insulation of the opaque part of the envelop is also studied. It is a little difficult to adjust the *U-value* of a wall directly in *EnnergyPlus*. For that reason, thickness of insulation layer was used to represent the insulation of opaque part of the envelop, which is an abstract value and has no influence on the geometry of the building. The solar absorptance of the outside layer of the wall is also studied, as it is related to the color or texture of the facade, which designers might be interested in.

Table 6-1 shows the ranges of all targeted parameters.

4 non-targeted factors were taken into consideration.

Cooling and heating setpoints have overwhelming influence on the AC energy consumption in building. In an office building, based on the AC design handbook, $24°C$ is used as cooling setpoint and $21°C$ as heating setpoint. However, duo to the sensor errors, occupants' behaviors, etc., there are surely uncertainties in the real AC setpoints. In this study, random values were used rather than constant setpoints.

Table 6-1: Targeted parameters

Parameter	Unit	Range	Description
WWR (E, S, W, N)		[0.2, 0.8]	Window to wall area ratio
SHGC (E, S, W, N)		[0.1, 0.9]	Solar heat gain coefficient of the glass
U-factor of windows (E, S, W, N)	W/m^2K	[0.5, 6]	Represent the heat transfer rate of the transparent part of the facade
Overhang scale (S)		(0, 2]	Overhang depth / window height
Fin width (E)	m	(0, 0.9]	The size of a fin on east-west axis
Fin depth (E)	m	(0, 0.9]	The size of a fin on south-north axis
Thickness of insulation (E, S, W, N)	m	(0, 0.2]	Represent the insulation of the opaque part of the facade
Solar absorptance of external surface (E, S, W, N)		[0.1, 0.9]	Related to the color of the opaque part of facade

The heat generation of electric machines, directly influencing the AC energy performance, varies quiet a lot from building to building, based on the business types. Surely, there are also uncertainties in the heat gain from people, lighting, etc. In this case study, we resolved all the uncertainties of internal heat gain in the electric equipment heat generation.

In addition to the necessary fresh air, natural ventilation is also considered. It has been found that by introducing ventilation in non-air-conditioned period, especially in summer night, can effectively reduce AC load. However, the effect of natural ventilation is quiet unstable, duo to the external wind and occupants' behaviors. We used a random air change rate to reflect the uncertainties in the natural ventilation.

In the Gaussian process regression, these non-targeted parameters were regarded as with uncertainty, as they could not be precisely predicted in early stage. The values and ranges of the non-targeted parameters could neither be controlled in dashboards. Table 6-2 shows the ranges of all non-targeted parameters.

Table 6-2: Non-targeted parameters

Parameter	Unit	Range	Description
Cooling set point	$°C$	[23, 30]	The room will be cooled if [room temperature >= Clsp]
Heating set point	$°C$	[23, 30]	The room will be heated if [room temperature <= Htsp]
Internal heat gain	W/m^2	[3, 15]	The heat generated by electric equipment in unit area
Air change rate	$1/hour$	[0, 6]	Natural ventilation in addition to the necessary air change

Both energy performance and thermal comfort were studied in this demonstration.

Hourly heating and cooling load were calculated using *EnergyPlus*. The 8760 values were summed up and divided by the floor area, converted into annual energy density, to evaluate the energy performance. Peak load is another important index that should be considered. Admittedly, as TMY weather data was used in this case study, the simulated results was adequate to evaluate the capacity of the HVAC system, but was able

to reflect the real peak loads to some extend. The top 10 heating/cooling load from the 8760 results were picked up. Their mean value was used to represent the heating/cooling peak load. for the reason that the occasionality in the very maximum and minimum values was quite big.

Operative temperature was used to study the thermal comfort. We defined the comfortable interval of operative temperature as, 20°C to 24°C in heating period; 21.5°C to 25°C in non-air-conditioned period; 23°C to 26°C in cooling period. The percentage of comfortable, hot and cold hours was used to represent the thermal comfort level of the building. Annual mean operative temperature, the mean value of all 8760 results, was also calculated. Besides, we picked up the top 10 and bottom 10 operative temperature from 8760 results and use their mean values to study the extreme situations.

Table 6-3 shows the criteria.

Table 6-3: Criteria

Parameter	**Unit**	**Description**
Annual cooling energy density	kWh/m^2	The cooling energy consumed by unit area. COP is 3.2
Annual heating energy density	kWh/m^2	The heating energy consumed by unit area. COP is 3.5
Cooling peak	kW	The average value of the top 10 hourly cooling rate from 8760 results
Heating peak	kW	The average value of the top 10 hourly heating rate from 8760 results
Annual mean operative temperature	°C	The average value of hourly operative temperature of all 8760 results
Hot operative temperature	°C	The average value of the top 10 hourly operative temperature from 8760 results
Cold operative temperature	°C	The average value of the bottom 10 hourly operative temperature from 8760 results
Comfort OT percentage	%	Cooling period: [23, 26]; non-AC period: [21.5, 25]; Heating period: [20, 24]
Hot OT percentage	%	Cooling period: (0, 23); non-AC period: (0, 21.5); Heating period: (0, 20)
Cold OT percentage	%	Cooling period: (26, 100); non-AC period: (25, 100); Heating period: (24, 100)

6.1.3 Simulation configurations

EnergyPlus 8.8 was used to execute the simulations. The template of office building in climate zone 3A from ASHRAE 90.1 2010 was used, which includes the constructions and material properties; people, lighting and electricity equipment configurations; as well as all kinds of schedules. This template from ASHRAE is based on the lifestyle in U.S. Some modifications have been done in this template to make it closer to the situations in Japan.

The office rooms are air-conditioned from 6:00 to 22:00. The Clsp and Htsp, 2 non-targeted parameters, were decided randomly. From 22:00 to 6:00 in the next morning, the Clsp is $0°C$ and the Htsp is $100°C$, which is equivalent to non-air-conditioned.

The people's activity level was $132w/person$ throughout the whole day, following the template value. The density of people was enhanced to $0.125person/m^2$ ($8m^2/person$) from the original $0.05person/m^2$, a American style. The office rooms were occupied from 6:00 to 22:00. Figure 6-5 shows the occupancy curve in one day.

The lighting power density is $9.68w/m^2$, following the template, controlled by the schedule illustrated in figure 6-6. The equipment heat generation density was another non-parameter controlled by the schedule illustrated in figure 6-7, which also reflected the uncertainties in people and lighting.

Depends on the air-conditioning design codes in Japan, the minimum air change rate in office space should be no less than 6 times, which is about $18m^3/m^2h$. We suppose that heat recovery is installed, so we set the mechanical ventilation volume $9m^3/m^2h$ to simply reproduce that 50% of heat is recovered. Additionally, we also introduce natural ventilation into the room. The ventilation volume is simply represented with constant ACR, from 0 to 6, which is one of the non-targeted parameters. The natural ventilation is controlled by external air temperature. The office rooms will be naturally ventilated if the outside temperature is higher than (Htsp+Clsp)/2 and lower than Clsp, no matter the room is occupied or not.

Table 6-4 shows the simulation configurations. Besides, the weekend is not considered. The weekday schedules are used throughout the whole year.

Table 6-4: Simulation configurations

Item	**Unit**	**Base value**	**Schedule**
Cooling setpoint	$°C$	[23, 30]	6:00 to 22:00
Heating setpoint	$°C$	[16, 23]	6:00 to 22:00
People	W/m^2	16.5	Figure 6-5
Electric equipment	W/m^2	[3, 15]	Figure 6-7
Lighting	W/m^2	9.68	Figure 6-6
Mechanical ventilation	m^3/m^2h	9	6:00 to 22:00
Natural ventilation	$1/hour$	[0, 6]	(Htsp+Clpt)/2 <= Out temp <= Clsp

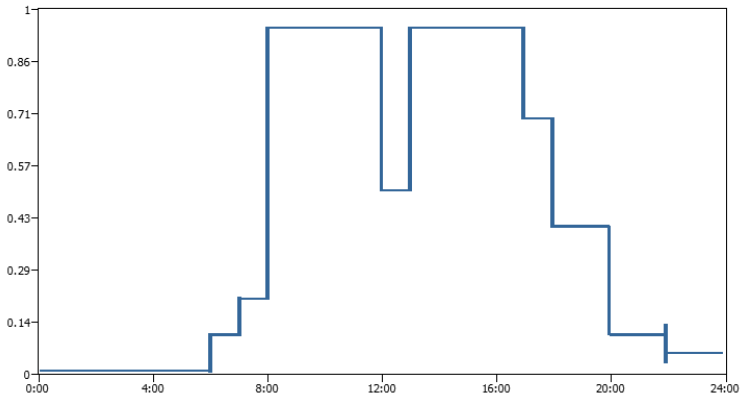


Figure 6-5: Occupancy schedule curve

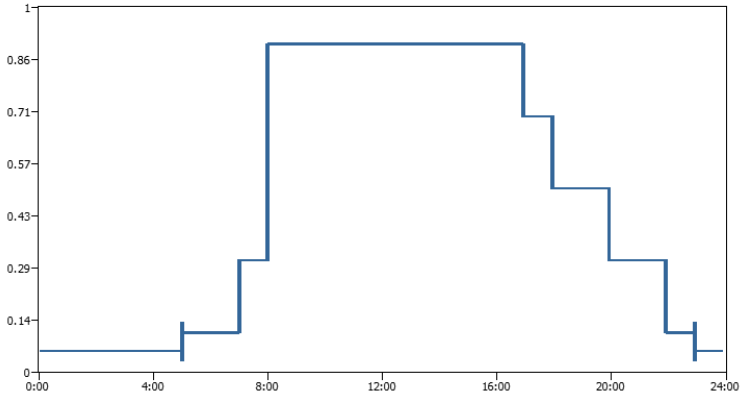


Figure 6-6: Lighting schedule curve

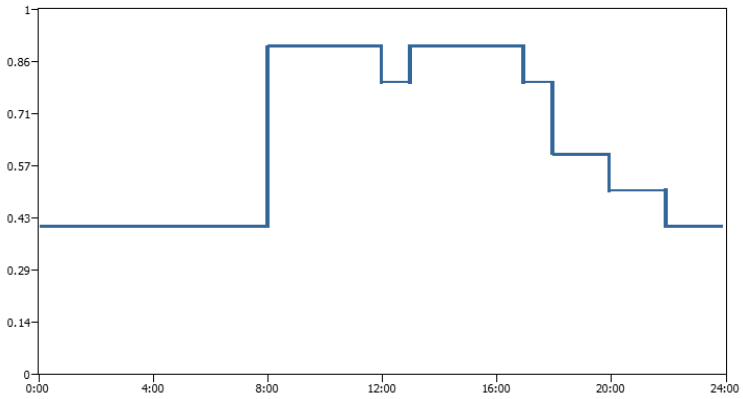


Figure 6-7: Equipment schedule curve

The constructions (Table 6-5&6-6) followed the template, besides the thickness of *Mass Wall Insulation* and the solar absorptance of outside layer are variable.

Table 6-5: Construction of external walls

	External wall			
	Outside	**Mid Layers**		**Inside**
Material	Stucco	Concrete	Mass Wall Insulation	Gypsum
Roughness	Smooth	Rough	MediumRough	Smooth
Thickness$[m]$	0.0253	0.2032	**(0, 0.2]**	0.0127
Conductivity$[W/mK]$	0.69	1.311	0.049	0.16
Density$[kg/m^3]$	1858	2240	265	784.9
Specific Heat$[J/kgK]$	837	836.8	836.8	830
Solar Absorptance	**[0.1, 0.9]**	0.7	0.7	0.4

Table 6-6: Construction of internal surfaces

	Floor/Ceiling			**Internal wall**		
	Floor surf		**Ceiling surf**	**Symmetrical**		
Material	Lightweight concrete	Air space	Acoustic tile	Gypsum	Air space	Gypsum
Roughness	Mid Rough		Smooth	Mid Smooth		Mid Smooth
Thickness$[m]$	0.1016	0.05	0.0191	0.019	0.04	0.019
Conductivity$[W/mK]$	0.53	0.278	0.06	0.16	0.278	0.16
Density$[kg/m^3]$	1280		368	800		800
Specific Heat$[J/kgK]$	840		590	1090		1090
Solar Absorptance	0.5		0.3	0.4		0.4

6.1.4 Climate Data

Meteonorm[73] was used to get the EPW files used in the simulations. The TMY data of Tokyo, representing the historical climate record, was exported. We call this unmarked climate data 'TMY' in the following sections. Climate change was considered in this case study. We also exported the predicted climate data of 2020, 2050 and 2080 using *Meteonorm*. A1B scenario was selected, a scenario included in the 5th assessment report from IPCC[74], which described a future that economy was rapidly developing but environment protecting was also considered. The prediction data from HadCM3, *Hadley Centre* Coupled Model version 3

6.2 Parametric modeling

We firstly generated a empty model of office building in climate zone 3a in *OpenStudio*[77], a modeling tool for *EnergyPlus*. Though *OpenStudio* is powerful at editing geometries, as it is *SketchUp* based, in this case study, what we needed is the template configurations that will be imported when we start a new *OpenStudio* project. A IDF file without geometric information was exported and modified in *IDFEditor* as discribed in section 6.1.3.

One floor with the standard plan from the building was extracted to do the calculation. We built a BIM model in *Autodesk Revit*. The coordinates of the vertices of the surfaces was extracted and exported into a CSV file in *Dynamo*, a parametric modeling plug-in. (Figure 6-8) Actually, at the very beginning, we intended to build the Building energy model (BEM) fully within Revit. Then we soon found it was too difficult at current stage. Admittedly, compared to building geometries in *OpenStudio/SketchUp*, there was no advantages to do it in *Revit*. In this research, we just tried to establish a start point of the integration of BIM and parametric study of energy performance.

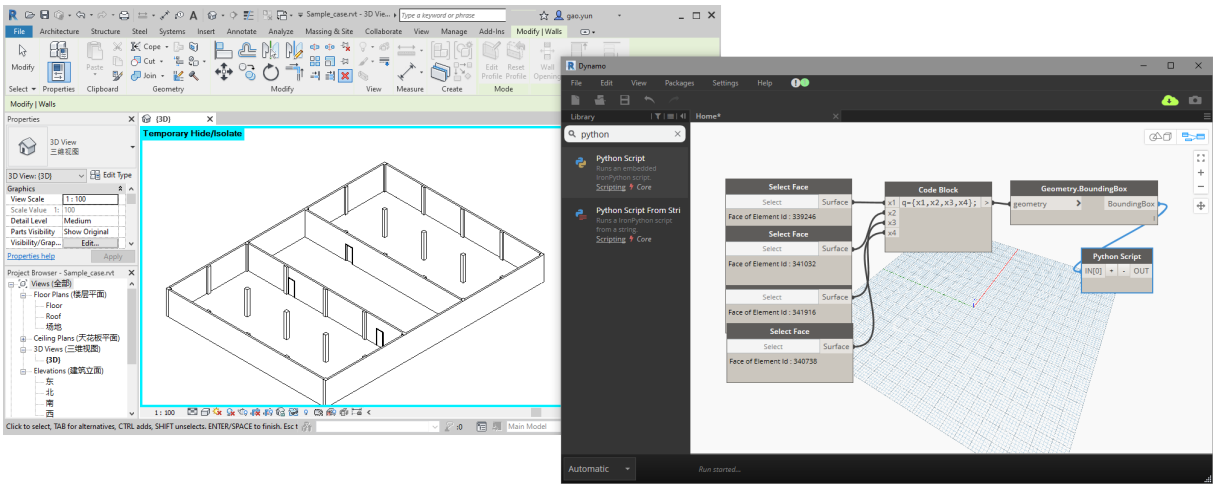


Figure 6-8: Revit model and Dynamo program

EPPiX was used to add geometries, based on the CSV file exported by *Dynamo*, into the modified template IDF file. 3 closed spaces were formed. As we just employed one standard floor in simulation, the floors and ceilings were set adiabatic. The heat transfer through the floors and ceilings was blocked, while their thermal mass was considered. The base IDF file was finished. There is another way to build a single-floored energy model in *EnergyPlus*. If the volume a zone is manually configured, it not necessary to make the zone closed by geometries. So that the adiabatic surfaces can be ignored. This process could not be adopted in this case study. As operative temperature is studied, the components of floors and ceilings are necessary.

Samples were generated using Latin Hypercube sampling. The value of each parameter was within [0, 1] in this step. The samples would be extended in sensitivity and interaction analysis. (see section 4.3.2 & 4.4.2) The sampled value of a parameter was remapped from [0, 1] to the range of that parameter. Before

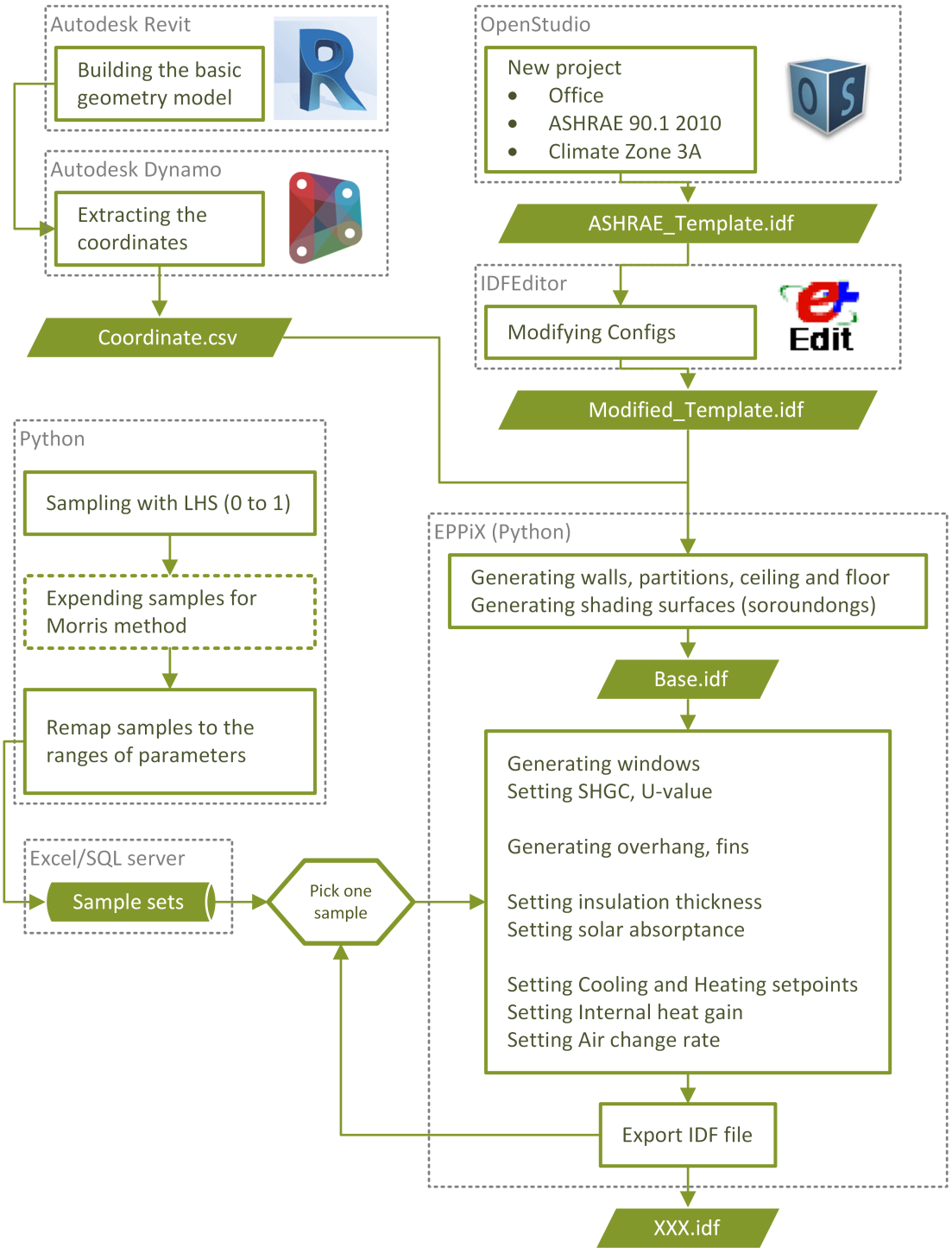


Figure 6-9: Parametric modeling flowchart

generating models, all the remapped samples were saved in database, *Excel* or *SQL Server*, which made it more smooth to do analysis in the latter phases. The data volume was far beyond that proper for a CSV file.

The program generated a IDF file for each row in the database, using *EPPiX*. As illustrated in figure 6-10, An opening was made by linearly scale the parent surface along the center point by $\sqrt{WWR}$. Then the height of the opening was calculated and multiplied by the overhang scale, to get the overhang depth. A horizontal overhang was attached to the upper border of the south opening. The width of the overhang is equal to the south facade, as illustrated in figure 6-11. 12 fins were attached to the east facade of each office room, but not the core room. 24 fins were generated in total. Other configurations were also renewed as described in section 6.1.2. Then an IDF file was exported. Figure 6-9 illustrates the process of parametric modeling.

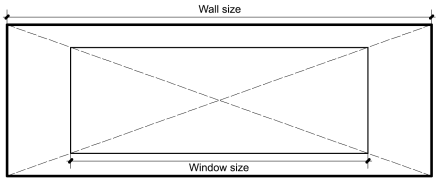


Figure 6-10: Overhang section

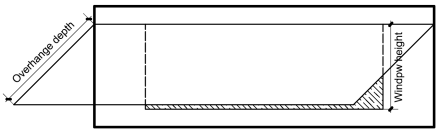


Figure 6-11: Fin plan

6.3 Sensitivity analysis and screening parameters

The sensitivity of each parameter over each criterion was analyzed using Morris methods. We firstly sampled 20 initial cases using LHS. Each case was expended into 28 cases as there are 27 parameters. The TMY EPW file was used in this step.

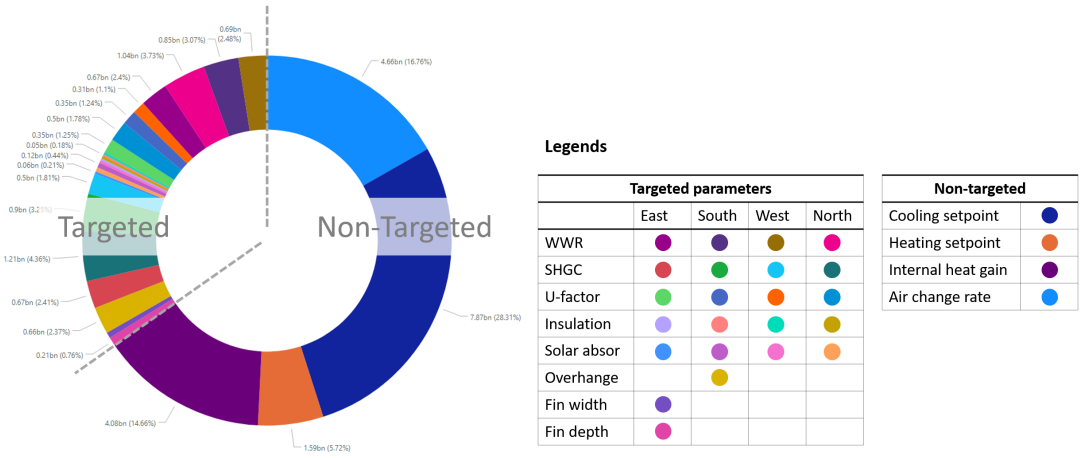


Figure 6-12: Sensitivities over annual AC load

As illustrated in figure 6-12, the sensitivities of non-targeted parameters were overwhelms that of targeted ones, which would disturb the data analysis and parameter screening. (This also indicated that the real green building was actually green life style.∧ω∧) On the other hand, designers cannot really control them. For these reasons, we did not take the non-parameters into consideration when determining the threshold of sensitivity of each criterion, otherwise all the targeted parameters might be less important and skipped in latter phases. For each criterion, the threshold was set to the half of the average value of the sensitivities of targeted parameters. (Equation 6.1)

$$Threshold^{ck} = 0.5 \times \sum_{i=1}^{23} \frac{S_{ti}^{ck}}{23} \tag{6.1}$$

Where:

$Thres_{ck}$ = The threshold of sensitivity over the k^{th} criterion

$Sens_{ti}^{ck}$ = The global sensitivity of the i^{th} targeted parameter over the k^{th} criterion

A parameter would be skipped in further phases if its sensitivity was lower than the threshold for every criterion. In another word, if one of the sensitivities exceeded the corresponding threshold, that parameter should be considered in latter phases. Table 6-7 shows the sensitivities and the thresholds. The cell of a less important parameter was displayed in light gray.

As expected, the sensitives of WWR and overhang scale on south facade is quite high, especially over cooling load, which indicates that the designer should be very carefully with the south openings.

The size of the fins are not very important in this case. The sensitivities of the fin width fail to pass the thresholds, which means that once there were fins, no matter the width, they worked. The sensitivities of fin depth over cooling load and peak exceed the thresholds a little. It can be a good news for the designer, as the most important appearance elements, fins, are less important to the energy consumption but necessary, so that they can be creative when designing the east facade.

It is found that, in this case the solar absorptance on the annual AC load was almost negligible, which means that the designer could feel free to decide the color and the material of the external surface. From the data, we can find that the sensitivities of insulation over cooling and heating, load and peak, are much less than the thresholds, but the insulation of the south facade influences the operative temperature.

It was beyond expectation that the insulation's sensitivities over the AC load were lower than the thresholds, which seemed incompatible with common sense. We thought it might because excessive insulation did not benefit the energy performance, as the thermal transfer had been low enough with insulation level that met the codes. We did a additional analysis on the sensitivities of insulation to verify this hypothesis. Besides the insulation thickness in all orientations, other parameters were constantly fixed to the middle values of their ranges. The four insulation-thickness parameters shared the same values. 100 samples were evenly taken within the interval of [2cm, 20cm]. The step length for Morris analysis is 2cm. Figure 6-13 and 6-

Table 6-7: Sensitivities of parameters over each criterion

Parameters	Total load	Heating load	Cooling load	Heating peak	Cooling peak	Operative temp
			East facade			
WWR E	6.68E+08	2.61E+08	4.07E+08	1.88E+06	7.19E+05	3.42E-03
SHGC E	6.70E+08	5.38E+07	6.16E+08	2.06E+05	6.22E+05	1.92E-03
U-factor E	3.46E+08	2.52E+08	9.38E+07	1.75E+06	2.15E+05	3.46E-03
Fin width E	1.38E+08	1.77E+07	1.20E+08	7.75E+04	1.25E+05	1.32E-03
Fin depth E	2.1E+08	1.29E+07	1.97E+08	1.00E+05	2.01E+05	1.25E-03
Insulation E	5.18E+07	3.49E+07	1.69E+07	3.25E+05	3.82E+04	1.00E-03
Solar absorpt E	5.84E+07	9.97E+06	4.85E+07	1.17E+05	4.23E+04	1.12E-03
			South facade			
WWR S	8.54E+08	2.23E+08	6.30E+08	1.60E+06	8.09E+05	5.84E-03
SHGC S	9.04E+08	1.45E+08	7.58E+08	3.57E+05	6.64E+05	3.85E-03
U-factor S	3.46E+08	2.44E+08	1.02E+08	1.56E+06	2.26E+05	4.77E-03
Overhang scale	6.59E+08	8.96E+07	5.70E+08	2.44E+05	3.69E+05	3.56E-03
Insulation S	4.89E+07	3.55E+07	1.34E+07	3.32E+05	3.59E+04	1.60E-03
Solar absorpt S	1.23E+08	3.49E+07	8.77E+07	1.25E+05	6.09E+04	1.27E-03
			West facade			
WWR W	6.91E+08	2.59E+08	4.32E+08	1.85E+06	7.95E+05	3.30E-03
SHGC W	5.03E+08	5.62E+07	4.46E+08	2.63E+05	5.32E+05	1.77E-03
U-factor W	3.06E+08	2.07E+08	9.93E+07	1.46E+06	2.39E+05	2.66E-03
Insulation W	5.42E+07	3.88E+07	1.54E+07	3.11E+05	3.29E+04	8.30E-04
Solar absorpt W	6.28E+07	1.64E+07	4.65E+07	9.65E+04	3.21E+04	1.03E-03
			North facade			
WWR N	1.04E+09	2.71E+08	7.65E+08	1.83E+06	9.60E+05	4.94E-03
SHGC N	1.21E+09	1.16E+08	1.10E+09	4.65E+05	1.09E+06	2.79E-03
U-factor N	4.96E+08	2.99E+08	1.97E+08	1.79E+06	2.64E+05	3.86E-03
Insulation N	6.52E+07	4.56E+07	1.95E+07	3.31E+05	3.42E+04	1.33E-03
Solar absorpt N	1.06E+08	2.22E+07	8.41E+07	1.25E+05	5.80E+04	6.42E-04
			Non-targeted			
Cooling setpoint	7.87E+09	2.61E+08	7.61E+09	7.17E+05	2.54E+06	4.73E-02
Heating setpoint	1.59E+09	1.30E+09	2.89E+08	5.50E+06	1.48E+03	3.50E-02
Internal heat gain	4.08E+09	7.17E+08	3.36E+09	2.79E+06	2.46E+06	5.72E-03
Air change rate	4.66E+09	2.50E+08	4.41E+09	1.54E+05	1.67E+05	1.43E-02
Threshold	4.18E+08	5.97E+07	1.49E+08	3.74E+05	1.77E+05	1.34E-03

Legend:
Less important

Sensitivity is an unit-less relative amount

14 illustrates that, with the insulation getting higher, the sensitivities over the AC load get lower. After the thickness getting 5cm, meeting the ASHRAE codes, the sensitivities get lower than the thresholds. After 10cm, the sensitivities gets close to 0, which means there is no longer benefit from add insulation.

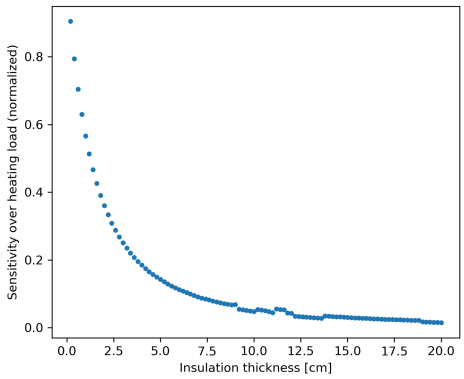


Figure 6-13: Sensitivity over heating load

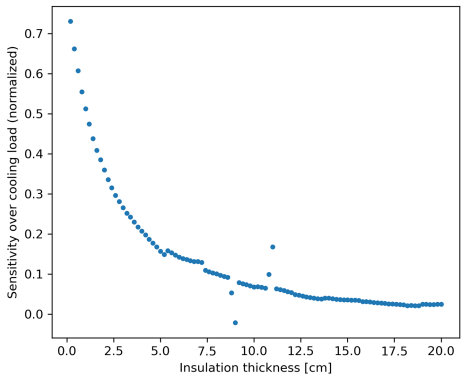


Figure 6-14: Sensitivity over cooling load

We also found that the sensitivities of parameters on the north facade were far more than expected, even more than that on the south. The north facade is studied in section 6.8

Another thing is that heating setpoint has high sensitivity over cooling load, as well as that cooling setpoint has high sensitivity over heating load. The reason might be the natural ventilation. Illustrated in table 6-7, the air change rate plays an important role in the AC load. On the other hand, the control of natural ventilation is decided by the AC setpoint. The AC setpoints were not the targeted parameters, so that we did not look deeper into this phenomenon.

20 cases, though enough to screen the parameters, were too rough to visualize the sensitivities in latter phases. We did additional sensitivity analysis for the left 19 parameters. 250 cases were generated using LHS. Each case was expanded into 20 cases. In this step, we also took climate change into consideration. Each case was analyzed under 4 climates, TMY, 2020, 2050 and 2080 from *Meteonorm*. 20 thousand times of simulations were executed in total. A relational database was established to record the sensitivity data, which included 6 tables. The first table included the values of the 19 parameters of all cases. The other 5 tables corresponded to the 5 criteria. Each table included the sensitivities of the 19 parameters over that criterion. The illustration of database structure can be found in figure 6-24.

6.4 Interaction analysis and grouping parameters

The interactions between the left 19 parameters over each criterion were analyzed using Expanded Morris method. 20 initial cases were generated using LHS, and each case was expanded into 191 cases. The TMY climate data was used.

The results of interaction analysis were relative values without physical meanings, which is difficult to read. Normalization was necessary. By common sense, WWR and U-factor on the same orientation should be strongly interactive. The average value of the interactions between WWR and U-factor on the same orientations over each criteria was used as the baseline value (Equation 6.2). The interaction values were normalized by being divided by the corresponding baseline values (Equation 6.3). The threshold for each criterion was set to 0.5.

$$Baseline^{ck} = \sum \frac{Inter^{ck}_{w^o,u^o}}{4} \tag{6.2}$$

$$NorInter^{ck}_{i,j} = \frac{Inter^{ck}_{i,j}}{Baseline^{ck}} \tag{6.3}$$

$$Threshold^{ck} = 0.5 \tag{6.4}$$

Where:

o	$\in$	[East, South, West, North], Orientation
w^o	=	WWR parameter on orientation o
u^o	=	U-factor parameter on orientation o
$Inter^{ck}_{w^o,u^o}$	=	Interaction value between w^o and u^o over the k^{th} criterion
$Baseline^{ck}$	=	Baseline value of interactions over the k^{th} criterion
$Inter^{ck}_{i,j}$	=	Interaction value between i^{th} and j^{th} parameters over the k^{th} criterion
$NorInter^{ck}_{i,j}$	=	Normalized interaction value between i^{th} and j^{th} parameters over the k^{th} criterion
$Thres_{ck}$	=	The threshold of sensitivity over the k^{th} criterion

The normalized data was putted in table 6-8, 6-9 and 6-10. Figure 6-15 illustrates the interactions between parameters over all criteria.

Consistent with common sense, the parameters of each orientation strongly interactive with each other, except the insulation on South. The interactions between parameters belonging to different orientations are quite weak. All the targeted parameters are interactive with 2 or more non-targeted parameters, over a least on criteria.

Table 6-8: Normalized interaction matrices over AC load

Cooling load

	Fin depth	WWR E	U-f E	SHGC E	Over-hang	WWR S	Thick S	U-f S	SHGC S	WWRW	U-f W	SHGC W	WWR N	U-f N	SHGC N	Clsp	Htsp	In heat	Air ch
Fin depth	—	1.27	0.21	1.76	0.14	0.14	0.09	0.09	0.07	0.09	0.09	0.11	0.1	0.12	0.1	0.84	0.12	0.23	0.86
WWR E	True	—	1.0	4.01	0.26	0.31	0.07	0.17	0.34	0.18	0.17	0.23	0.29	0.12	0.3	2.02	0.17	0.34	1.55
U-factor E	False	True	—	1.04	0.16	0.22	0.09	0.17	0.25	0.21	0.24	0.21	0.24	0.16	0.25	0.47	0.16	0.52	2.4
SHGC E	True	True	True	—	0.11	0.2	0.06	0.2	0.17	0.17	0.16	0.14	0.19	0.16	0.11	1.7	0.15	0.46	1.9
Overhang	False	False	False	False	—	5.35	0.1	0.55	5.84	0.41	0.32	0.18	0.13	0.11	0.1	2.24	0.27	0.74	6.5
WWR S	False	False	False	False	True	—	0.16	1.2	5.66	0.31	0.29	0.27	0.15	0.12	0.12	2.8	0.21	0.69	4.19
Thick S	False	False	False	False	False	False	—	0.07	0.08	0.06	0.1	0.11	0.14	0.11	0.08	0.19	0.06	0.1	0.33
U-factor S	False	False	False	False	True	True	False	—	2.13	0.22	0.22	0.23	0.13	0.17	0.13	0.52	0.15	0.61	3.51
SHGC S	False	False	False	False	True	True	False	True	—	0.36	0.26	0.15	0.09	0.09	0.09	2.49	0.2	0.84	4.63
WWR W	False	False	False	False	False	False	False	False	False	—	0.98	4.49	0.35	0.14	0.3	2.14	0.21	0.56	2.18
U-factor W	False	False	False	False	False	False	False	False	False	True	—	1.0	0.26	0.15	0.29	0.62	0.18	0.56	2.85
SHGC W	False	False	False	False	False	False	False	False	False	True	True	—	0.23	0.13	0.09	1.88	0.18	0.54	2.41
WWR N	False	False	False	False	False	False	False	False	False	False	False	False	—	0.82	7.79	3.35	0.19	0.66	3.73
U-factor N	False	False	False	False	False	False	False	False	False	False	False	False	True	—	2.1	0.34	0.12	0.57	2.97
SHGC N	False	False	False	False	False	False	False	False	False	False	False	False	True	True	—	3.34	0.13	0.94	4.57
Clsp	True	True	False	True	True	True	False	True	True	True	True	True	True	False	True	—	2.61	11.99	3.9
Htsp	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	True	—	0.5	1.39
In heat	False	False	True	False	True	True	False	True	True	True	True	True	True	True	True	True	True	—	17.34
Air change	True	True	True	True	True	True	False	True	True	True	True	True	True	True	True	True	True	True	—

Heating load

	Fin depth	WWR E	U-f E	SHGC E	Over-hang	WWR S	Thick S	U-f S	SHGC S	WWRW	U-f W	SHGC W	WWR N	U-f N	SHGC N	Clsp	Htsp	In heat	Air ch
Fin depth	—	0.2	0.17	0.27	0.31	0.23	0.09	0.25	0.33	0.25	0.34	0.23	0.14	0.1	0.16	0.2	0.27	0.32	0.24
WWR E	False	—	1.06	0.36	0.25	0.3	0.18	0.21	0.2	0.21	0.27	0.26	0.23	0.22	0.28	0.23	0.96	0.58	0.19
U-factor E	False	True	—	0.24	0.22	0.36	0.28	0.33	0.28	0.37	0.27	0.26	0.2	0.19	0.15	0.21	0.95	0.64	0.39
SHGC E	False	False	False	—	0.2	0.16	0.25	0.15	0.37	0.26	0.27	0.23	0.2	0.13	0.33	0.29	0.41	0.39	0.29
Overhang	False	False	False	False	—	0.56	0.22	0.25	0.56	0.25	0.31	0.25	0.21	0.16	0.16	0.16	0.38	0.37	0.31
WWR S	False	False	False	False	True	—	0.2	0.86	0.69	0.28	0.34	0.18	0.25	0.27	0.18	0.16	0.81	0.52	0.33
Thick S	False	False	False	False	False	False	—	0.12	0.18	0.16	0.15	0.17	0.21	0.22	0.32	0.18	0.2	0.27	0.2
U-factor S	False	False	False	False	False	True	False	—	0.36	0.25	0.18	0.19	0.13	0.13	0.18	0.18	0.87	0.53	0.38
SHGC S	False	False	False	False	True	True	False	False	—	0.38	0.2	0.15	0.19	0.15	0.14	0.14	0.59	0.64	0.24
WWR W	False	False	False	False	False	False	False	False	False	—	1.07	0.5	0.3	0.23	0.2	0.19	0.85	0.64	0.26
U-factor W	False	False	False	False	False	False	False	False	False	True	—	0.25	0.25	0.25	0.28	0.19	0.88	0.64	0.3
SHGC W	False	False	False	False	False	False	False	False	False	True	False	—	0.24	0.18	0.22	0.16	0.33	0.47	0.29
WWR N	False	False	False	False	False	False	False	False	False	False	False	False	—	1.01	0.72	0.15	1.0	0.57	0.28
U-factor N	False	False	False	False	False	False	False	False	False	False	False	False	True	—	0.32	0.19	0.88	0.47	0.29
SHGC N	False	False	False	False	False	False	False	False	False	False	False	False	True	False	—	0.23	0.63	0.51	0.37
Clsp	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	—	8.88	0.74	4.43
Htsp	False	True	True	False	False	True	False	True	True	True	True	False	True	True	True	True	—	1.84	2.84
In heat	False	True	True	False	False	True	False	True	True	True	True	False	True	False	True	True	True	—	0.4
Air change	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	True	True	False	—

Table 6-9: Normalized interaction matrices over AC peak

Cooling peak

	Fin depth	WWR E	U-f E	SHGC E	Over-hang	WWR S	Thick S	U-f S	SHGC S	WWRW	U-f W	SHGC W	WWR N	U-f N	SHGC N	Clsp	Htsp	In heat	Air ch
Fin depth	—	1.79	0.26	2.3	0.06	0.06	0.03	0.07	0.07	0.05	0.05	0.06	0.05	0.07	0.07	0.38	0.05	0.1	0.25
WWR E	True	—	1.13	5.1	0.1	0.13	0.02	0.07	0.13	0.13	0.06	0.07	0.21	0.07	0.2	1.95	0.07	0.26	0.26
U-factor E	False	True	—	1.43	0.09	0.08	0.03	0.05	0.09	0.06	0.03	0.09	0.12	0.06	0.14	1.21	0.06	0.19	0.29
SHGC E	True	True	True	—	0.09	0.1	0.02	0.08	0.12	0.09	0.07	0.09	0.12	0.08	0.13	0.66	0.08	0.17	0.41
Overhang	False	False	False	False	—	4.16	0.05	0.31	4.8	0.16	0.12	0.1	0.14	0.06	0.09	0.63	0.15	0.14	0.56
WWR S	False	False	False	False	True	—	0.18	0.91	5.76	0.17	0.08	0.11	0.09	0.07	0.12	1.9	0.08	0.25	0.4
Thick S	False	False	False	False	False	False	—	0.04	0.04	0.03	0.03	0.04	0.04	0.04	0.03	0.17	0.03	0.06	0.07
U-factor S	False	False	False	False	False	True	False	—	2.36	0.07	0.05	0.09	0.05	0.05	0.05	1.21	0.04	0.18	0.4
SHGC S	False	False	False	False	True	True	False	True	—	0.14	0.12	0.11	0.09	0.05	0.14	0.7	0.06	0.15	0.49
WWR W	False	False	False	False	False	False	False	False	False	—	1.14	5.18	0.17	0.08	0.16	1.77	0.06	0.29	0.2
U-factor W	False	False	False	False	False	False	False	False	False	True	—	1.34	0.12	0.06	0.15	1.27	0.05	0.17	0.24
SHGC W	False	False	False	False	False	False	False	False	False	True	True	—	0.12	0.07	0.11	0.45	0.06	0.15	0.3
WWR N	False	False	False	False	False	False	False	False	False	False	False	False	—	0.82	9.06	2.04	0.1	0.29	0.35
U-factor N	False	False	False	False	False	False	False	False	False	False	False	False	True	—	2.86	1.07	0.06	0.17	0.33
SHGC N	False	False	False	False	False	False	False	False	False	False	False	False	True	True	—	0.92	0.05	0.22	0.69
Clsp	False	True	True	True	True	True	False	True	True	True	True	False	True	True	True	—	0.25	2.15	7.03
Htsp	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	—	0.11	0.39
In heat	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	True	False	—	1.35
Air change	False	False	False	False	True	False	False	False	False	False	False	False	False	False	True	True	False	True	—

Heating peak

	Fin depth	WWR E	U-f E	SHGC E	Over-hang	WWR S	Thick S	U-f S	SHGC S	WWRW	U-f W	SHGC W	WWR N	U-f N	SHGC N	Clsp	Htsp	In heat	Air ch
Fin depth	—	0.25	0.2	0.18	0.35	0.25	0.13	0.31	0.28	0.25	0.34	0.27	0.18	0.16	0.15	0.18	0.26	0.45	0.27
WWR E	False	—	1.12	0.29	0.21	0.38	0.21	0.26	0.3	0.29	0.29	0.24	0.3	0.28	0.28	0.3	0.89	0.6	0.24
U-factor E	False	True	—	0.22	0.26	0.42	0.33	0.45	0.24	0.36	0.3	0.28	0.28	0.2	0.17	0.22	0.62	0.42	0.41
SHGC E	False	False	False	—	0.2	0.22	0.27	0.17	0.38	0.35	0.25	0.19	0.3	0.16	0.31	0.29	0.33	0.33	0.28
Overhang	False	False	False	False	—	0.56	0.25	0.33	0.37	0.29	0.32	0.3	0.2	0.23	0.22	0.21	0.39	0.46	0.36
WWR S	False	False	False	False	True	—	0.29	0.85	0.4	0.33	0.45	0.21	0.29	0.25	0.15	0.24	0.69	0.53	0.45
Thick S	False	False	False	False	False	False	—	0.16	0.23	0.24	0.21	0.24	0.29	0.23	0.29	0.16	0.27	0.22	0.24
U-factor S	False	False	False	False	False	True	False	—	0.33	0.29	0.22	0.25	0.15	0.15	0.19	0.21	0.89	0.64	0.37
SHGC S	False	False	False	False	False	False	False	False	—	0.35	0.2	0.17	0.18	0.16	0.18	0.14	0.59	0.58	0.29
WWR W	False	False	False	False	False	False	False	False	False	—	0.96	0.43	0.28	0.23	0.26	0.25	0.82	0.51	0.27
U-factor W	False	False	False	False	False	False	False	False	False	True	—	0.26	0.28	0.23	0.2	0.26	0.6	0.4	0.27
SHGC W	False	False	False	False	False	False	False	False	False	False	False	—	0.31	0.21	0.22	0.16	0.27	0.39	0.44
WWR N	False	False	False	False	False	False	False	False	False	False	False	False	—	1.08	0.61	0.23	0.81	0.53	0.46
U-factor N	False	False	False	False	False	False	False	False	False	False	False	False	True	—	0.35	0.22	0.7	0.42	0.49
SHGC N	False	False	False	False	False	False	False	False	False	False	False	False	True	False	—	0.23	0.43	0.37	0.45
Clsp	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	—	3.63	0.82	2.07
Htsp	False	True	True	False	False	True	False	True	True	True	True	False	True	True	False	True	—	1.56	1.43
In heat	False	True	False	False	False	True	False	True	True	True	False	False	True	False	False	True	True	—	0.61
Air change	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	True	True	True	—

Table 6-10: Normalized interaction matrices over operative temperature

	Fin depth	WWR E	U-f E	SHGC E	Over-hang	WWR S	Thick S	U-f S	SHGC S	WWRW	U-f W	SHGC W	WWR N	U-f N	SHGC N	Clsp	Htsp	In heat	Air ch
Fin depth	—-	0.29	0.14	0.44	0.16	0.2	0.2	0.22	0.23	0.19	0.24	0.17	0.14	0.14	0.16	0.19	0.2	0.19	0.21
WWR E	False	—-	1.12	0.75	0.24	0.23	0.22	0.27	0.27	0.24	0.28	0.19	0.25	0.2	0.18	0.49	0.32	0.28	0.32
U-factor E	False	True	—-	0.25	0.22	0.17	0.21	0.26	0.3	0.19	0.36	0.24	0.26	0.21	0.13	0.57	0.34	0.29	0.41
SHGC E	False	True	False	—-	0.28	0.24	0.21	0.24	0.25	0.23	0.22	0.22	0.23	0.18	0.2	0.37	0.27	0.24	0.23
Overhang	False	False	False	False	—-	1.09	0.2	0.26	1.23	0.24	0.19	0.19	0.25	0.17	0.25	0.42	0.3	0.37	0.41
WWR S	False	False	False	False	True	—-	0.17	0.99	1.25	0.23	0.31	0.31	0.2	0.23	0.26	0.45	0.33	0.46	0.41
Thick S	False	False	False	False	False	False	—-	0.21	0.2	0.2	0.19	0.28	0.23	0.2	0.18	0.2	0.22	0.14	0.14
U-factor S	False	False	False	False	False	True	False	—-	0.68	0.26	0.31	0.17	0.19	0.21	0.21	0.63	0.34	0.38	0.4
SHGC S	False	False	False	False	True	True	False	True	—-	0.28	0.25	0.25	0.15	0.16	0.18	0.52	0.33	0.4	0.28
WWR W	False	False	False	False	False	False	False	False	False	—-	0.96	0.92	0.2	0.18	0.29	0.58	0.28	0.3	0.37
U-factor W	False	False	False	False	False	False	False	False	False	True	—-	0.37	0.19	0.23	0.24	0.55	0.35	0.24	0.45
SHGC W	False	False	False	False	False	False	False	False	False	True	False	—-	0.26	0.16	0.23	0.36	0.29	0.24	0.19
WWR N	False	False	False	False	False	False	False	False	False	False	False	False	—-	0.93	1.27	0.5	0.39	0.42	0.66
U-factor N	False	False	False	False	False	False	False	False	False	False	False	False	True	—-	0.51	0.58	0.35	0.31	0.6
SHGC N	False	False	False	False	False	False	False	False	False	False	False	False	True	True	—-	0.58	0.36	0.33	0.41
Clsp	False	False	True	False	False	False	False	True	True	True	True	False	True	True	True	—-	0.57	1.8	3.08
Htsp	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	True	—-	1.01	2.74
In heat	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	True	True	—-	0.98
Air change	False	False	False	False	False	False	False	False	False	False	False	False	True	True	False	True	True	True	—-

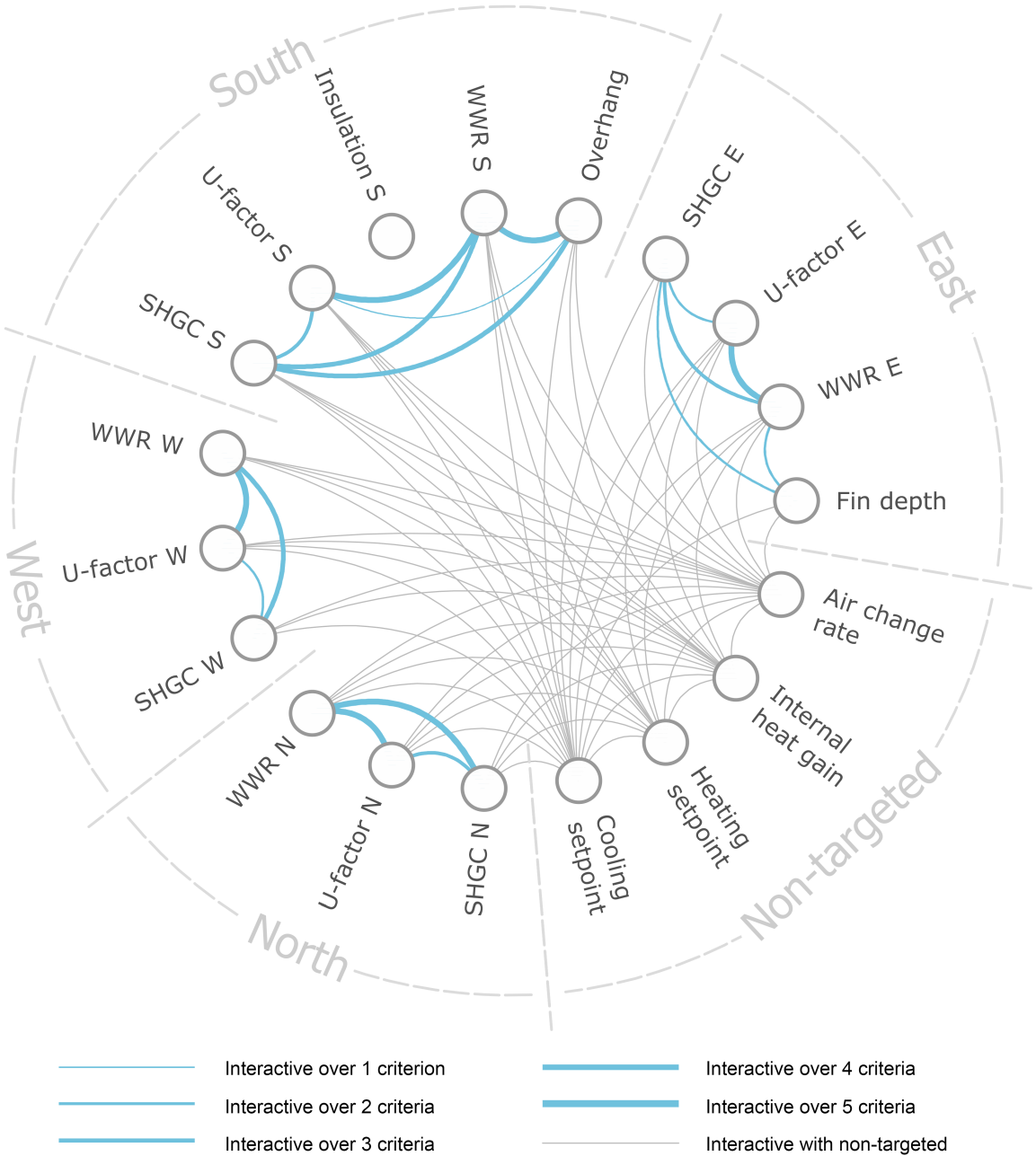


Figure 6-15: Interactions between parameters

The next step is to group the parameters. Principally, to make sure that the size of each group is minimized, there two rules to group the parameters based on the interaction matrices,

1 If a parameter is interactive with all parameters in a group, that parameter should join the group.

2 All parameters in one group should be interactive pairwise.

8 groups were made based on this 2 principles (table 6-11).

Table 6-11: Minimized groups

Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7	Group 8
Fin depth	WWR E	Overhang scale	WWR S	WWR W	WWR W	WWR N	Insulation S
WWR E	U-factor E	WWR S	U-factor S	U-factor W	U-factor W	U-factor N	
SHGC E	ClSp	U-factor S	SHGC S	SHGC W	ClSp	SHGC N	
ClSp	HtSp	SHGC S	ClSp	ClSp	HtSp	ClSp	
ACR	In HG	ClSp	HtSp	In HG	In HG	HtSp	
	ACR	In HG	In HG	ACR	ACR	In HG	
		ACR	ACR			ACR	

However, too many groups would cause inconvenience in latter phase. As meta-modeling would be carried out, the dimension of a group (the number of parameters in a group) is not critical. We adjusted the grouping into 5 groups (table 6-12). Besides 1 group that only included insulation on south, each of other groups included the parameters on one orientation and the 4 non-targeted parameters. Admittedly, it was quite a boring grouping solution.

Table 6-12: Adjusted groups

Group 1	Group 2	Group 3	Group 4	Group 5
Fin depth	Overhang scale	WWR W	WWR N	Insulation S
WWR E	WWR S	U-factor W	U-factor N	
U-factor E	U-factor S	SHGC W	SHGC N	
SHGC E	SHGC S	Cooling setpoint	Cooling setpoint	
Cooling setpoint	Cooling setpoint	Heating setpoint	Heating setpoint	
Heating setpoint	Heating setpoint	Internal heat gain	Internal heat gain	
Internal heat gain	Internal heat gain	Air change rate	Air change rate	
Air change rate	Air change rate			

6.5 Meta-modeling

In this case study, we imagined that the design team decided to do studies on the east and south facades, as they were related to the appearance of this building. For the west and north facade, the design team decided to use the base values from the codes for parameters.

In this step, we trained GPR models separately for the 1^{st} and 2^{nd} parameter groups, which respectively included the parameters of east and south facades, using STMN (section 5.6.2) method. Then combined the predictions of the same criterion form two GPR models using PCP (section 4.6).

6.5.1 Adjustment of the ranges of non-targeted parameters

From the sensitivity analysis (figure 6-12), it can be found that the sensitivities of non-targeted parameters overwhelmed that of the targeted parameters. As explained in chapter 4, the sensitivity of a parameter also reflects the uncertainty in the results that is caused by this parameter. In chapter 5, we have proved that high uncertainties caused by non-targeted parameters could lead to the failures of training meta-models. For these reasons, we adjusted the ranges of non-targeted parameters to depress the uncertainties that they caused in the values of criteria.

We adjusted the ranges into normal distributions. However, without any limitations, there are chances that the randomly generated heating setpoint is higher than cooling setpoint. What's more, there are also chances that extremely high or low values were generated. As the variance of a group of data is sensitive to the extreme values, the uncertainties in the simulation results could be too large. So, we limited the range of a normal distribution, which is called truncated normal distribution. Equation 6.5 shows the probability density function (PDF) of $TN(\mu, \sigma, b, t)$.

$$PDF_{tn}(x) = \frac{1}{\sigma\sqrt{2\pi}} exp(-\frac{(x-\mu)^2}{2\sigma^2}) + (2 + erf\frac{b-\mu}{\sigma\sqrt{2}} + erf\frac{\mu-t}{\sigma\sqrt{2}})/2(t-b) \tag{6.5}$$

Where:

μ = Mean value of the original normal distribution
σ = Standard deviation (scale coef) of the original normal distribution
b = Bottom value of the limitation range
t = Top value of the limitation range

As illustrated in figure 6-16, a truncated normal distribution can be regarded as a normal distribution, with the parts beyond the limitation range cut, moved up a bit. The area of the added 'basement' is just equal to that of the cut 'wings'. By the way, though the PDF of the truncated normal distribution looks complicated, there is no necessary to use mathematically difficult method to generate random numbers. Using the keyword

'while' in your codes can make it easy, without any mathematical mistakes. Though there is actually a chance that the program freezes, which infinitely approaches 0.

Here is a *C* flavored example,

```
float clsp = 0;
float htsp = 0;

while(clsp < b_c or clsp > t_c)
{
    clsp = random number ~ Normal(μ_c, σ_c);
}

while(htsp < b_h or htsp > t_h)
{
    htsp = random number ~ Normal(μ_h, σ_h);
}

return clsp, htsp;
```

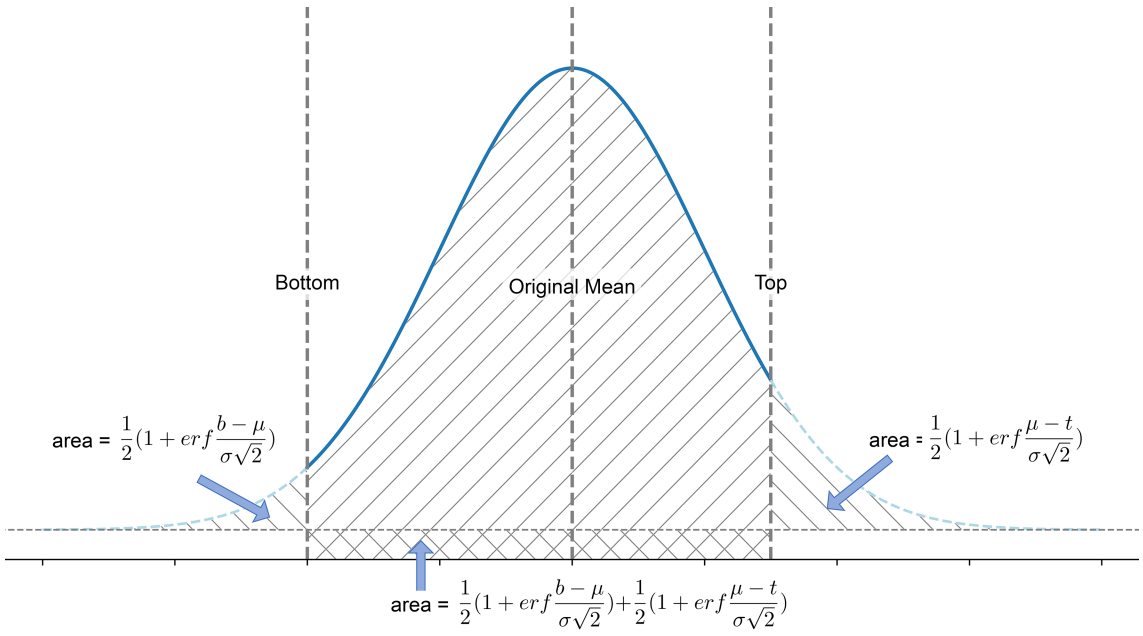


Figure 6-16: Truncated normal distribution

Table 6-13 shows the adjustment of the ranges of non-targeted parameters.

Table 6-13: Simulation configurations

Non-targeted parameter	Unit	Original range	Adjusted range
Cooling set point	$^{\circ}C$	[23, 30]	$LN(24, 1, 23, 25)$
Heating set point	$^{\circ}C$	[16, 23]	$LN(20, 1, 19, 21)$
Internal heat gain	W/m^2	[3, 15]	$LN(9, 2, 7, 11)$
Air change rate	$1/hour$	[0, 6]	$LN(3, 1, 2, 4)$

6.5.2 Selection of sampling methods

Though we had adjusted the ranges of non-targeted parameters, the sensitivities of them were still high. '**S**ingle sample of **T**argeted parameters paired with **M**ultiple samples of **N**on-targeted parameters' (STMN, section 5.6.2) method was used to sampling the parameters and training GPR models.

For each group, we randomly generated 500 samples of targeted parameters using LHS, 400 for training and 100 for testing. Then, for each sample, 20 samples of non-targeted samples were generated. So that, 10000 cases were generated for each group. We called this collection of samples *East/South STMN sampled*.

For comparison, we also generated 2000 cases for each group using LHS commonly, to certify the effectiveness STMN method. All these cases are for training. We called this collection of samples *East/South commonly sampled*. For testing, we used the STMN sampled cases.

We also combined 1^{st} and 2^{nd} groups, then generated 500 samples (400 training, 100 testing) including the targeted parameters of both east and south facade, to certify the **P**rediction **C**ombination **P**olynomial (PCP, section 4.6) method. Each sample were extended in to 20 cases using STMN. We called these samples *East+South STMN sampled*.

Table 6-14 shows the sampling methods.

All these cases were converted to IDF files and simulated using *EnergyPlus* with TMY climate data. R^2_U analysis was also did to show the relative uncertainties caused by non-targeted parameters in cooling and heating load. As illustrated in figure 6-17, the relative uncertainties were still quite huge, even with ranges of non-targeted parameters shrunk, which also indicated the necessity of STMN method. We firstly trained GPR models of Cooling load and Heating load for each collection of samples, using RBF kernel functions, as a trial. Then, we combined predictions from *East STMN trained* and *South STMN trained* models using PCP method (equation 6.10). As there were no common targeted parameter shared by Group 1 and 2, the correction term between $f(T^{g1}, C^{other})$ and $f(T^{g2}, C^{other})$ is $f(C^{all})$ (equation 6.6). The mean value of $Pd_e(0.5)$ and $Pd_s(0.5)$ was used to represent $f(C^{all})$ (equation 6.9). Theatrically, it is better to take the covariance between Pd_e and Pd_s into consideration, as the non-targeted parameters are shared by Group 1 and 2. In this case study, as we aimed at the overall tendency, it was not a big problem to ignore the covariance.

Table 6-14: Sampling methods

Group 1				Group 2				Group 3,4	Non-targeted			
Fin depth	**WWR E**	**U-f E**	**SHGC E**	**Overhang**	**WWR S**	**U-f S**	**SHGC S**	**Other paras**	**Clsp**	**Htsp**	**In heat**	**ACR**
East STMN sampled ($i \sim [0,500), j \sim [0,20)$)												
$rand_i$	$rand_i$	$rand_i$	$rand_i$	0.5	0.5	0.5	0.5	0.5	$rand_{i,0}$ $rand_{i,1}$ ⋮ $rand_{i,j}$	$rand_{i,0}$ $rand_{i,1}$ ⋮ $rand_{i,j}$	$rand_{i,0}$ $rand_{i,1}$ ⋮ $rand_{i,j}$	$rand_{i,0}$ $rand_{i,1}$ ⋮ $rand_{i,j}$
South STMN sampled												
0.5	0.5	0.5	0.5	$rand_i$	$rand_i$	$rand_i$	$rand_i$	0.5	$rand_{i,0}$ $rand_{i,1}$ ⋮ $rand_{i,j}$	$rand_{i,0}$ $rand_{i,1}$ ⋮ $rand_{i,j}$	$rand_{i,0}$ $rand_{i,1}$ ⋮ $rand_{i,j}$	$rand_{i,0}$ $rand_{i,1}$ ⋮ $rand_{i,j}$
East Commonly sampled												
$rand_i$	$rand_i$	$rand_i$	$rand_i$	0.5	0.5	0.5	0.5	0.5	$rand_i$	$rand_i$	$rand_i$	$rand_i$
South Commonly sampled												
0.5	0.5	0.5	0.5	$rand_i$	$rand_i$	$rand_i$	$rand_i$	0.5	$rand_i$	$rand_i$	$rand_i$	$rand_i$
East+South STMN sampled												
$rand_i$	$rand_i$	$rand_i$	$rand_i$	$rand_i$	$rand_i$	$rand_i$	$rand_i$	0.5	$rand_{i,0}$ $rand_{i,1}$ ⋮ $rand_{i,j}$	$rand_{i,0}$ $rand_{i,1}$ ⋮ $rand_{i,j}$	$rand_{i,0}$ $rand_{i,1}$ ⋮ $rand_{i,j}$	$rand_{i,0}$ $rand_{i,1}$ ⋮ $rand_{i,j}$

$$\because f(T_i^{g1}, T_i^{g2}, 0.5^{other}) = f(T_i^{g1}, 0.5^{other}) - f(0.5^{all}) + f(T_i^{g2}, 0.5^{other}) \tag{6.6}$$

$$Pd_e(T_i^{g1}) \rightarrow f(T^{g1}, C^{other}) \tag{6.7}$$

$$Pd_s(T_i^{g2}) \rightarrow f(T^{g2}, C^{other}) \tag{6.8}$$

$$(Pd_e(0.5) + Pd_s(0.5))/2 \rightarrow f(C^{all}) \tag{6.9}$$

$$\therefore Pd_{pcp}(T_i^{g1}, T_i^{g2}) = Pd_e(T_i^{g1}) - (Pd_e(0.5) + Pd_s(0.5))/2 + Pd_s(T_i^{g2}) \tag{6.10}$$

Where:

$f(T_i^{g1}, T_i^{g2}, 0.5^{other})$	:	Real value of a criterion when targeted parameters in Group 1 and 2 are T_i^{g1}, T_i^{g2} and other targeted parameters are fixed to 0.5
$f(T_i^{g1}, 0.5^{other})$	:	Real value of a criterion when targeted parameters in Group 1 are T_i^{g1} and other targeted parameters are fixed to 0.5
$f(T_i^{g2}, 0.5^{other})$	:	Real value of a criterion when targeted parameters in Group 2 are T_i^{g2} and other targeted parameters are fixed to 0.5
$f(0.5^{all})$	:	Real value of a criterion when all targeted parameters re fixed to 0.5
Pd_{pcp}	:	Combined prediction using PCP method
Pd_e	:	Predicted value of a criterion by GPR East STMN trained
Pd_s	:	Predicted value of a criterion by GPR South STMN trained
T^{g1}	:	Targeted parameters in Group 1 (East)
T^{g2}	:	Targeted parameters in Group 2 (South)
0.5	:	Parameters are fixed to the mid values of their ranges

We used the testing samples from *East/South STMN sampled* collection to test the *East/South STMN trained* and *East/South commonly trained* GPR models. The targeted parameters of each sample were fed into the GPR models. The predicted mean value were compared to the mean value of the 20 simulated results of each case (figure 6-18 and 6-20), as well as the predicted and simulated variances (figure 6-19 and 6-21). R_E^2 was used to evaluate the error in each GPR model.

Likewise, samples from *East+South STMN sampled* collection were used to test *East+South PCP combined* and *East+South STMN trained* models.

From the figures 6-18 to 6-21, we can see that the *STMN trained* models outperforms those commonly sampled and trained. The *PCP combined* models work well as expectation. The *East+South STMN trained* models also works perfectly, which was a little surprising. STMN trained and PCP combined models were able to correctly predicted the mean value of each sample set.

The predictions of variance were a little poor, but the overall tendency was not bad. Additionally, the mean values of all predicted variances are almost the same as those of simulated variances. *East+South STMN trained* models performed a little better than *East+South PCP combined* ones. We thought that it was very hard to say the errors were in predictions or simulated values. Each simulated variance were calculated

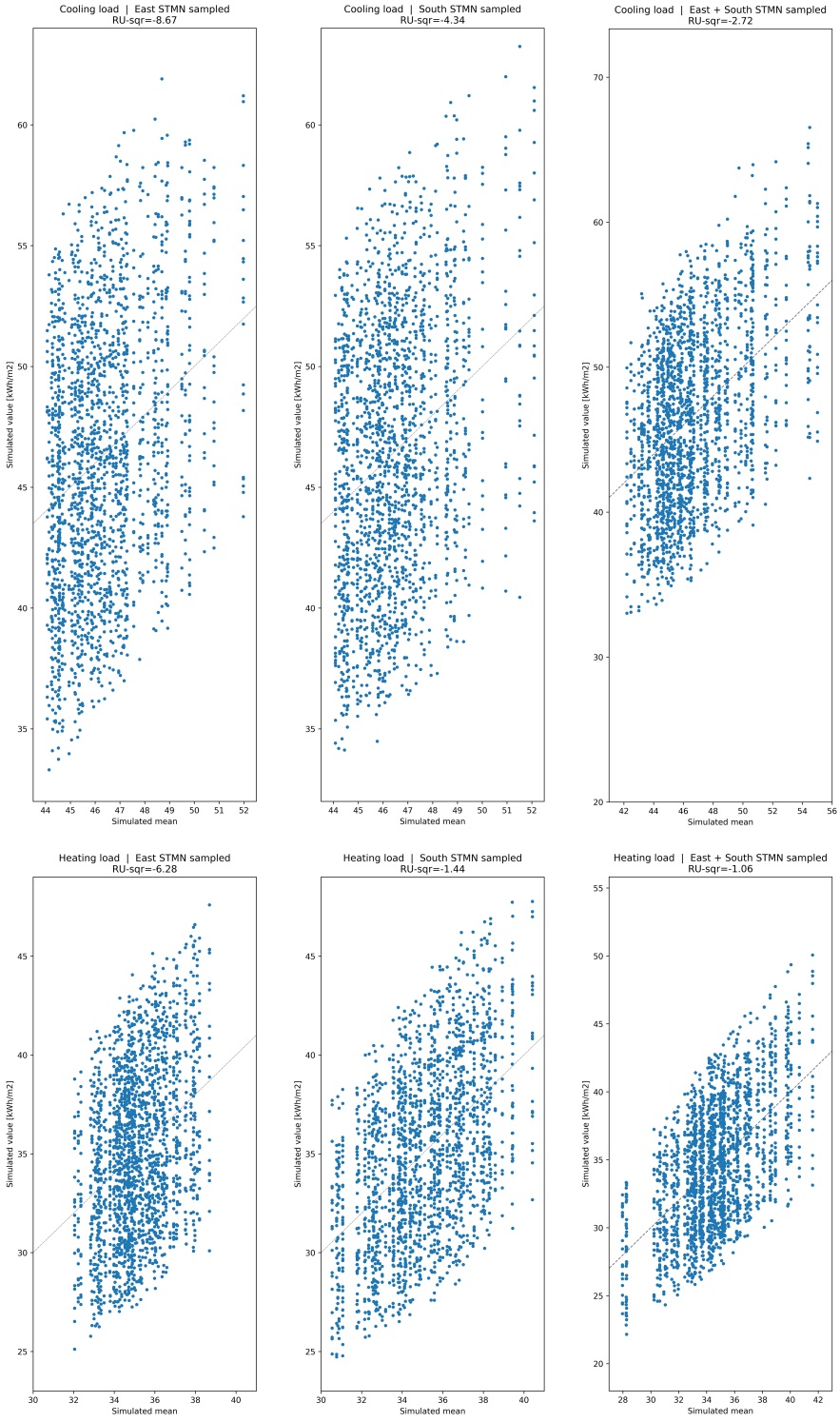


Figure 6-17: Relative uncertainties in Cooling and Heating load

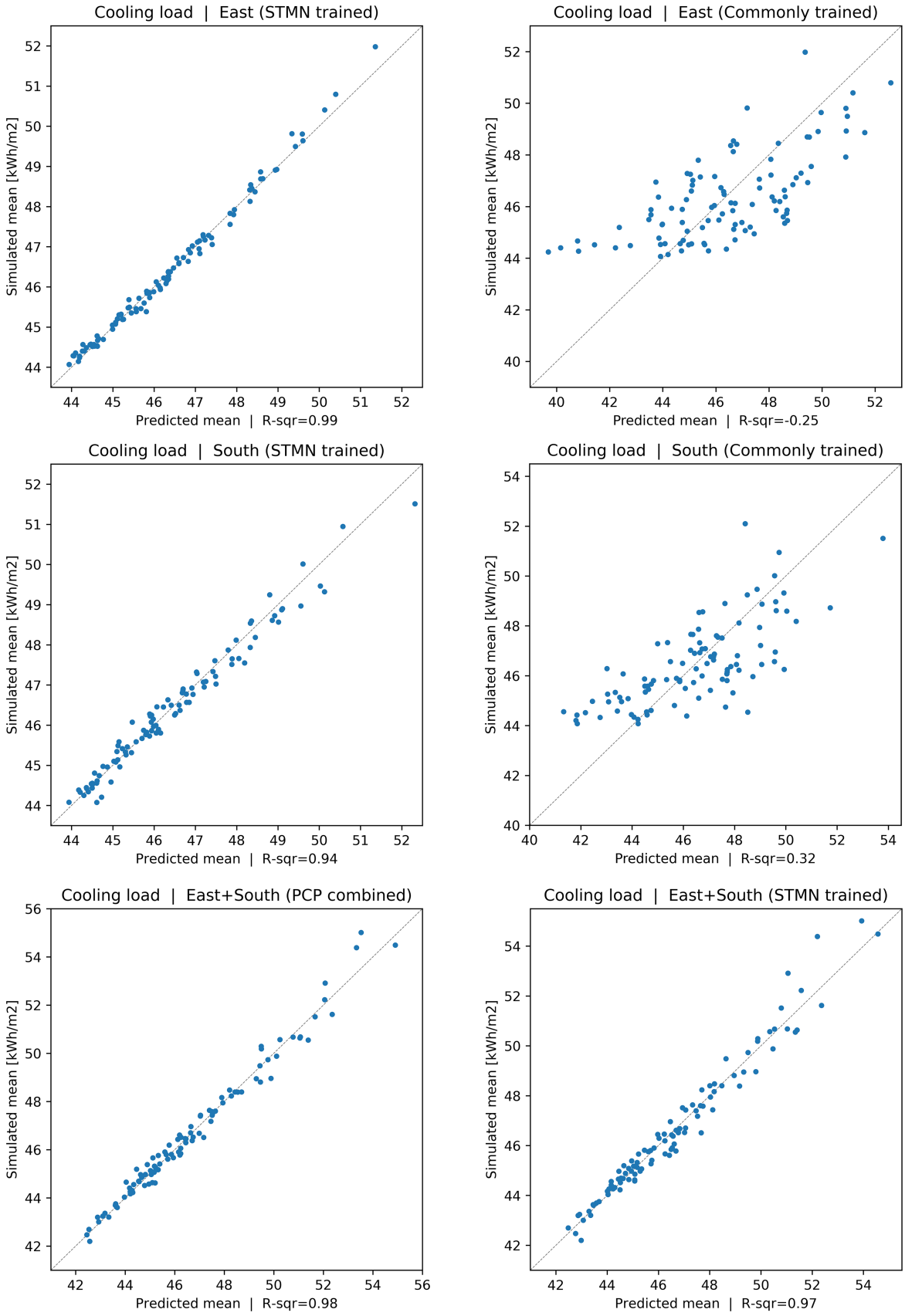


Figure 6-18: Testing of GPR models of Cooling load (Mean, μ)

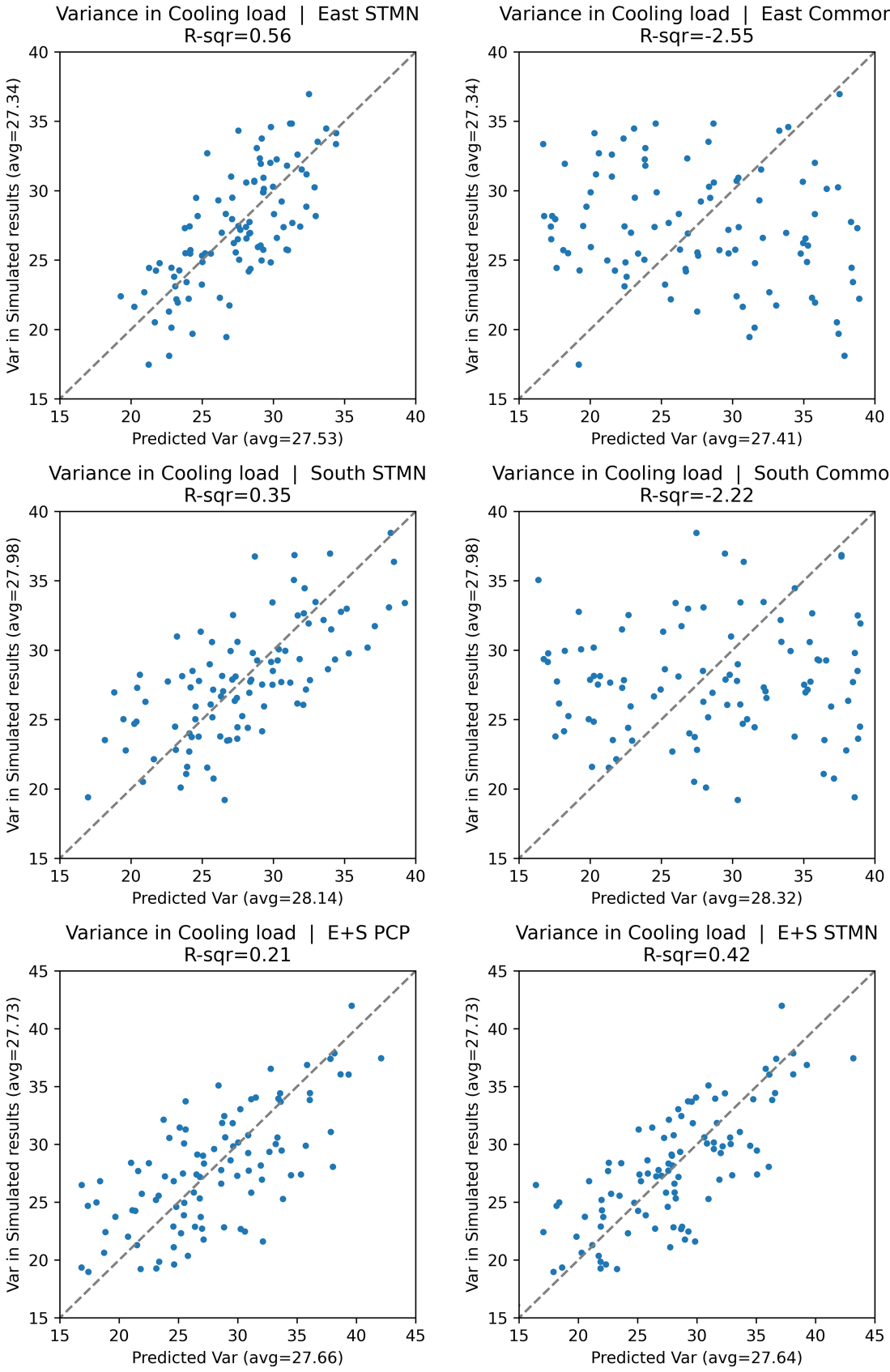


Figure 6-19: Testing of GPR models of Cooling load (Variance, σ^2)

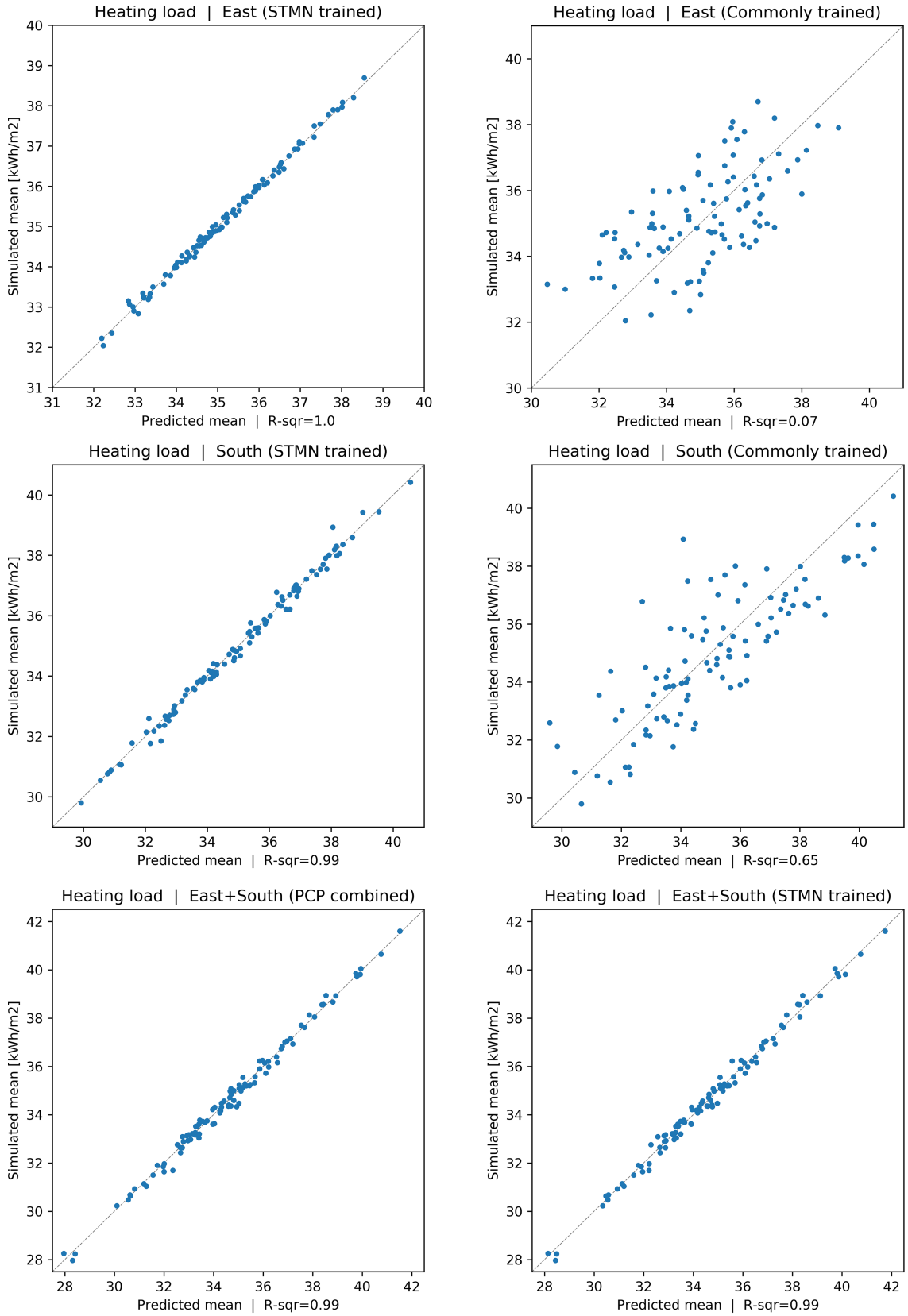


Figure 6-20: Testing of GPR models of Heating load (Mean, μ)

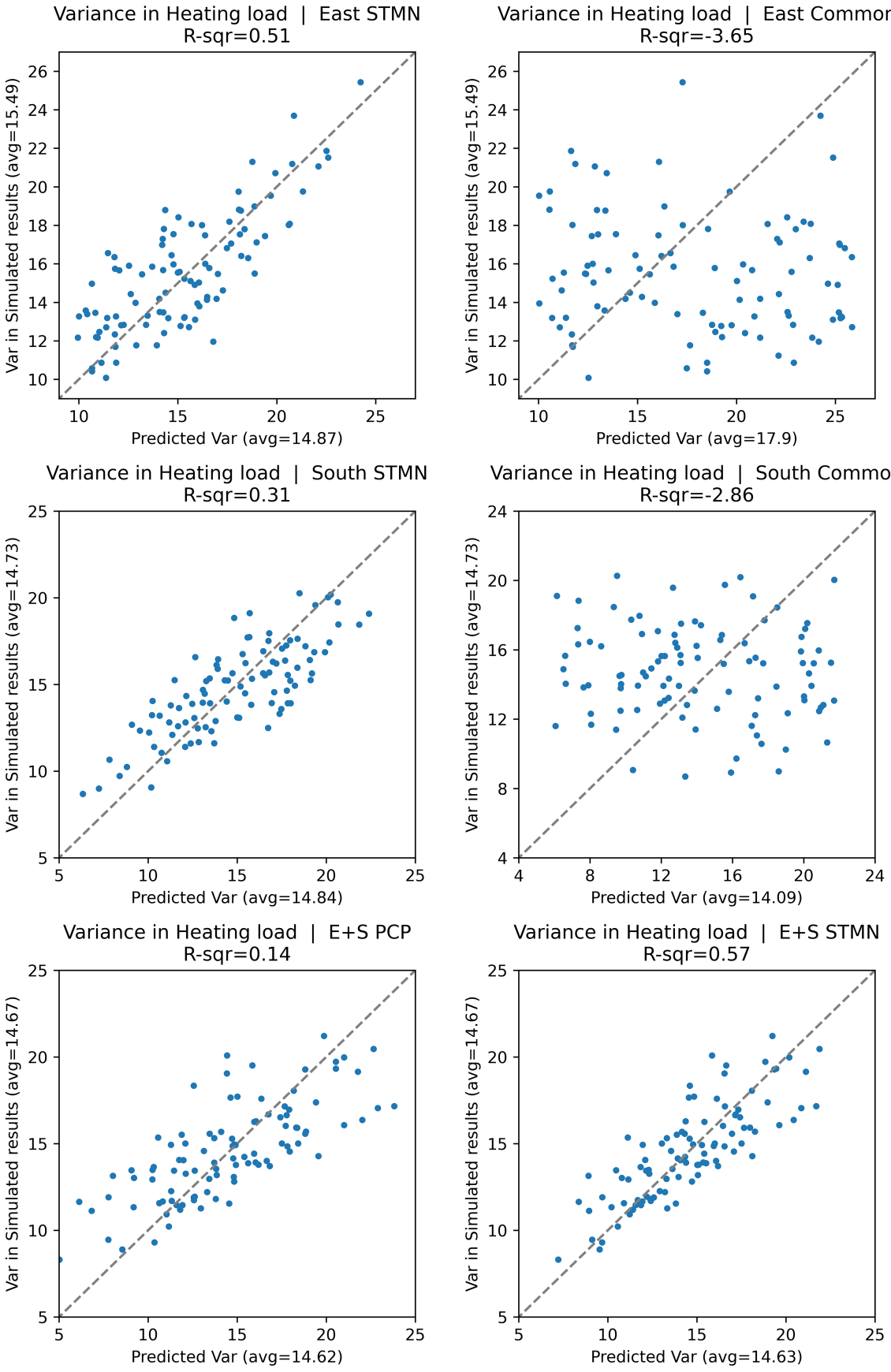


Figure 6-21: Testing of GPR models of Heating load (Variance, σ^2)

based on 20 samples. As the value of variance is very sensitive to those few extreme values in the data set, 20 sample were not enough to give out a reliable variance.

Finally, we decided to continue with the *East+South STMN* training method. The predictions of variances were a little better. Meanwhile, the time costing to train these models was as half as that of getting *PCP combined* models. The performance good enough. On the other hand, there was no covariance problem.

6.5.3 Training GPR models for all criteria and all climate

We executed simulations for all cases in *East+South STMN sampled* collection, using *EnergyPlus*, with 2020, 2050, 2080 climate data. *STMN trained* GPR models were made for all criteria and all climate.

'True R^2 error' method was used to test the trained GPR models. We also compared the simulated and predicted variances of cooling and heating loads of each sample of targeted-parameters. The mean values of simulated and predicted variances were very close. So that we could say that the predicted distributions of cooling and heating loads well reflected the uncertainties caused by non-targeted parameters. The testing results of all GPR models were gathered in table 6-15.

For other criteria, as they were not our main targets in this case study, we only used the mean values of each sample to train the GPR models. Figure 6-22 shows the testing results for TMY.

Table 6-15: Testing results

	Cooling load				Heating load			
	$R^2_E(\mu)$	$R^2_E(\sigma^2)$	Sim var	Pre var	$R^2_E(\mu)$	$R^2_E(\sigma^2)$	Sim var	Pre var
TMY	0.97	0.42	27.73	28.61	0.99	0.57	14.67	14.92
2020	0.98	0.53	28.34	27.98	0.99	0.52	14.77	14.82
2050	0.99	0.39	28.56	28.28	0.98	0.48	15.28	14.75
2080	0.96	0.48	29.32	29.93	0.99	0.55	14.29	14.81
	Cl pk	Ht pk	Comf	Cold	Heat	OT mean	OT cold	OT hot
TMY	0.97	1.0	0.97	0.99	0.95	1.0	1.0	0.93
2020	0.96	1.0	0.96	0.99	0.95	1.0	1.0	0.94
2050	0.97	1.0	0.97	0.99	0.95	0.99	1.0	0.93
2080	0.99	0.99	0.97	1.0	0.95	1.0	1.0	0.96

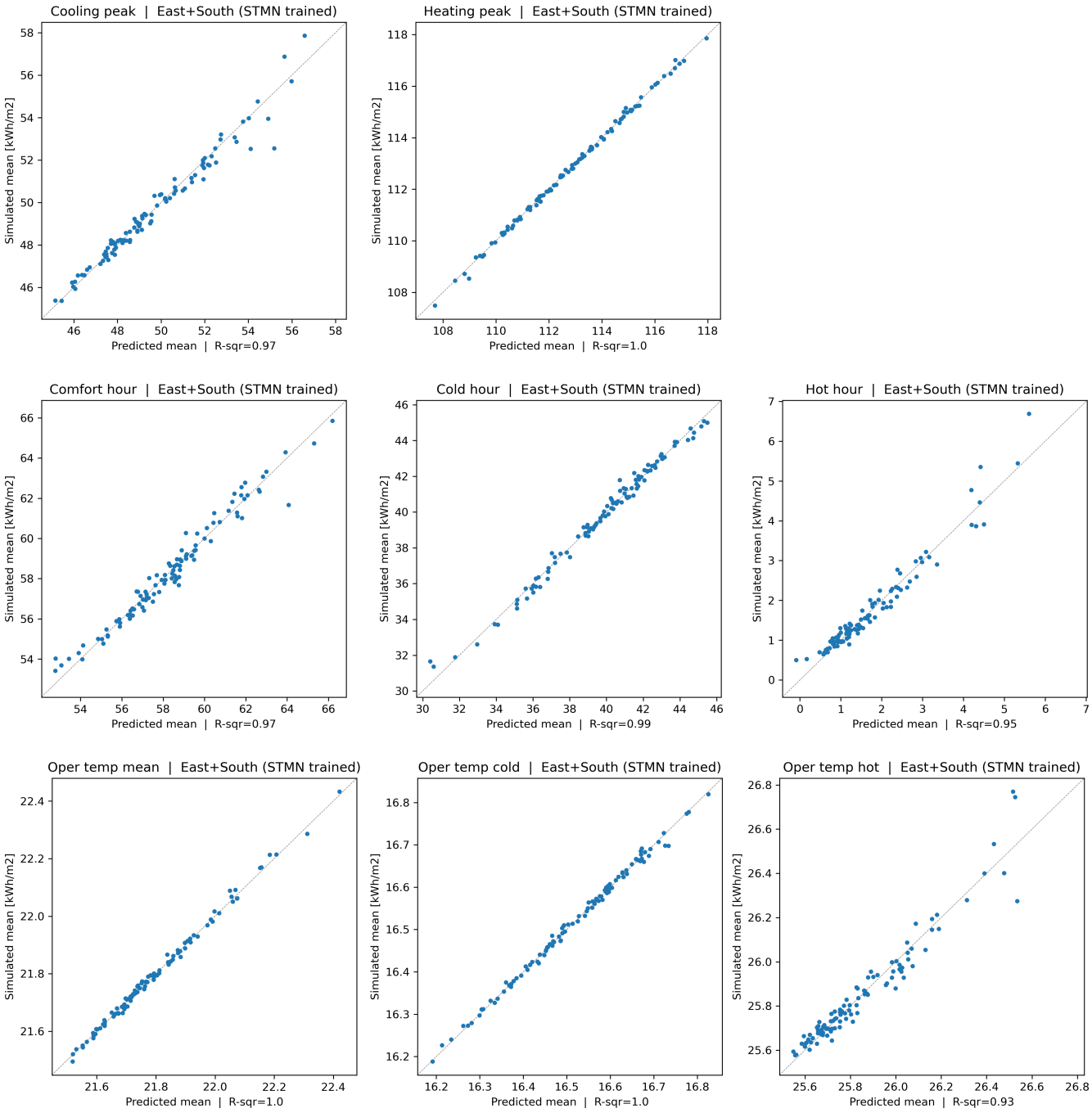


Figure 6-22: Testing of GPR models of other criteria (TMY)

6.5.4 Combining Cooling and heating load to AC electric consumption

As we planed to visualize the annual electric energy consumption by AC, one more prediction model was needed for each climate year. The distribution of the AC electric consumption can be inferred from that of cooling and heating loads. However, there are also situations that the covariance of cooling and heating loads of the same sample of targeted parameters should be considered.

Firstly, we executed the Pearson correlation analysis to see whether the distribution of cooling and heating loads of the same sample were independent. All samples in the *East+South STMN sampled* collection were analyzed. Each sample included 20 sub-samples and 20 pairs of simulated cooling and heating load. The Pearson correlation coefficients were calculated. The average PCC is 0.18, which showed that the cooling and heating loads were independent.

The correlations between cooling and heating of 20 samples are illustrated in figure 7-4

For that reason, we can use the convolution of the predicted cooling and heating distribution as that of the AC electric consumption (equation 6.13). The COP of cooling and heating is respectively 3.2 and 3.5.

$$Cl_i \sim N(\mu_i^c, \sigma_i^c) \tag{6.11}$$

$$Hl_i \sim N(\mu_i^h, \sigma_i^h) \tag{6.12}$$

$$AC_i \sim N\left(\frac{\mu_i^c}{3.2} + \frac{\mu_i^h}{3.5}, \sqrt{\left(\frac{\sigma_i^c}{3.2}\right)^2 + \left(\frac{\sigma_i^h}{3.5}\right)^2}\right) \tag{6.13}$$

Where:

Cl_i	:	Predicted normal distribution of cooling load
μ_i^c	:	Mean value of predicted distribution of cooling load
σ_i^c	:	Standard deviation value of predicted distribution of cooling load
Hl_i	:	Predicted normal distribution of heating load
μ_i^h	:	Mean value of predicted distribution of heating load
σ_i^h	:	Standard deviation value of predicted distribution of heating load
AC_i	:	Normal distribution of AC electric consumption

6.6 Data visualization

6.6.1 Making predictions and database

200,000 cases were sampled using LHS, and then remapped. Each case included targeted parameters on east and south, in the 1^{st} and 2^{nd} groups. All these cases were fed into the trained GPR models. For each criterion, 200,000 probability distributions were predicted.

There were 2 kinds probability distribution in this case study. The first one was the predicted distribution of each criterion of each case, which was called single-case distribution in this research. The other one is the distribution of all the 200,000 single-case distributions, which was called overall distribution. It was almost impossible to visualize the all the 200,000 single-case distributions in 2D diagrams. As an alternative solution, we extracted the mean value μ and the boundaries of the 68.2% confidence interval, $\mu - \sigma$ and $\mu + \sigma$, to represent a single-case distribution (Figure 6-23). Then the distributions of all the μ, $\mu - \sigma$ and $\mu + \sigma$ would be displayed in the dashboard.

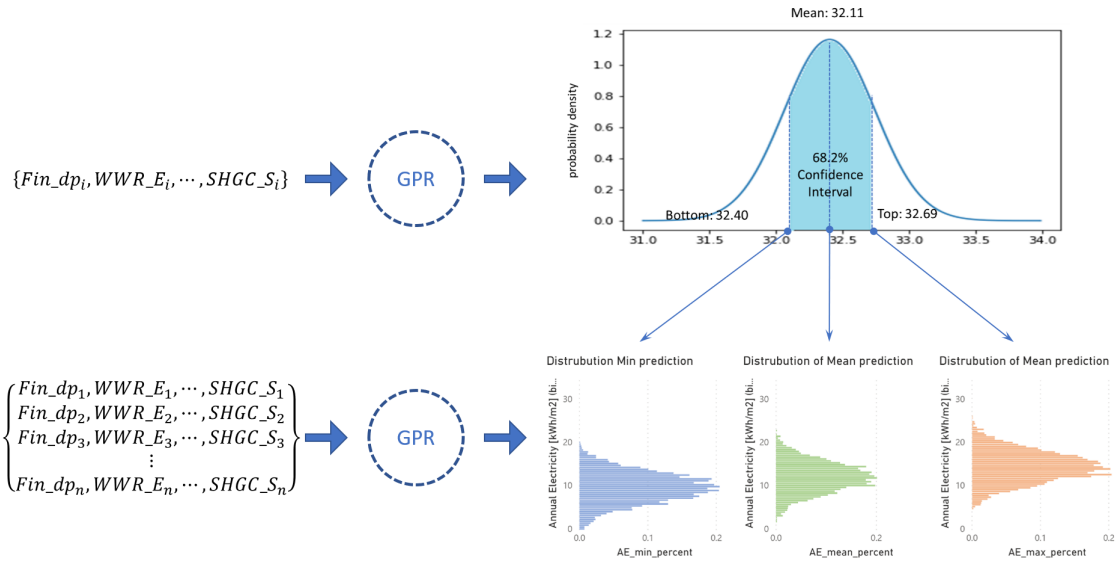


Figure 6-23: Parametric modeling flowchart

The remapped 200,000 cases were stored in a table, which is called sample table in this case study. The key of the table is the [Sample ID] column. The μ values of all criteria, of all cases and all years, were stored in a table called criterion table. The criteria tables also included [Sample ID] and [EPW year] columns, but no keys. The criterion tables including $\mu - \sigma$ and $\mu + \sigma$ values of heating load and cooling load were also made respectively. Another table called climate year table was made, including only 4 rows, used to filter the data of a certain year.

The data of the 250 cases of sensitivity analysis that we did in the last of section 6.3, was also included in

this database. The values of the targeted parameters are joined in the sample table. The sensitivities of all parameters over one criterion were stored in one table, which was called sensitivity table.

The sample table was related to the criterion/sensitivity tables by [Sample ID] single-directionally, from sample table to criterion/sensitivity tables. The climate year table was also related to the criterion/sensitivity tables by [EPW year] single-directionally, from climate year table to criterion/sensitivity tables. We also included the climate data from *Meteonorm* into the database, to make a dashboard that illustrated the impact of climate change on architecture design. Figure 6-24 illustrates the relations in this database. Admittedly, as a SQL database, the relations seems a little strange. That was because the way that Power BI filtered data and displayed the column names.

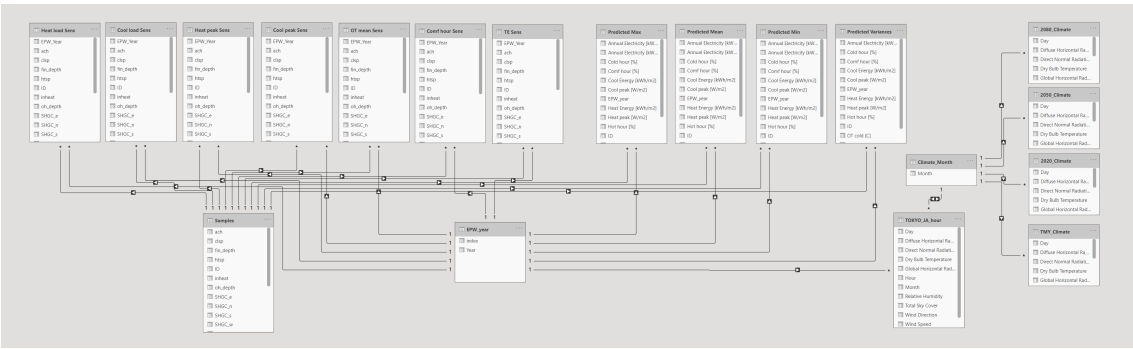


Figure 6-24: SQL database relations

6.6.2 Making dashboards

Micro soft Power BI was used to visualize the data and make design dashboards. For demonstration, we made three dashboards. Designers can adjust the parameters by moving the sliders of filters or check the check boxes. Power BI will filter the cases that whose parameters are within the user-set ranges and picked up the corresponding criterion or sensitivity values according to the [Sample ID] and [EPW year] columns. So that the variation of criteria and sensitivities will be reflected in real-time.

In the energy density and thermal comfort dashboard (figure 6-25), designers could move the sliders on the left side to adjust the ranges of parameters. The EPW years could also be selected by checking the checkbox on the upper left corner. Power BI would then filter out the corresponding cases. For each case, 3 indices of energy density, μ, $\mu - \sigma$ and $\mu + \sigma$, were displayed, named ‘Min prediction’, ‘Mean prediction’ and ‘Max prediction’ in this dashboard. The average values of these 3 indices of all the filtered cases are displayed on the upper right in the form of a speedometer. Furthermore, we also displayed the distributions of the 3 indices of all the filtered cases, which can help designers better understand the possibilities and risks of their decisions. A pie diagram of comfort is located on the bottom of this page, indicating the proportions of comfort, cold and hot hours. The minimum, average and maxim values of cooling/heating peaks are also

displayed on the lower right corner. As explained in section 6.1.2, these indices are not for HVAC sizing but just references.

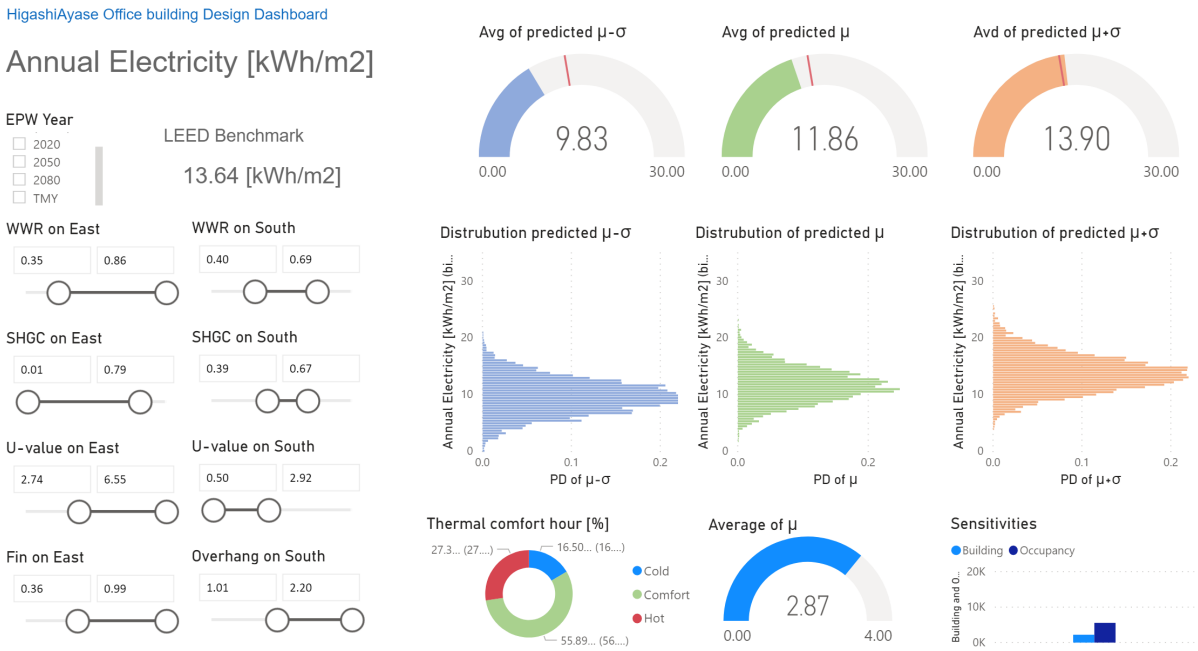


Figure 6-25: Energy density and thermal comfort dashboard

The sensitivity of each parameter will change with the variation of the value of that parameter, as well as that of other parameters due to the interactions. Figure 6-26 illustrates the dashboard that displays the sensitivities of all parameter overs each criterion with a pie diagram. The displayed sensitivity pies surely also vary with the sliders on the left adjusted. With this dashboard, the designers can be kept informed that which parameter is important, which could be an important hint for the designers. At this stage, duo to the limitation of Power BI, the sliders in this page were not linked to those in the energy density dashboard, otherwise it would be perfect.

As a building would last for decades, climate change should also be considered in design stage. A climate change dashboard (figure 6-27) was also made. We thought that showing the influence from climate change on the sensitivities of parameters, rather than that on the energy consumption or thermal comfort, can better help designers understand the impact on the buildings, as this research mainly focused on the design stage. The variation of air temperature and solar radiation, which cause most impact on the buildings in the climate change phenomenon, was displayed on the upper half of this page. The lower part shows the variation of the sensitivity of each parameter over the annual energy density criterion. Users could find that the shading-related parameters get more important but the insulation-related parameters get less important, so that the design team can pay more attention to the shading.

Figure 6-26: Sensitivity dashboard

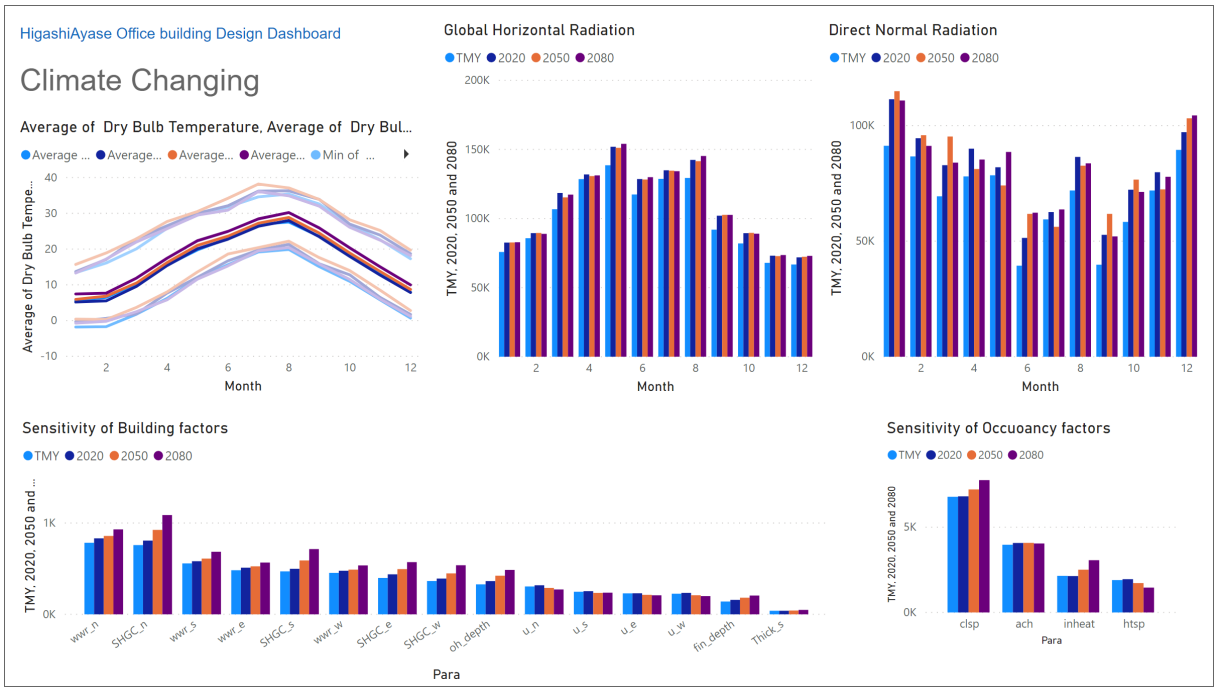


Figure 6-27: Sensitivity dashboard

6.7 Adding prediction

It was found in this case study that, when designers trying to explore the design space using the energy density dashboard, if they limited the parameters to quite small ranges, the displayed distributions became unsmooth and unstable, as illustrated in figure 6-28.

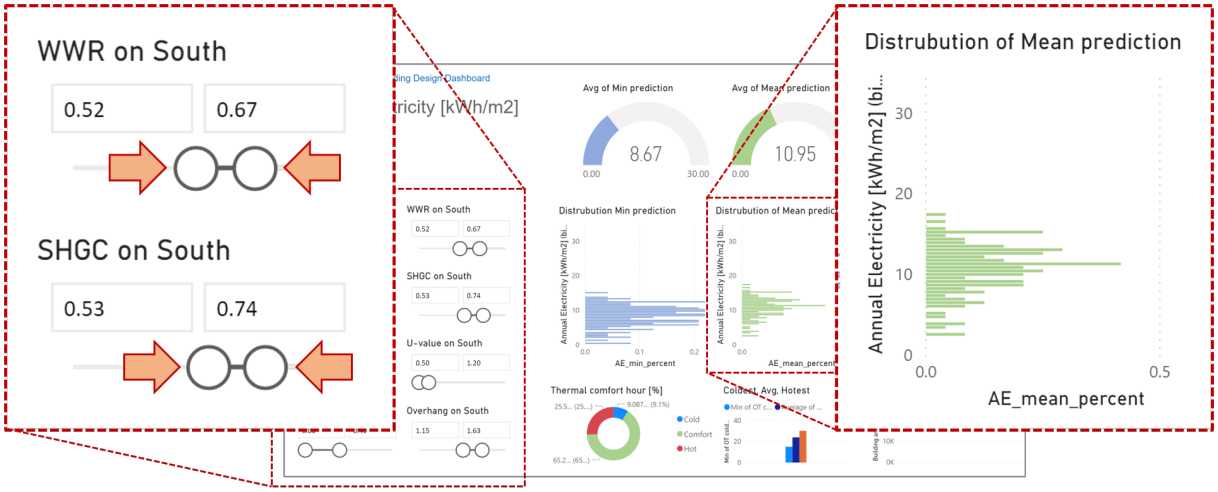


Figure 6-28: Sensitivity dashboard

This is because, with the ranges of parameters getting smaller, the filtered cases in the database gets fewer. As these cases are generated using Monte Carlo method, small sample size results in occasional results and unsmooth distribution curves. On the other hand, Power BI is database based visualization tool, which can neither utilize the trained GPR models, nor even do interpolations.

We imagined that the design team decided to use double glass window, but whether use Low-E glass or not had not been determined. Also, WWRs and shadings were decided to be mid-leveled. Table 6-16 shows the ranges of parameters, which the design team wanted to look deeper into.

We generated 10,000 additional cases using LHS within the shrunk parameter ranges, and added them

Table 6-16: Shrunk parameter ranges

Orientation	WWR	U-factor	SHGC	Fin depth	Overhang scale
East	[0.3, 0.6]	[0.7, 1.7]	[0.46, 0.77]	(0, 0.3]	
East	[0.3, 0.6]	[0.7, 1.7]	[0.46, 0.77]		[1, 1.5]

into the database. These additional cases were fed into the trained GPR of each criterion and each climate. Likewise, the μ, $\mu - \sigma$ and $\mu + \sigma$ values were extracted and added into the database. As illustrated in figure 6-29, with added data, the distribution got quite smooth. The design team could do more specific tests on each parameter.

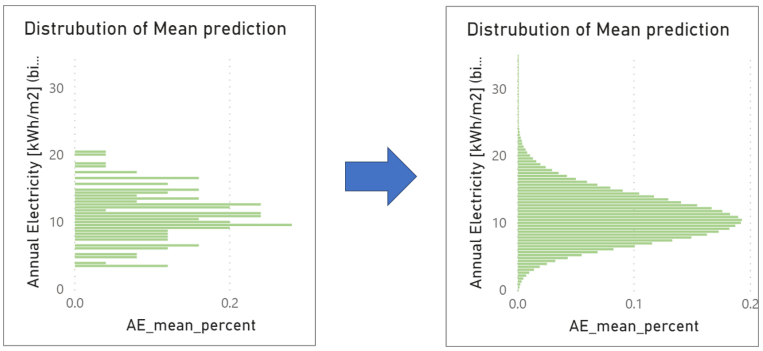


Figure 6-29: Sensitivity dashboard

6.8 Adding simulations and dashboards

From the sensitivity page (figure 6-26), we found that the sensitivity of the north opening was quite high, even higher than those on other orientations, which was a blind spot. Especially, the sensitivity of SHGC of north opening over cooling load was much higher than we expected, which was almost 2 times of that on south.

We though it was reasonable because this building was 15 degrees East of South, the solar radiation might get into the room in the summer afternoon. By the way, it also reminded us that maybe we should do sun path or facade accumulated radiation analysis before conceptual design. The design team decided to attach fins to the north facade. Figure 6-30 illustrates the fins on the north facade. the distance between two adjacent fins was 2m, the with of a fin was within the range of (0m, 1.5m].

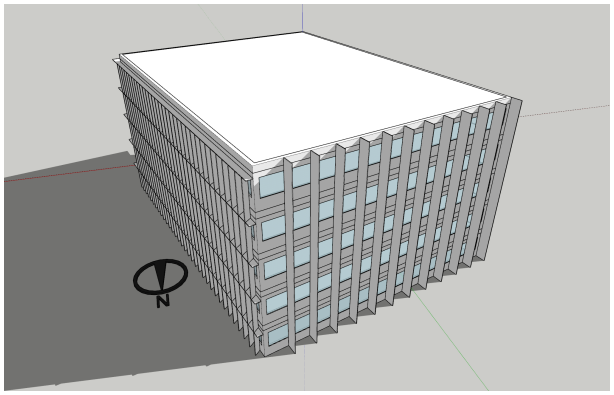


Figure 6-30: Sensitivity dashboard

We did other 300 samples of WWR, SHGC, U-value of the north opening, the newly added fin width on north using LHS. The values of the other parameters were fixed to the mid values of the their ranges. Then each sample was expanded into 20 cases using STMN method. 6000 IDF files were generated parametrically and simulated with the four EPW files. GPR models were trained.

10,000 predictions were made. The samples, as well as the μ, $\mu - \sigma$ and $\mu + \sigma$ values, were added into the database. A new dashboard was made (figure 6-31).

As the whole process was highly automated, it took just several hours to run additional simulations, train new meta-models and make new dashboard.

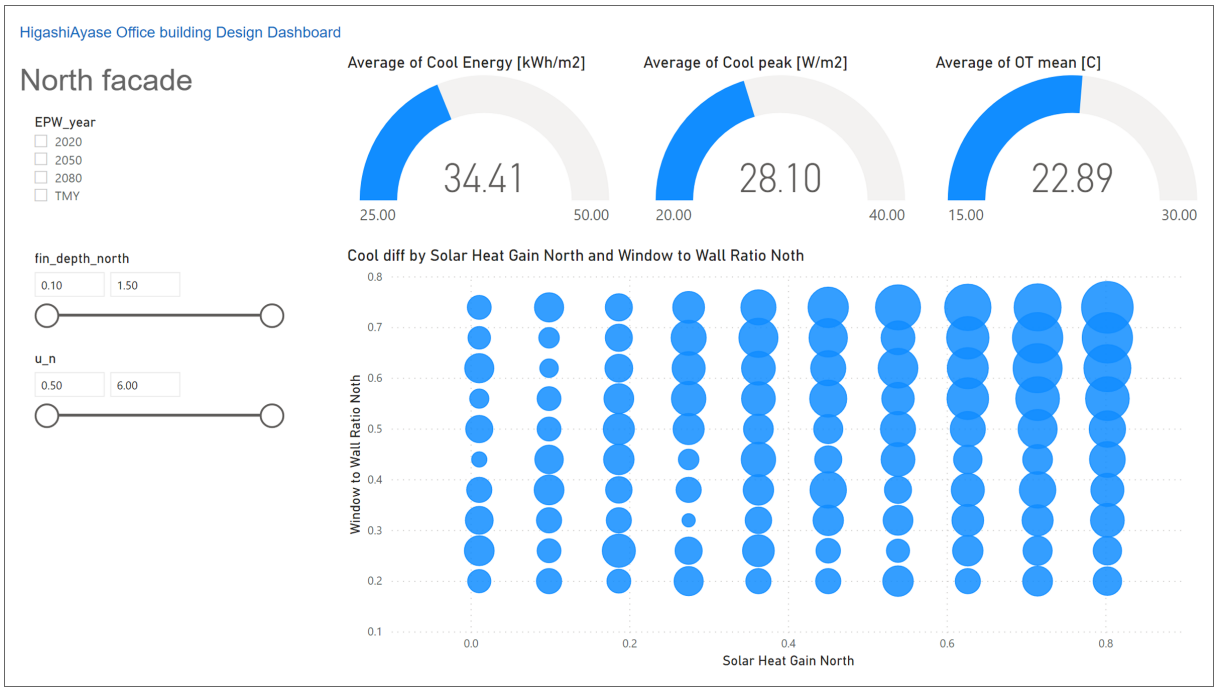


Figure 6-31: Sensitivity dashboard

6.9 Results

In this section, we demonstrated the proposed framework for the early design stage with a case study.

6.9.1 Efficiency of this framework

This first thing we have proved by this case study is the efficiency of the proposed framework.

The CPU of the computer we used in this demonstration is an 12-core one. The whole computer is 1000$ level, which is consumer-grade and widely used in architecture design studio. As we didn't use GPU in this case study, so the cost of the GPU can be ignored.

There were 27 parameters at the beginning. To execute the sensitivity analysis using Morris method, the simulations were executed ($(27+1)\times 20$ =)560 times, which took about 5 minutes. The number of parameters reduced from 27 to 19 based on their sensitivities. Then the interactions between the 19 parameters were analyzed using Extended Morris method. The simulations were executed ($19(19-1)/2+(19-1)+2\times 20$ =)3820 times, which took about 20 minutes. Based on the interaction, the parameters were separated into 5 groups. In this way, the 27-dimensional problem has been simplified into two 8-dimensional, two 7-dimensional and one 1-dimensional problems, which helped the designers avoid the curse of dimensionality.

To train and test GPR models, ignoring the sample collections used for comparison, 10,000 IDF files were generated and simulated with the climate data of TMY, 2020, 2050 and 2080. The simulations were executed 40,000 times, which took about 5 hours. With the trained GP models, the energy performance and thermal comfort of other 200,000 cases were predicted within about 2 minutes.

Using the proposed process, a database containing 200,0000 cases could be made within two days. Comparatively, it takes more than one month to simulate so many cases even using a high-end workstation.

The whole process was highly automated. We rearranged the Python codes into a template. When users try to adapt this framework, it will take them several hours to establish a base IDF file and modify the parametric modeling codes that generate IDF files, because the parameters of interest will be different from project to project. The codes of sensitivity analysis, interactions analysis, meta-modeling and database generation can be directly used with just a little modification in parameters and executed full-automatically. As described before, it will takes about 8 hours to execute the simulations. After database made, it will take several hours to make dashboards. So, it can be said that, for a new project, it will take about 3 days to go over the whole process (table 6-17). Actually, when writing this thesis, as the simulation configurations were adjusted, we redid all the simulations from the very beginning, which took us 2 days.

Table 6-17: Time costs of the whole process

Phases	**Work contents**	**Ways**	**Time costs**
Pre-process	Base IDF establishment Parametric modeling codes modification	Manually	8 hours
Dimension reduction	Sensitivity analysis Interaction analysis	Automatically	8 hours
Meta-modeling	GPR training Database generation		
Post process	Dashboards making	Manually	8 hours

Though it can be ensured that, by introducing this framework, the efficiency get much higher than a traditional trails-and-errors process. However, it is almost impossible to quantify the efficiency or time cost in a trails-and-errors process. We imagined the situation that trails-and-errors process was adopted (table 6-18) and did a qualitative comparison. In this situation, the design team spent several days but just touched quite a few design options. A lot of time was expended in communications.

Table 6-18: Time costs of the whole process

Period		**Executor**	**Work contents**
1^{st} day	Morning	Designers	Propose design options
	Afternoon	Engineers	Analyze energy performance
2^{nd} day	Morning	Designers + Engineers	Meeting
	Afternoon	Designers	Modify design options
3^{rd} day	Morning	Engineers	Analyze energy performance
	Afternoon	Designers + Engineers	Meeting
	Overtime	Designers	Prepare presentation
4^{th} day	Morning	Design team + client	Meeting
	Afternoon	Designers	Modify design options
...	...	...	...

There is another problem that the time cost result in this section is made under the situation of 27 parameters. Would the time cost keep acceptable with the number of parameters (dimension *d*) increasing. The times cost of sensitivity analysis using Morris method is linear to the number of parameter (equation 4.35), while that of interaction analysis using Expanded Morris method is square. The time cost of collecting training data is very difficult to quantify, as it is closely related to the smoothness of landscape of results. Even with the same *d*, the necessary sample sizes could be extremely different for different problems. On the other hand, it is also possible that the necessary sample sizes are similar while the *d* varies. We suggest designers to execute the sensitivity analysis, even with a very *d*, as the time cost increases linearly which would not loss control. Then, if the estimated time cost is too high, designers can reduce *d* based the sensitivity of each parameter.

6.9.2 Uncertainties and Gaussian Process Regression

In this case study, we defined 4 non-targeted parameters, cooling and heating setpoints, internal heat gain and air change rate of natural ventilation. In meta-modeling, these non-targeted parameters were actually uncertainty, rather than parameters, in the early design stage. With training data collected using STMN method, Gaussian process regression showed its robustness when dealing with noisy data. The trained GPR models successfully predicted the uncertainties (variances) caused by the non-targeted parameters. By outputting the boundaries of the confidence interval, designers can also notice the possibility and risk of each decision, rather than only watch on the mean value.

6.9.3 Supports for communication and decision-making

Using dashboards, the design team can efficiently test different combination of parameters and avoiding doing duplicated works in traditional trail-error process. This process can also make communication between architects and engineers more smooth. They test their ideas with the dashboards and get real-time feedback when negotiating, avoiding quarreling without solid proofs. The dashboards can also benefit the meetings with clients, as they can be accessed by iPads or smart phones through Power BI mobile application or web browser.

The results of sensitivity analysis can remind the design team the importance of parameters. By understanding the results of interaction analysis, the design team can know which parameters should be studied together and avoid that the decisions made early phases influenced by those in latter phases.

By widely exploring the decision space and showing the sensitivities, a blind point, the shadings on the north façade, was found.

Chapter 7

Conclusion and future work

In this research, we pointed out several currently exiting problems in the early stage of architecture environmental design. Through literature review, it was found that Monte Carlo method, sensitivity analysis, meta-modeling and interactive data visualization have been utilized by researchers to solve these problems. But these technics are still fragmented. Therefore, we proposed a framework that integrated all the technics above to help designers make decisions. With this proposed framework, a design team can widely explore the decision space make dashboards very efficiently.

Interaction analysis has been introduced into this research to separate parameters into groups to reduces the dimension. The simulated or predicted criteria could be combined using PCP method. GPR has been used as the meta-modeling method in this research. A trained GPR is able to predict the distribution of a criterion, which is very robust at regressing uncertain data. In order to perform decent training when the uncertainties caused by NTPs is too high, we proposed STMN method.

Certainly, there are several limitations in this research. Correspondingly, we proposed future work in the last section. We noticed problem of the correlations between criteria, so that we proposed coregionalized GPR as future work. We also planed to integrate Robust optimization with GPR and GA. In the near future, we will also positively introduce this framework into real projects.

7.1 Conclusion

At the beginning of this thesis, by revealing the gaps between calculated and measured energy performance in buildings, we led to the topic of uncertainty. These gaps arise duo to the uncertainties in the input parameters of simulations. Besides the uncertainty issues, there were also several other problems in current design process in the early stage.

We concluded 5 problems that we intended to solve for the early stage of architecture environmental design in this research:

- How to keep the calculation cost at a low level?
- How to take uncertainties into consideration in the early design stage?
- How to fully explore the design possibilities?
- How to utilize the data to support decision-making and benefit communications?
- How to analyze the interactions and avoid 'looped' work?

Through literature review, we found that several technics in the fields of computer science and statistic have been introduced into the building energy field to solve these problems. Sensitivity analysis has been used to quantify the importance of parameters and reduce the number of parameters. Informing designers with the importance of parameters can also help them make decisions. Several kinds of meta-modeling methods have been utilized to regress simulation data. Many researchers have employed Monte Carlo method to reproduce the uncertainties in the inputs and quantify the uncertainties in the output. Meanwhile, several kinds of interactive data visualization method have been proposed by researchers to help designers in decision-making.

However, these technics were still fragments. A framework was needed to integrate all the technics above. Moreover, though the technics of interaction analysis have been well developed in statistic field, it has not been fully used in architecture environmental design. Gaussian process has been used as a meta-modeling method by researchers. Its robustness on regressing uncertain data, however, has not been fully utilized.

Therefore, in this research, we proposed an integrated decision-making framework that lined up sensitivity and interaction analysis, meta-modeling using GPR and interactive data visualization. Figure 7-1 is a simplified flow chart of this proposed framework.

As we mainly used statistic methods in this research, a large amount of data is needed, meanwhile simulations were the only resource of data in the early stage of data. Therefore, we began our working with developing tools that is able to generate models automatically and execute simulations in parallel. The achievements were EPPiX, Ultimate EP executor, GH2FD and PyFD.

Like other researchers, we used Morris method to analyze the sensitivities of parameters and screen them. We also introduced Expanded Morris Method into this research to analyze the interactions between parameters and separate them into groups. In this way, the number of parameters were dramatically reduced, so that we can **better explore the design possibilities with less samples**, which also **significantly reduced the time cost of simulations**. **Parameters in one group can be studied and decided together, to avoid being influenced by those in other groups and 'looped' work.**

Gaussian process regression is used as the meta-modeling method in this research. Rather than include the NTPs in the training data, we regarded the uncertainties in the simulation results caused by them as fluctuation. Only targeted parameters and simulation results were included in the training data. **In this way, a trained GPR model was able to predict the probability distribution of a criterion, which directly reflects the uncertainty in this criteria.** With research progressing, it was found that if the fluctuation in the training data was too big, larger than the variety caused by targeted parameters, a trained GPR model might fail to give out a decent prediction. Therefore, we proposed a stratified sampling method named STMN to perform decent GPR training with highly noisy data.

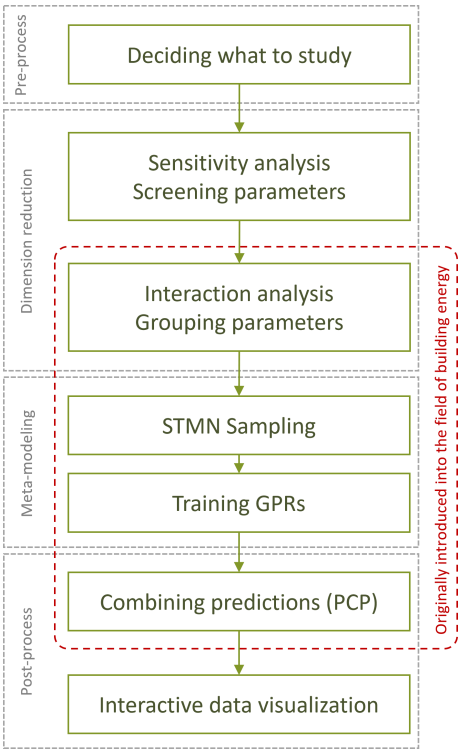


Figure 7-1: Proposed framework (simplified)

For the situations that parameters belonging different groups should be studied together, we proposed the PCP method to combine predictions from GPR models trained for different groups into one distribution.

Finally, with the trained GPR models, we generated a huge amount of predicted distributions and visualized them interactively in the form of dashboards. **Using these dashboard, designers could test their ideas and got real-time feedback when making decisions. These dashboards could also benefit communications.**

In Chapter 6, we demonstrated this proposed framework with an office building design project. The high efficiency of this framework has been well proved with this case study.

7.2 Limitations and future work

7.2.1 More criteria, more projects

In this research, we concentrated on energy performance and thermal environment in this research. However, energy performance is just a small part in architecture design. Other factors, such as construction cost. Theoretically, using this framework, designers can study any problems that can be quantified. If the input and output information can be quantified, and the calculation cost is not too high, sensitivity and interaction analysis, as well as training GPR. For those non-quantitable problems, designers can judge by themselves when using dashboards.

It is really a pity that we did not get a change to introduce this framework into a real project. We really hope to hear the feedback from professional architects. As I have decided to work as an architect for several year before teaching and researching in a university, I will positively introduce my researches into real projects.

7.2.2 Covariance, correlation and coregionalization

In this research, we mainly trained GPR models for different criteria independently. In most situations, for the same sample of targeted parameters, the distributions of different criteria, caused by the multi-samples of NTPs, are independent from each other. Nevertheless, in some situations, duo to the configurations of NTPs, the distributions of different criteria are no longer independent. The joint distribution of multiple criteria should be considered.

To explain this problem, we used the model from chapter 6 to make a demonstration. For the targeted parameters, the mid values of their ranges were used. The ranges of NTPs were adjusted, to see their influences on the distributions of cooling and heating load. In another word, all cases share the same building factors, but the NTPs are all different.

We made 3 sets of configurations, as shown in the upper part of table 7-1. In the 1^{st} set, the main uncertainty is the cooling and heating setpoints. The spans the ranges is up tp $4°C$. The variance in this loads caused by the AC setpoints dominates. In this 2^{nd} set, the spans of the AC setpoint ranges is $0.5°C$, while that of the internal heat gain was enlarged to $12W/m^2$. The main uncertainty in this set is the internal heat gain. In the 3^{nd} set, the main uncertainty infiltration rate, from 0 to 3.

For each set, 200 samples of NTPs were generated using LHS, based on which IDF files were generated parametrically and simulated with *EnergyPlus*. The annual cooling and heating loads were outputted.

Figure 7-2 illustrates the joint distributions of cooling and heating loads. The x coordinate of each point is the heating load of a case, while the y coordinate stands for the cooling load. From the first diagram, we can

Table 7-1: Configurations of covariance tests

		Set 1	**Set 2**	**Set 3**
Cooling setpoint	$°C$	$\mathbf{[22, 26]}$	$[25.75, 24.25]$	$[25.75, 24.25]$
Heating setpoint	$°C$	$\mathbf{[16, 20]}$	$[17.75, 18.25]$	$[17.75, 18.25]$
Internal heatgain	W/m^2	$[8, 10]$	$\mathbf{[3, 15]}$	$[8, 10]$
Infiltration rate	$1/h$	$[0.3, 0.5]$	$[0.3, 0.5]$	$\mathbf{[0, 3]}$

see that there is no obvious relationship between cooling and heating load, as the cooling and heating setpoints respectively determined the heating and cooling load, which is independent. When the main uncertainty is internal heat gain, the cooling and heating loads present a negative relationship, for the reason that high internal heat gain reduces heating load but meanwhile increases cooling load. From the last diagram, we can find that, in the case that the main uncertainty is the infiltration rate, the relationship between cooling and heating load is positive, as they both arise with the building getting more infiltration.

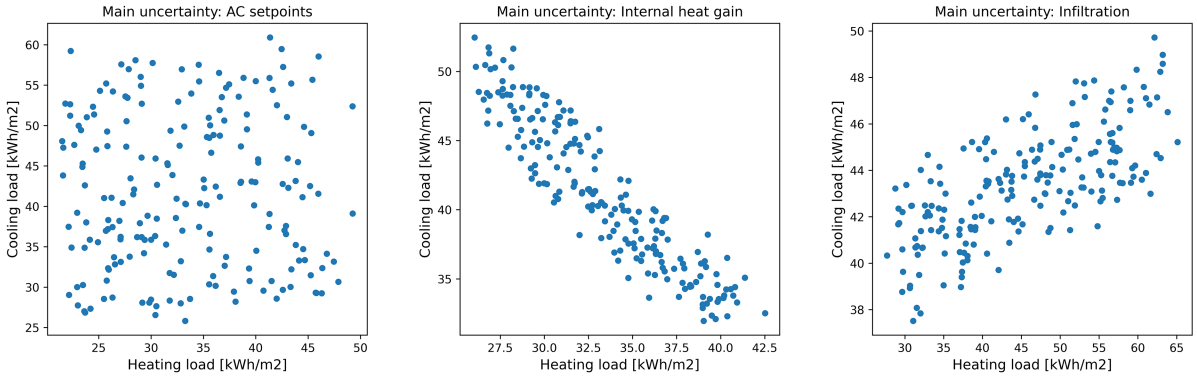


Figure 7-2: Joint distributions of heating and cooling load

In the situation that criteria are not independent, their covariance should be taken into consideration. The covariance of two variables is a measure of their joint variability. Equation 7.1 is the definition the covariance of v_1 and v_2. After some mathematical transformation, the formula to calculate covariance can be simplified into equation 7.1a.

$$
\begin{aligned}
Cov(v_1, v_2) &= \sum_{i=1}^{n} \frac{(v_1 i - \overline{v_1})(v_2 i - \overline{v_2})}{n} && (7.1)\\
&= \sum_{i=1}^{n} \frac{v_1 i \times v_2 i - \overline{v_1} \times v_2 i - v_1 i \times \overline{v_2} + \overline{v_1} \times \overline{v_2}}{n} \\
&= \sum_{i=1}^{n} \frac{v_1 i \times v_2 i}{n} - \overline{v_1} \sum_{i=1}^{n} \frac{v_2 i}{n} - \overline{v_2} \sum_{i=1}^{n} \frac{v_1 i}{n} + \overline{v_1} \times \overline{v_2} \\
&= \overline{v_1 \times v_2} - \overline{v_1} \times \overline{v_2} - \overline{v_1} \times \overline{v_2} + \overline{v_1} \times \overline{v_2} \\
&= \overline{v_1 \times v_2} - \overline{v_1} \times \overline{v_2} && (7.1a)
\end{aligned}
$$

From the definition formula of covariance (7.1), we can find that if v_1 and v_2 have the same tendency, when $v_1 i$ is larger/smaller than the mean value $\overline{v_1}$, the $v_2 i$ in the same position is more likely to be larger/smaller than $\overline{v_2}$, the covariance of v_1 and v_2 tends to be larger than 0. In contrast, if the tendencies of v_1 and v_2 are opposite, the value of their covariance is negative.

However, the value of the covariance is strongly influenced by the ranges and unit of the original data, so that it is difficult to correctly evaluate the correlation between two variables based on their covariance. In the case that we want to assess the correlation, **P**earson **C**orrelation **C**oefficient (PCC) should be used. The definition formula of PCC is as shown in equation 7.2. The PCC can be regarded as a normalized value of covariance of 2 variables. The value of PCC is always within -1 to 1. A PCC value larger than 0.5 indicates strong positive correlation, while a value less than -0.5 indicates strong negative correlation. If the PCC value is close to 0, the 2 variables are non-correlated.

$$\rho_{v_1,v_2} = \frac{Cov(v_1, v_2)}{\sigma_1 \sigma_2} \tag{7.2}$$

$$\rho_{v_1,v_2} \in [-1, 1] \tag{7.2a}$$

We also analyzed the covariance and correlation between cooling and heating loads for the 3 sets.

Table 7-2: Results of covariance tests

	Set 1	**Set 2**	**Set 3**
Main uncertainty	AC setpoints	Internal heat gain	Infiltration rate
Var(Cooling load)	87.94	28.23	5.44
Var(Heating load)	56.75	17.14	101.78
Covar(Cooling, Heating)	0.42	-20.34	16.4
Correlation coefficient	0.006	-0.92	0.68

It should be noticed that a PCC value dose not indicate the causality of the 2 variables. For example, though the PCC in set 2 is -0.92, indicating a very strong negative relationship between cooling and heating loads, the cooling load is certainly not the reason, nor the results, of the heating load. In contrast, a 0-valued PCC dose not mean that 2 variables are independent. For example, in the function $y = \sin(x) + fluctuation$ where $x \in [-3\pi, 3\pi]$, $fluctuation \in [-0.2, 0.2]$, the PCC value of x and y is 0. But x and y are obviously not independent.

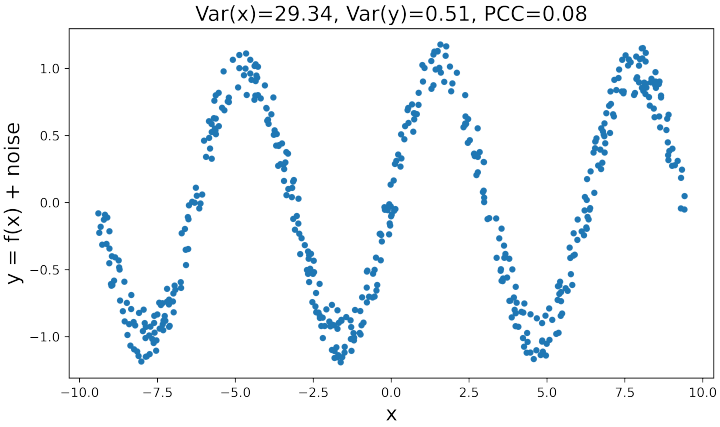


Figure 7-3: Correlation between x and sin (x)

We noticed that we had not analyzed the correlation between different criteria in the case study presented in the last chapter. Therefore, we picked up 20 sample sets from the simulation data of TMY climate, and analyzed the PCC between cooling and heating loads of each set to see whether we made mistakes.

Fortunately, most of the PCC values are less than 0.5. The average is 0.18, which means cooling and heating loads of a same sample of targeted parameters are not correlated. Therefore, there was no mistake to train GPR models of cooling and heating load independently, like what we did in the case study.

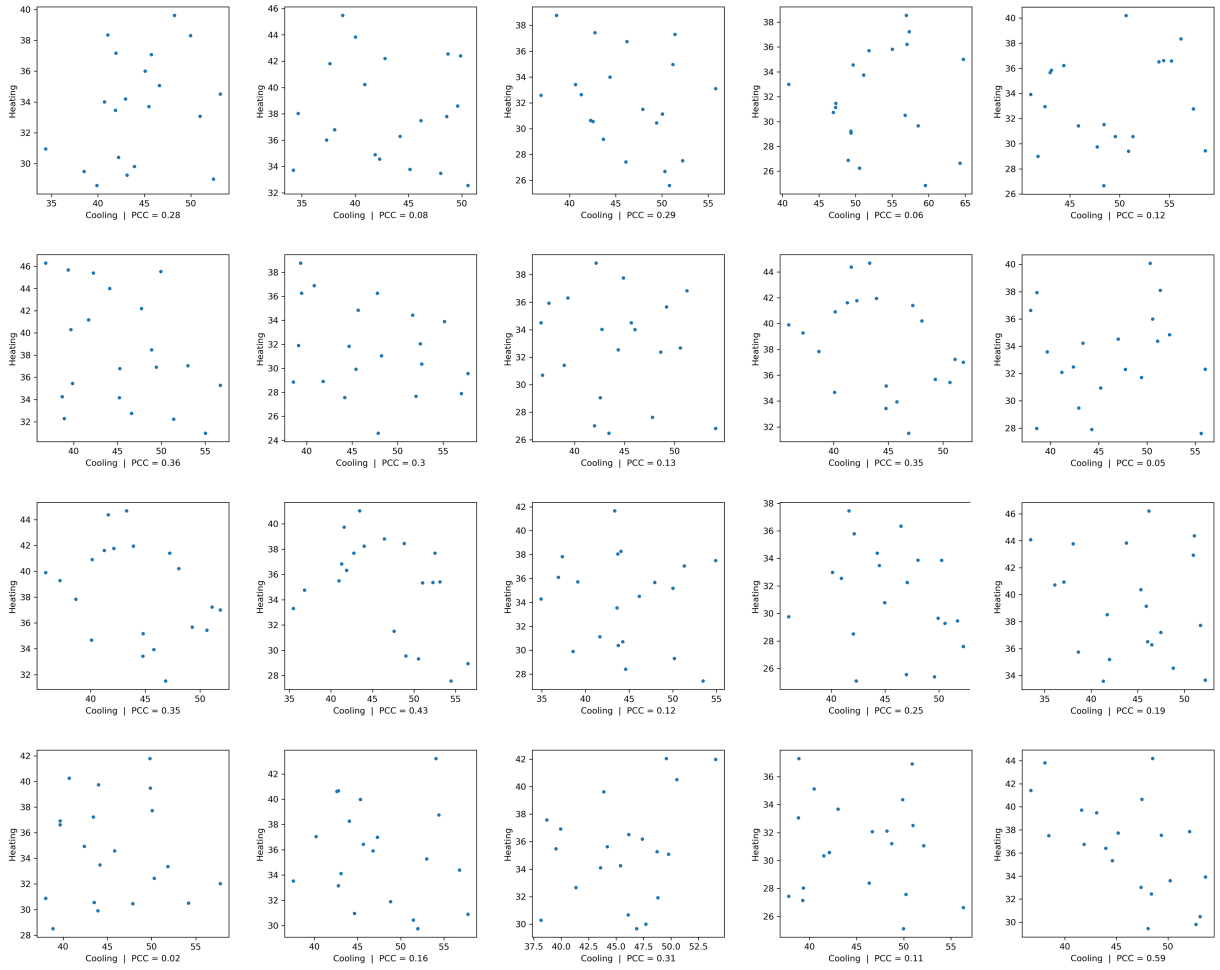


Figure 7-4: Correlations between cooling and heating

In the case that the correlation between two criteria is not negligible, a coregionalized regression model[76] is a solution to predict the joint distributions. In a coregionalized GPR model, relies on the use of multiple output kernels[19] to generate the covariance matrix of two matrix (equation 7.6).

$$X = \{x_1, x_2, x_3, \cdots, x_n\} \tag{7.3}$$

$$Y_i = f_i(X) \tag{7.4}$$

$$Y_j = f_j(X) \tag{7.5}$$

$$Cov(Y_i, Y_j) = k(X, X) \cdot B_{i,j} \tag{7.6}$$

$$= \begin{bmatrix} k(x_1, x_1) \times B_{i,j(1,1)} & \cdots & k(x_1, x_n) \times B_{i,j(1,n)} \\ \vdots & \ddots & \vdots \\ k(x_n, x_1) \times B_{i,j(n,1)} & \cdots & k(x_n, x_n) \times B_{i,j(n,n)} \end{bmatrix} \tag{7.6a}$$

Where:

Y_i, Y_j	:	The value array of the i^{th}/j^{th} criterion
$Cov(Y_i, Y_j)$	:	The covariance matrix of Y_i and Y_j generated using multiple output kernel
k	:	The kernel function
$B_{i,j}$	:	The coregionalization matrix

$B_{i,j}$ is the coregionalization matrix, which should be positive definite. Therefore, this coregionalization matrix can be defined as equation 7.7.

$$B_{i,j} = WW^T + \kappa I \tag{7.7}$$

$$I = \begin{bmatrix} 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 1 \end{bmatrix} \tag{7.8}$$

Where:

W	:	A parameter matrix
κ	:	A n-dimensional parameter vector
I	:	A n^2 diagonal matrix

7.2.3 Robust optimization

In the introduction section, we criticized that optimization is not a suitable decision-making method in the early stage of design. However, we still believe that, after designers studied the parameters using dashboards and shrunk the ranges of parameters, optimization can yet be a efficient way to finalize the value of parameters.

In the literature review, we introduced the research by Bamdad et al.[54] that 3 scenarios were assumed in the optimization procedure. Actually, this optimization method used in their research can be regarded as a kind of simplified **R**obust **O**ptimization (RO)[33]. In the 2^{nd} chapter, we introduced the concept of antifragile, which can be described as 'try to avoid failures under different conditions'. Similarly, the concept of RO can be described as 'optimize in the worst-case scenario', or 'try to improve the worst result in all scenarios'. RO is a kind of optimization method under uncertainty, which has been well developed in computer science field.

In this research, we trained GPR models able to predict probability distributions of the objective variables. A predicted distribution represent 'the probable values of the objective variables in all scenarios'. Therefore, GPR meta-modeling can naturally be integrated into RO procedure.

As future work, we plan to integrate GPR, GA and RO together to perform optimizations under uncertainty. We considered 2 ways to evaluate a candidate. The first one is to 'use the bottom of a certain confidence interval as the worst-case'. For example, if we use the bottom of 95% confidence interval, then we can say that in 95% cases, the value of the objective variable is better than the optimized result. Another way to 'use the Pareto frontier of mean and variance'. In this way, we can get two kinds of optimized results, 'a wide distribution with relatively good mean value' and 'a narrow distribution with relatively bad mean value'.

Bibliography

[1] Donald E. Knuth. “Structured Programming with go to Statements”. In: *ACM Computing Surveys* 6 (4 Dec. 1972), pp. 1–41. DOI: 10.1145/356635.356640.

[2] Christopher Alexander, Sara Ishikawa, and Murray Silverstein. *A Pattern Language: Towns, Buildings, Construction*. New York: Oxford University Press, Aug. 1977. ISBN: 0195019199. URL: http://www.amazon.fr/exec/obidos/ASIN/0195019199/citeulike04-21.

[3] 滝沢博. “標準問題の提案 (オフィス用標準問題)”. In: 第 *15* 回熱シンポジウム「伝熱解析の現状と課題」. 日本建築学会環境工学委員会熱分科会. Sept. 1985.

[4] 武宮 正樹. 宇宙流基本知識–武宮囲碁ワールド. Chiyoda city, Tokyo, JP: 日本棋院, Jan. 1989. ISBN: 4818202460.

[5] T. Ishigami and T. Homma. “An importance quantification technique in uncertainty analysis for computer models”. In: *[1990] Proceedings. First International Symposium on Uncertainty Modeling and Analysis*. 1990, pp. 398–403.

[6] Nebojsa Nakicenovic et al. *Special Report on Emissions Scenarios*. 1st ed. Cambridge, UK: Cambridge University Press, 2000. ISBN: 0-521-80493-0.

[7] Drury B. Crawley et al. “EnergyPlus: Creating a new-generation building energy simulation program”. In: *Energy and Buildings* 33.4 (2001), pp. 319–331. ISSN: 03787788. DOI: 10.1016/S0378-7788(00)00114-6.

[8] California Public Utilities Commission. *California long term energy efficiency strategic plan: achieving maximum energy savings in California for 2009 and beyond*. Sept. 2008. URL: https://www.cpuc.ca.gov/WorkArea/DownloadAsset.aspx?id=5305.

[9] Mark F. Jentsch, AbuBakr S. Bahaj, and Patrick A.B. James. “Climate change future proofing of buildings–Generation and assessment of building simulation weather files”. In: *Energy and Buildings* 40.12 (2008), pp. 2148–2168. ISSN: 0378-7788. DOI: https://doi.org/10.1016/j.enbuild.2008.06.005.

[10] Cathy Turner et al. *Energy Performance of LEED for New Construction Buildings*. Mar. 2008. URL: https://newbuildings.org/sites/default/files/Energy_Performance_of_LEED-NC_Buildings-Final_3-4-08b.pdf.

[11] Per Heiselberg et al. “Application of sensitivity analysis in design of sustainable buildings”. In: *Renewable Energy* 34.9 (Mar. 2009), pp. 2030–2036. ISSN: 09601481. DOI: 10.1016/j.renene.2009.02.016.

[12] Guy R. Newsham, Sandra Mancini, and Benjamin J. Birt. “Do LEED-certified buildings save energy? Yes, but…” In: *Energy and Buildings* 41.8 (Aug. 2009), pp. 897–905. ISSN: 0378-7788. DOI: 10.1016/J.ENBUILD.2009.03.014.

[13] Arno Schlueter and Frank Thesseling. “Building information model based energy/exergy performance assessment in early design stages”. In: *Automation in Construction* 18.2 (Feb. 2009), pp. 153–163. ISSN: 0926-5805. DOI: 10.1016/J.AUTCON.2008.07.003.

[14] John H. Scofield. “Do LEED-certified buildings save energy? Not really…” In: *Energy and Buildings* 41.12 (Dec. 2009), pp. 1386–1390. ISSN: 0378-7788. DOI: 10.1016/J.ENBUILD.2009.08.006.

[15] Harvard Instructional Physics Laboratory. *A Summary of Error Propagation*. (accessed Jun 29, 2019). 2010. URL: http://ipl.physics.harvard.edu/wp-uploads/2013/03/PS3_Error_Propagation_sp13.pdf.

[16] T. Ramesh, Ravi Prakash, and K. K. Shukla. “Life cycle energy analysis of buildings: An overview”. In: *Energy and Buildings* 42.10 (May 2010), pp. 1592–1600. ISSN: 03787788. DOI: 10.1016/j.enbuild.2010.05.007.

[17] Janelle S. Hygh et al. “Multivariate regression as an energy assessment tool in early building design”. In: *Building and Environment* 57 (Apr. 2012), pp. 165–175. ISSN: 03601323. DOI: 10.1016/j.buildenv.2012.04.021.

[18] Lynne Blundell. *Environmental upgrade agreements*. 2012. URL: https://sustainableaustraliafund.com.au/wp-content/uploads/2014/03/TFE_eBOOK_FINAL1.pdf.

[19] Neil D. Lawrence Mauricio A. Alvarez Lorenzo Rosasco. “Kernels for Vector-Valued Functions: a Review”. In: (Apr. 2012).

[20] Nassim Nicholas Taleb. *Antifragile: Things That Gain from Disorder*. 1st ed. York City, US: Random House Publishing Group, Nov. 2012. ISBN: 0679645276.

[21] Shady Attia et al. “Assessing gaps and needs for integrating building performance optimization tools in net zero energy buildings design”. In: *Energy and Buildings* 60 (May 2013), pp. 110–124. ISSN: 03787788. DOI: 10.1016/j.enbuild.2013.01.016. URL: http://linkinghub.elsevier.com/retrieve/pii/S0378778813000339.

[22] D. Gossard, B. Lartigue, and F. Thellier. “Multi-objective optimization of a building envelope for thermal performance using genetic algorithms and artificial neural network”. In: *Energy and Buildings* 67 (Dec. 2013), pp. 253–260. ISSN: 0378-7788. DOI: 10.1016/J.ENBUILD.2013.08.026.

[23] Neil Hirst. “Buildings and climate change: Summary for Decision Makers”. In: *Design and Management of Sustainable Built Environments*. Ed. by Runming Yao. London, UK: Springer, London, Mar. 2013. ISBN: 9781447147817. DOI: 10.1007/978-1-4471-4781-7_2.

[24] Sanja Stevanović. “Optimization of passive solar design strategies: A review”. In: *Renewable and Sustainable Energy Reviews* 25 (2013), pp. 177–196. ISSN: 13640321. DOI: 10.1016/j.rser.2013.04.028.

[25] Ehsan Asadi et al. “Multi-objective optimization for building retrofit: A model using genetic algorithm and artificial neural network and an application”. In: *Energy and Buildings* 81 (Oct. 2014), pp. 444–456. ISSN: 0378-7788. DOI: 10.1016/J.ENBUILD.2014.06.009.

[26] Ward Cunningham. *Premature Optimization*. (accessed Sep 30, 2020). 2014. URL: https://wiki.c2.com/?PrematureOptimization.

[27] Daniel Daly, Paul Cooper, and Zhenjun Ma. "Understanding the risks and uncertainties introduced by common assumptions in energy simulations for Australian commercial buildings". In: *Energy and Buildings* 75 (June 2014), pp. 382–393. ISSN: 0378-7788. DOI: 10.1016/J.ENBUILD.2014.02.028.

[28] D. Garcia Sanchez et al. "Application of sensitivity analysis in building energy simulations: Combining first- and second-order elementary effects methods". In: *Energy and Buildings* 68 (Jan. 2014), pp. 741–750. ISSN: 0378-7788. DOI: 10.1016/J.ENBUILD.2012.08.048.

[29] Arthur Santos Silva and Enedir Ghisi. "Uncertainty analysis of the computer model in building performance simulation". In: *Energy and Buildings* 76 (2014), pp. 258–269. ISSN: 0378-7788. DOI: https://doi.org/10.1016/j.enbuild.2014.02.070.

[30] Liesje Van Gelder, Hans Janssen, and Staf Roels. "Probabilistic design and analysis of building performances: Methodology and application example". In: *Energy and Buildings* 79 (Aug. 2014), pp. 202–211. ISSN: 0378-7788. DOI: 10.1016/J.ENBUILD.2014.04.042.

[31] Autodesk. *Insight - High performance and sustainable building design analysis*. (access Dec 25, 2019). 2015. URL: https://insight360.autodesk.com/oneenergy.

[32] China Academy of Building Research. *Design Standard for Energy Efficiency of Public Buildings*. 1st ed. Beijing, PRC: China Architecture and Building Press, Feb. 2015. ISBN: 1511226431.

[33] Bram L. Gorissen, Ihsan Yanikoğlu, and Dick den Hertog. "A practical guide to robust optimization". In: *Omega (United Kingdom)* 53 (2015), pp. 124–137. ISSN: 03050483. DOI: 10.1016/j.omega.2014.12.006. arXiv: 1501.02634.

[34] Farzad Jalaei and Ahmad Jrade. "Integrating building information modeling (BIM) and LEED system at the conceptual design stage of sustainable buildings". In: *Sustainable Cities and Society* 18 (Nov. 2015), pp. 95–107. ISSN: 22106707. DOI: 10.1016/j.scs.2015.06.007.

[35] Emanuele Naboni et al. "Defining The Energy Saving Potential of Architectural Design". In: *Energy Procedia* 83 (Dec. 2015), pp. 140–146. ISSN: 1876-6102. DOI: 10.1016/J.EGYPRO.2015.12.204.

[36] Anh Tuan Nguyen and Sigrid Reiter. "A performance comparison of sensitivity analysis methods for building energy models". In: *Building Simulation* 8.6 (June 2015), pp. 651–664. ISSN: 19968744. DOI: 10.1007/s12273-015-0245-4.

[37] Mohammad Rahmani Asl et al. "BPOpt: A framework for BIM-based performance optimization". In: *Energy and Buildings* 108 (Sept. 2015), pp. 401–412. ISSN: 03787788. DOI: 10.1016/j.enbuild.2015.09.011.

[38] Yongjun Sun, Pei Huang, and Gongsheng Huang. "A multi-criteria system design optimization for net zero energy buildings under uncertainties". In: *Energy and Buildings* 97 (June 2015), pp. 196–204. ISSN: 0378-7788. DOI: 10.1016/J.ENBUILD.2015.04.008. URL: https://www.sciencedirect.com/science/article/pii/S0378778815002856.

[39] Lai Wei et al. "Comparative Study on Machine Learning for Urban Building Energy Analysis". In: *Procedia Engineering* 121 (Jan. 2015), pp. 285–292. ISSN: 1877-7058. DOI: 10.1016/J.PROENG.2015.08.1070.

[40] 経済産業省. *Definition of ZEH and future measures proposed by the ZEH Roadmap Examination Committee*. Dec. 2015. URL: https://www.enecho.meti.go.jp/category/saving_and_new/saving/zeh_report/pdf/report_160212_en.pdf.

[41] Edith Cherry and John Petronis. *Architectural Programming | WBDG - Whole Building Design Guide*. (accessed Dec 25, 2019). 2016. URL: https://www.wbdg.org/design-disciplines/architectural-programming.

[42] Mingya Zhu et al. "An alternative method to predict future weather data for building energy demand simulation under global climate change". In: *Energy and Buildings* 113 (Feb. 2016), pp. 74–86. ISSN: 0378-7788. DOI: 10.1016/J.ENBUILD.2015.12.020.

[43] U.S. Energy Information Administration. *International energy outlook 2017*. Sept. 2017. URL: https://www.eia.gov/outlooks/ieo/pdf/0484(2017).pdf.

[44] Jianli Chen et al. "Uncertainty analysis of thermal comfort in a prototypical naturally ventilated office building and its implications compared to deterministic simulation". In: *Energy and Buildings* 146 (July 2017), pp. 283–294. ISSN: 0378-7788. DOI: 10.1016/J.ENBUILD.2017.04.068.

[45] ESCO・エネルギーマネジメント推進協議会. 【新版】*ESCO* のススメ. 2017. URL: https://www.jaesco.or.jp/esco-energy-management/esco/esco-about-esco/.

[46] Ghjuvan Antone Faggianelli, Laurent Mora, and Rania Merheb. "Uncertainty quantification for Energy Savings Performance Contracting: Application to an office building". In: *Energy and Buildings* 152 (Oct. 2017), pp. 61–72. ISSN: 03787788. DOI: 10.1016/j.enbuild.2017.07.022.

[47] Advanced Knowledge Laboratory Inc. *FlowDesigner* オートメーションとは | *FlowDesigner* オートメーション. (accessed Sep 30, 2020). 2017. URL: https://flowdesigner.jp/automation/about/.

[48] Advanced Knowledge Laboratory Inc. 製品情報｜流体シミュレーション・*FlowDesigner* のアドバンスドナレッジ研究所（*AKL*）. (accessed Sep 30, 2020). 2017. URL: http://www.akl.co.jp/products/flowdesigner/.

[49] Torben Østergård, Rasmus L. Jensen, and Steffen E. Maagaard. "Early Building Design: Informed decision-making by exploring multidimensional design space using sensitivity analysis". In: *Energy and Buildings* 142 (May 2017), pp. 8–22. ISSN: 0378-7788. DOI: 10.1016/J.ENBUILD.2017.02.059.

[50] Inyoung Park et al. "Climate change prediction based a long short-term memory neural network". In: *4th World Conference on CLIMATE CHANGE*. International Association for Engineering Geology and the Environment. Oct. 2017.

[51] THE EUROPEAN PARLIAMENT and THE COUNCIL OF THE EUROPEAN UNION. *EUR-Lex - 32012L0027 - EN - EUR-Lex*. (accessed Jun 28, 2020). Oct. 2017. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32012L0027.

[52] Liping Wang, Xiaohong Liu, and Hunter Brown. "Prediction of the impacts of climate change on energy consumption for a medium-size office building with two climate models". In: *Energy and Buildings* 157 (Dec. 2017), pp. 218–226. ISSN: 0378-7788. DOI: 10.1016/J.ENBUILD.2017.01.007.

[53] Shady Attia. *Net Zero Energy Buildings: Concepts, Frameworks and Roadmap for Project Analysis and Implementation*. 1st ed. Oxford, UK: Butterworth-Heinemann, Mar. 2018. ISBN: 9780128124611.

[54] Keivan Bamdad et al. “Building energy optimisation under uncertainty using ACOMV algorithm”. In: *Energy and Buildings* 167 (May 2018), pp. 322–333. ISSN: 0378-7788. DOI: 10.1016/J.ENBUILD.2018.02.053.

[55] Maria Ferrara, Elisa Sirombo, and Enrico Fabrizio. “Automated optimization for the integrated design process: the energy, thermal and visual comfort nexus”. In: *Energy and Buildings* 168 (June 2018), pp. 413–427. ISSN: 0378-7788. DOI: 10.1016/J.ENBUILD.2018.03.039.

[56] Richard Gagnon, Louis Gosselin, and Stéphanie Decker. “Sensitivity analysis of energy performance and thermal comfort throughout building design process”. In: *Energy and Buildings* 164 (Apr. 2018), pp. 278–294. ISSN: 0378-7788. DOI: 10.1016/J.ENBUILD.2017.12.066.

[57] Torben Østergård, Rasmus Lund Jensen, and Steffen Enersen Maagaard. “A comparison of six metamodeling techniques applied to building performance simulations”. In: *Applied Energy* 211.October 2017 (Oct. 2018), pp. 89–103. ISSN: 03062619. DOI: 10.1016/j.apenergy.2017.10.102.

[58] Javier M. Rey-Hernández et al. “Modelling the long-term effect of climate change on a zero energy and carbon dioxide building through energy efficiency and renewables”. In: *Energy and Buildings* 174 (Sept. 2018), pp. 85–96. ISSN: 03787788. DOI: 10.1016/j.enbuild.2018.06.006.

[59] Lisa Rivalin et al. “A comparison of methods for uncertainty and sensitivity analysis applied to the energy performance of new commercial buildings”. In: *Energy and Buildings* 166 (May 2018), pp. 489–504. ISSN: 03787788. DOI: 10.1016/j.enbuild.2018.02.021.

[60] Jingyi Wang et al. “The uncertainty of subjective thermal comfort measurement”. In: *Energy and Buildings* 181 (2018), pp. 38–49. ISSN: 03787788. DOI: 10.1016/j.enbuild.2018.09.041.

[61] ASHRAE. *Standard 90.1 – Energy Standard for Buildings Except Low-Rise Residential Buildings*. Atlanta, Georgia, US, 2019. URL: https://www.ashrae.org/technical-resources/bookstore/standard-90-1.

[62] U.S. Green Building Council. *LEED certification for new buildings*. (accessed Jun 22, 2020). 2019. URL: https://www.usgbc.org/leed/rating-systems/new-buildings.

[63] Martin Prignon et al. “A method to quantify uncertainties in airtightness measurements: Zero-flow and envelope pressure”. In: *Energy and Buildings* 188-189 (2019), pp. 12–24. ISSN: 0378-7788. DOI: https://doi.org/10.1016/j.enbuild.2019.02.006.

[64] CORE studio. *Design Explorer 2*. (access Dec 25, 2019). Thornton Tomasetti. 2019. URL: https://tt-acm.github.io/DesignExplorer/.

[65] National Renewable Energy Laboratory (NREL). *EnergyPlus | EnergyPlus*. (accessed Sep 30, 2020). 2020. URL: https://energyplus.net/.

[66] Yanmeng Chen. “A study of the assessment method of insultation and shading’s energy saving effect under climate change using simulations”. MA thesis. The Unversity of Tokyo, 2020.

[67] U.S. Green Building Council. *Why LEED certification*. (accessed Jun 22, 2020). 2020. URL: https://www.usgbc.org/leed/why-leed.

[68] eu.ESCO. *About eu.esco*. (accessed Jun 29, 2020). 2020. URL: https://www.euesco.org/about-eu.esco/index.html.

[69] Hannes Harter et al. “Uncertainty Analysis of Life Cycle Energy Assessment in Early Stages of Design”. In: *Energy and Buildings* 208 (2020), p. 109635. ISSN: 03787788. DOI: 10.1016/j.enbuild.2019.109635.

[70] TAKAGI SHUTA OFFICE LLC. *Wind Analysis 風環境解析 GH2FD | Building Environment Design.com*. (accessed Sep 30, 2020). 2020. URL: http://building-env.com/archives/2047.

[71] Wikipedia. *Typical meteorological year - Wikipedia*. (accessed Jun 29, 2020). Mar. 2020. URL: https://en.wikipedia.org/wiki/Typical_meteorological_year.

[72] Ying Zhang et al. “Study on model uncertainty of water source heat pump and impact on decision making”. In: *Energy and Buildings* 216 (2020), p. 109950. ISSN: 03787788. DOI: 10.1016/j.enbuild.2020.109950.

[73] Meteotest AG. *Features - Meteonorm (en)*. (accessed Dec 25, 2019). URL: https://meteonorm.com/en/meteonorm-features.

[74] The Intergovernmental Panel on Climate Change. *Fifth Assessment Report −IPCC*. (accessed Dec 25, 2019). URL: https://www.ipcc.ch/assessment-report/ar5/.

[75] The Intergovernmental Panel on Climate Change. *Fourth Assessment Report −IPCC*. (accessed Dec 25, 2019). URL: https://www.ipcc.ch/assessment-report/ar4/.

[76] Alexander G. de G. Matthews James Hensman. *Multi-output Gaussian processes in GPflow*. (accessed Sep 30, 2020). URL: https://gpflow.readthedocs.io/en/master/notebooks/advanced/multioutput.html.

[77] National Renewable Energy Laboratory. *OpenStudio | OpenStudio*. (accessed Dec 25, 2019). URL: https://www.openstudio.net/.

[78] ENERGY STAR. *What is energy use intensity (EUI)?* (accessed Jun 22, 2020). URL: https://www.energystar.gov/buildings/facility-owners-and-managers/existing-buildings/use-portfolio-manager/understand-metrics/what-energy.

[79] Wikipedia. *Curse of dimensionality - Wikipedia*. (accessed Aug 20, 2020). URL: https://en.wikipedia.org/wiki/Curse_of_dimensionality.

[80] Wikipedia. *HadCM3 - Wikipedia*. (accessed Dec 25, 2019). URL: https://en.wikipedia.org/wiki/HadCM3.

[81] Wikipedia. *k-means clustering - Wikipedia*. (accessed Dec 25, 2019). URL: https://en.wikipedia.org/wiki/K-means_clustering.

[82] Wikipedia. *Linear regression - Wikipedia*. (accessed Dec 25, 2019). URL: https://en.wikipedia.org/wiki/Linear_regression.

[83] Wikipedia. *Multivariate normal distribution - Wikipedia*. (accessed Dec 25, 2019). URL: https://en.wikipedia.org/wiki/Multivariate_normal_distribution#Conditional_distributions.

[84] Wikipedia. *Polygon mesh - Wikipedia*. (accessed Aug 20, 2020). URL: https://en.wikipedia.org/wiki/Polygon_mesh.

[85] Wikipedia. *Support vector machine - Wikipedia*. (accessed Dec 25, 2019). URL: https://en.wikipedia.org/wiki/Support_vector_machine.

Acknowledgements

I would like to, and should, express my sincere thanks to my advisor, Professor Mae, for all the guidance, all the advises, all the support and all the great opportunities he has provided. I should admit that it was not easy do a designer-oriented research in this architecture environment group. Without Prof. Mae's support and guidance, it would be impossible for me to insist on my research topic. I should also thank Prof. Mae for taking me to the workshops and conference in U.S., and organizing GGA, so that I could get contact with the top designers and researchers. I should also thank him for involving me into the YKK projects. The YKK Indonesian project has become the origin point of my research. Thanks to Prof. Mae's efforts, I could get the change to cooperate with AKL, Takenaka, and meet research from other universities and institutes. All of these would become the most precious experiences in my life. Also, I would like to express my thank to Prof. Mae for his regards in daily life.

I would like to pass my gratitude to Mr. Taniguchi for his advice on my research, as well as the encouragements from him when I was troubled by depression.

I would like also send my thanks to the seniors in the lab, Mr. Yasin and Miss. Prudsamon, for their help in my daily and study life. I would like to thank them for leading me into the Ph.D life.

I would like to convey my thanks to my lab mates. The discussions with them inspired my research. I really appreciate your interests on my research topic, which has been my precious source of sense of achievement in these years.

Last but not least, to my beloved, mother, father, cousins and all my family, for their backing of my studying aboard life.

And a special thank to my girlfriend, who helped to revise the spelling and grammar of my manuscripts.