RMIT UNIVERSITY

An artificial intelligence platform for design optimization and data analysis: Application for fire and ventilation problems

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

Nan Li

B.E. & M.E., Tianjin University, China

School of Engineering

College of Science, Engineering and Health

RMIT University

July 2019

Declaration

I certify that except where due acknowledgement has been made, the work is that of the author alone; the work has not been submitted previously, in whole or in part, to qualify for any other academic award; the content of the thesis is the result of work which has been carried out since the official commencement date of the approved research program; any editorial work, paid or unpaid, carried out by a third party is acknowledged; and, ethics procedures and guidelines have been followed.

I acknowledge the support I have received for my research through the provision of an Australian Government Research Training Program Scholarship.

Nan Li

10th July 2019

School of Engineering,
College of Science, Engineering and Health,
RMIT University

Acknowledgements

I would like to thank my primary supervisor Associate Professor Sherman Chi Pok Cheung, for his constant guidance and mentoring throughout my PhD study period. He is the best supervisor who has not only directed my academic program, but also gave me lots of advice on my personal life. I would also like to thank my other supervisors Professor Xiaodong Li and Professor Jiyuan Tu. Their expertise in their own research area have been always very helpful to my research program. Particularly, I am grateful to Professor Tu for his support of both my research and life.

My deepest thanks go to my grandparents for their countless love and care given to me. Finally, my deep and sincere gratitude to my wife, Mrs Run Cao, for her continuous love, help, understanding and support.

Publications during Candidature

Peer Reviewed Journal Publications:

1. **Li, N.,** Cheung, S. C., Li, X. & Tu, J. (2017). Multi-objective optimization of HVAC system using NSPSO and Kriging algorithms—A case study. *Building Simulation,* 10(5), 769-781. DOI: 10.1007/s12273-017-0352-5, ***IF = 1.673, Q1***

2. **Li, N.,** Lee, E.W.M., Cheung, S. C. & Tu, J. (2019). Multi-fidelity surrogate algorithm for fire origin determination in compartment fires. *Engineering with Computers*. DOI: 10.1007/s00366-019-00738-9, ***IF = 1.951, Q1***

3. **Li, N.,** Yang, L., Li, X., Tu, J. & Cheung, S. C. (2019). Multi-objective optimization design of high-speed train cabin ventilation system using particle swarm optimization and Multi-fidelity Kriging. *Building and Environment*, 155, 161-174. DOI: 10.1016/j.buildenv.2019.03.021, ***IF = 4.539, Q1***

4. **Li, N.,** Shang, Y., Tu, J. & Cheung, S. C. (2019). A new method to cut powerline wildfire risk caused by vegetation conduction ignition in Victoria, Australia, using Convolutional Neural Networks. *International Journal of Wildland Fire*. (under review). ***IF = 2.445, Q1***

5. Li, G., **Li, N.,** Wen, C. & Ding, S. (2018). Investigation and modeling of flank wear process of different PCD tools in cutting titanium alloy Ti6Al4V. *The International Journal of Advanced Manufacturing Technology,* 95(1-4), 719-733. DOI: 10.1007/s00170-017-1222-0, ***IF = 2.601, Q1***

6. Yi, S., **Li, N.,** Solanki, S., Mo, J. & Ding, S. (2019). Effects of graphene oxide nanofluids on cutting temperature and force in machining

Ti-6Al-4V. *The International Journal of Advanced Manufacturing Technology*.
DOI: 10.1007/s00170-019-03625-1, ***IF = 2.601, Q1***

7. Li, G., Yi, S., **Li, N.,** Wen, C. & Ding, S. (2019). Quantitative analysis of cooling and lubricating effects of graphene oxide nanofluids in machining titanium alloy Ti6Al4V. *Journal of Materials Processing Technology*.
   DOI: 10.1016/j.jmatprotec.2019.04.035, ***IF = 3.647, Q1***

Conference Publications:

1. **Li, N.,** Cheung, S. C., Li, X. & Tu, J. (2015). Multi-objective optimization of thermal comfort and energy consumption in a typical office room using CFD and NSM-PSO. *The 21st International Congress on Modelling and Simulation (MODSIM 2015)*, Gold Coast, Australia, November 2015.

2. **Li, N.,** Cheung, S. C., Li, X. & Tu, J. (2015). Development of a multi-objective design optimization platform using NSM-PSO and CFD for heating and ventilation applications. *The 11th International Conference on CFD in the Minerals and Process Industries*, Melbourne, Australia, December 2015.

Credits

The following publications have been used as portion of the materials in this thesis:

Chapter 3

Li, N., Cheung, S. C., Li, X. & Tu, J. (2017). Multi-objective optimization of HVAC system using NSPSO and Kriging algorithms—A case study. *Building Simulation,* 10(5), 769-781. ***IF = 1.673, Q1***

Li, N., Cheung, S. C., Li, X. & Tu, J. (2015). Multi-objective optimization of thermal comfort and energy consumption in a typical office room using CFD and NSM-PSO. *The 21st International Congress on Modelling and Simulation (MODSIM 2015)*, Gold Coast, Australia, November 2015.

Chapter 4

Li, N., Lee, E.W.M., Cheung, S. C. & Tu, J. (2019). Multi-fidelity surrogate algorithm for fire origin determination in compartment fires. *Engineering with Computers*. ***IF = 1.951, Q1***

Chapter 5

Li, N., Yang, L., Li, X., Tu, J. & Cheung, S. C. (2019). Multi-objective optimization design of high-speed train cabin ventilation system using particle swarm optimization and Multi-fidelity Kriging. *Building and Environment,* 155, 161-174. ***IF = 4.539, Q1***.

Chapter 6

Li, N., Shang, Y., Tu, J. & Cheung, S. C. (2019). A new method to cut powerline wildfire risk caused by vegetation conduction ignition in Victoria, Australia, using Convolutional Neural Networks. *International Journal of Wildland Fire*. (under review). ***IF = 2.445, Q1***

Contents

List of Figures

List of Tables

Nomenclature

Symbol definition

c_p	specific heat capacity of air (J kg^{-1} K^{-1})
D_n	crowding distance of n^{th} particle in PSO
D_φ	kinematic diffusivity (m^2 s^{-1})
E	energy consumption (W)
f_{cl}	basic clothing insulation
g	gravitational acceleration (m s^{-2})
h	specific enthalpy of air (J kg^{-1})
h_c	convection heat transfer coefficient (W m^{-2} K^{-1})
k_{air}	thermal conductivity of air (W m^{-1} K^{-1})
m	mass flow rate (kg s^{-1})
M	metabolic energy production (W s^{-1})
p_i	personal best position in PSO
p_g	global best position in PSO
P	air pressure (Pa)
P_a	vapour pressure of water in ambient air (Pa)
$\dot{Q}$	thermal load (W)
$S_\emptyset$	source term
t_a	ambient air temperature (K)
$\overline{t_s}$	mean radiant temperature (K)

t_{cl} average skin temperature (K)

u fluid velocity (m s^{-1})

V inlet velocity (m s^{-1})

W rate of mechanical work (W s^{-1})

x current position of particle in PSO

Z measurement process in Kriging

Subscripts

f function

g global

i input

l local

o output

PMV predicted mean vote

Greek symbols

ε random errors in Kriging process

ϕ scalar variable for governing equation

φ concentration of contaminant (kg m^{-3})

ρ regression coefficient in Multi-fidelity Kriging process

λ weighting coefficients in Kriging

Γ diffusion coefficient (m^2 s^{-1})

v moving velocity of particle in PSO

ω parameters of PSO

Abstract

This thesis focuses on the development of novel multi-objective software platforms to assist engineering design and investigation, especially for simulation-based indoor environment problems, which always involve multiple evaluation criteria. In addition, this thesis aims to develop new methods to reduce the computational cost associated with the design process.

In modern building design, engineers are constantly facing challenging to find an optimal design to maintain a high level of thermal comfort and indoor air quality for occupants while minimizing the system energy consumption. Over the past decades, several algorithms have been proposed and developed for optimizing the heating, ventilation and air conditioning (HVAC) system for indoor environment. Nevertheless, the majority of these optimization algorithms are focused on single objective optimization procedures and require a large training sample for surrogate modelling. For multi-objective HVAC design problems, previous studies introduced an arbitrary weighting factor to combine all design objectives into one single objective function. The near-optimal solutions were however sensitive to the chosen value of the weighting factor.

In another hand, the computational cost is very heavy in the computer-aided investigation process of reverse engineering problems. Computational Fluid Dynamics (CFD) aided fire investigation is one of the reverse engineering. With the significant growth of the world population, our cities are becoming more and more crowding. In this situation, any fire occurring would cause severe consequences, including property damage and human injuries or even deaths. In assessing the fire cause, the fire origin determination is a crucial step identifying the origin of fire outbreak and the sequential fire and smoke propagation. Traditionally, fire investigators relied upon the visible fire damages at the fire scene to determine the location of fire originated based on their own professional experience. The fire investigation process is however subject to the expert interpretation inherently embedded in the qualitative analyses.

In addition, we are living in an era of big data, where lots amount of data are generating every day, especially in engineering field. Traditional analysis methods are not suitable to handle large amount of data quickly and accurately. In contrast, new techniques such as machine learning are able to deal with big data and extract data features.

The main body of this thesis is composed of seven chapters, and the details of each chapter are as the followings:

The research background and a comprehensive literature review are described in the first two chapters where the research gaps found in the existing literatures are discussed. From Chapter 3 to Chapter 6, the main contributions of this research are demonstrated. In Chapter 3, a nondominated sorting-based particle swarm optimization (NSPSO) algorithm together with the Kriging method to perform optimization for the HVAC system design of a typical office room was developed. In addition, an adaptive sampling procedure was also introduced to enable the optimization platform to adjust the sampling point and resolution in constructing the training sample. Chapter 4 presents a Multi-fidelity Kriging algorithm to quantitatively determine the fire origin based on the soot deposition patterns predicted by the numerical simulations, which provides an unbiased and fast methodology to assist the fire investigation. A comprehensive multi-objective optimization platform of the ventilation system inside a typical high-speed train (HST) cabin is discussed in Chapter 5, where the NSPSO and the Multi-fidelity Kriging were combined together to reduce computational cost. Chapter 6 demonstrates a successful application of convolutional neural networks (CNN) in vegetation feature analysis to help cut powerline wildfire risk caused by vegetation conduction ignition. Finally, all the contributions in this research are summarised in Chapter 7.

Chapter 1
Introduction

1.1 Background and Motivation

In recent decades, with the development of the computer science, the computer-aided engineering (CAE) techniques have been widely applied to assist industrial design process. Traditionally, the design process is a cycle of trial and error which includes lots of experimental tests and thus consumes huge amount of resources. Nowadays, the usage of computational simulation techniques helps to shorten the design period by replacing the time-consuming experimental tests with relatively fast numerical simulations. However, the improvement of replacing experiments with simulations only is limited without changing the conventional trial and error design scheme. We need a smarter design scheme for modern engineering design to explore more solutions and to reduce design cost in the meanwhile.

In the modern indoor environment design process, computational fluid dynamics (CFD) techniques have been widely adopted to analyze air distribution and flow characteristics. In comparison to the traditional design cycle, CFD simulations offer a faster and more economical way for engineers to carry out parametric studies, leading towards a more desirable system design. However, the parametric analysis using CFD technique is inherently discrete in the design space with pre-selected design variable values, and the best optimum solution could be "hidden" in the discretized domain. In other words, the accuracy of a near-optimum solution depends largely on the "resolution" of the test matrix where significant computational costs are required. In addition, in traditional non-adaptive experiments, decisions such as how to sample during an experiment are made and fixed in advance. However, the main disadvantage of uniform sampling is high resource cost, because the sampling density must be uniformly high everywhere in order to meet the sampling requirements in some particular areas. On the other hand, most

of the heating, ventilation and air conditioning (HVAC) system design normally involves multi-objective considerations. Design indices such as predicted mean vote (PMV), percentage dissatisfied of draft (PD), age of air, CO_2 concentration and energy cost are commonly considered in literatures. It is challenging for traditional methods to handle multiple design criteria in an efficient way.

The numerical simulation technology makes it feasible to solve reverse engineering problem, such as the building fire investigation. Fire origin determination is a challenging task in forensic sciences owing to the complex transient fire behaviour and its inherent nature in destroying evidence. One of the key hypothesis of the fire investigator is the correct identification of the fire origin. For most of the fire scenarios, the determination of fire origin is crucial for an accurate and reliable fire cause assessment. Today, with the development of computational technology, the application of fire simulation for fire scene reconstruction has gained measured success in the past decades. Nonetheless, fire reconstruction studies are efficient in forward engineering processes where fire consequences are predicted based on known fire size and location. Fire origin determinations are unfortunately reserve engineering processes where fire consequences or damages are known except the fire size or locations. As each CFD simulation requires considerable computational time and resources, it is still impractical relying on CFD solely for fire investigation; especially for fire origin determination where a huge amount of simulations are required to support the reversed engineering process. An efficient way of creating simulation cases and using simulation results are required to make such a kind of CAE method to solve reverse engineering problems.

Apart from the indoor building fires, the wildland fires can pose significant consequences on human life, assets and ecologic environment. In the past decade, wildfires have caused huge losses of human lives and properties around the world including in Greece (Amiridis et al. 2012), the United States (Keeley et al. 2009), Russia (Vivchar 2011) and Australia (Chafer et al. 2004, Stephenson et al. 2013, Collins et al. 2015). Especially, in Australia, over the past century, 20% of the total building

losses in nature hazards arise from bushfires (Mcaneney et al. 2009). Understanding the ignition process of the wildfire is essential to reducing the risk of future wildland fires. Huge amounts of data have been collected from both historical fire field and laboratory experimental tests in purpose of analysing the reason of causing the wildland fire and avoiding the future occurrence of wildland fire. However, the traditional analysis methods are not suitable to handle large amount of data quickly and accurately. Developing new analysis methods is essential to help engineers better understand the ignition process in wildland fires and to reduce the risk of future wildland fires.

1.2 Research Gaps and Objectives

When designing a multi-objective optimization ventilation system, in most previous works (Laverge and Janssens 2013, Li et al. 2013), a single objective function was constructed by aggregating several design indices using pre-defined weighting factors. One particular disadvantage of this method is that the optimal solution could be sensitive to the values of the weighting factors. In other words, different values of weighting factors could result in substantially different solutions. The weighting factors must be therefore chosen carefully based on subjective factors such as engineering or expert judgements. Furthermore, the optimization procedure gives only one near-optimal solution where there is no flexibility for the designer to strike a balance or "trade-off" of the conflicting parameters. It is nearly impractical to explore all the possible solutions by trying different weights.

In addition, in constructing the sample data, most studies from previous literatures adopt uniform sampling where the CFD simulations are uniformly distributed within the design range (Zhou and Haghighat 2009a, Li et al. 2013). The main disadvantage of uniform sampling is its high cost, as the sampling density must be uniformly high everywhere in order to meet the sampling requirements in some particular areas. Nevertheless, in most practical cases, the final optimal solutions are normally concentrated in a certain region within the design space.

Therefore, substantial computational time could be wasted in constructing the sample data for some virtually redundant samples which are far away from the optimal solution.

On the other hand, in previous literatures, researchers have been struggled to strike a balance between the simulation accuracy and the computational cost (Lin et al. 2005, Lin et al. 2009, Zhou and Haghighat 2009a). For example, in most previous literatures, the simplified manikin models such as rectangle or cylinder blocks are used in CFD simulations to study the indoor environment (Lin et al. 2009, Yuan et al. 1999, Li et al. 2013). Using simplified manikin model is unable to get accurate CFD predictions, especially when we are targeting on flow zones near the manikin. A few of other researchers applied high-resolution 3D scanned thermal manikin models to improve the CFD accuracy. However, it would dramatically increase the cost of computational resources at the same time. Although the usage of parallel computing could help to reduce the total computational time, it still poses a significant burden to the computational resources as multiple CPU cores have to be occupied to do parallel computing. There is a lack of a method which is capable to save computational cost without sacrificing the simulation accuracy. In addition, a data analysis tool which can handle a large amount of data quickly and accurately, thus can be automated and industrialized, is yet to be developed.

In attempting to fill the aforementioned research gaps found in literatures, the specific objectives of this thesis are:

- To develop a new design scheme to achieve multi-objective optimization without having to use any weighting factors, and is capable to explore the whole design space, to provide multiple solutions.
- To develop a response-adaptive sampling approach, which allocates the sampling points in an adaptive way to save computational resources.

- To develop a new surrogate model of CFD simulation, which is able to save computational cost without sacrificing the simulation accuracy.
- To develop a new data analysis tool, which is able to handle a large amount of data quickly and accurately, by using machine learning technology.

1.3 Thesis Outline

This thesis is composed of seven chapters. The topic of each chapter is outlined below:

Chapter 2 provides a comprehensive literature review in relation to the existing CFD studies, especially in the area of indoor ventilation system design. Different surrogate models and optimization schemes are reviewed in this chapter. The drawbacks of reviewed methods are also discussed. The reviewed literature lays a solid foundation for the research outcomes in the following chapters.

Chapter 3 presents a nondominated sorting-based particle swarm optimization (NSPSO) algorithm together with the Kriging method to perform optimization for the HVAC system design of a typical office room. In addition, an adaptive sampling procedure was also introduced to enable the optimization platform to adjust the sampling point and resolution in constructing the training sample.

Chapter 4 presents a Multi-fidelity Kriging algorithm to quantitatively determine the fire origin based on the soot deposition patterns predicted by the numerical simulations, which provides an unbiased and fast methodology to assist the fire investigation.

Chapter 5 illustrates a comprehensive multi-objective optimization platform of the ventilation system inside a typical high-speed train (HST) cabin, where the NSPSO and the Multi-fidelity Kriging were combined together to reduce computational cost.

Chapter 6 demonstrates a successful application of convolutional neural networks (CNN) in vegetation feature analysis to help cut powerline wildfire risk caused by vegetation conduction ignition.

Chapter 7 concludes and highlights all the conclusions and contributions.

1.4 Contribution

Based on the research gaps and research objectives discussed before, this thesis contributes to the following major outcomes:

a) Developed a new multi-objective optimization design scheme for HVAC system design, which is free of using weighting factors, and is capable to explore the whole design space, to provide multiple solutions.
b) Developed an adaptive CFD sampling approach, which is able to allocate the sampling points in an adaptive way according to the distribution of optimal solutions, helping to save computational resources dramatically.
c) Developed a multi-fidelity surrogate model of CFD simulation, which is able to save computational cost without sacrificing the simulation accuracy.
d) Developed a new data analysis tool, which is capable to handle a large amount of data quickly and accurately, by using machine learning technology.

Chapter 2
Literature Review

This chapter presents a review of the techniques and methodologies in the literatures related to this thesis. The review covers three main aspects:

1) Indoor ventilation system design. In this section, the technology and approaches appeared in current literatures related to CFD-assisted indoor ventilation system design is described. The review includes the validation of CFD simulation, the surrogate methods of CFD simulations, and the assessments of thermal comfort and indoor air quality.
2) Evolutionary algorithms development. In this section, the development history of evolutionary algorithms (EAs) is reviewed. The algorithms include genetic algorithm (GA), particle swarm optimization (PSO), nondominated sorting genetic algorithm (NSGA) and nondominated sorting particle swarm optimization (NSPSO).
3) Machine learning techniques. In this section, the machine learning techniques used for image classification and recognition are demonstrated.

2.1 Indoor ventilation system design

2.1.1 CFD modelling and validation

In the modern indoor environment design, to achieve a higher system performance, computational fluid dynamics (CFD) simulation tools such as ANSYS Fluent, StarCCM+ and OpenFOAM have been widely adopted to analyze air distribution and flow characteristics and their relationship in thermal comfort and energy consumption (Ravikumar and Prakash 2009,

Cardinale et al. 2010, Hiyama et al. 2010, Kochetov et al. 2015, Gangisetti et al. 2016).

Having been developed for more than 70 years, the computational fluid dynamic (CFD) has been accepted as a reliable tool to solve and analyze problems that involve fluid flows, as long as the simulation model is well generated (including geometry, mesh, boundary conditions, turbulent model, etc.). CFD is a branch of fluid mechanics that uses numerical methods to solve the governing equations of fluid dynamics. These equations called Navier–Stokes equations (shown in Figure 2.1 (Tu et al. 2008)) include mass conservation, momentum equation, energy equation and additional equations for turbulent flow.

Mass Conservation

$$\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}+\frac{\partial w}{\partial z}=0$$

Momentum Equations

$$\frac{\partial u}{\partial t}+u\frac{\partial u}{\partial x}+v\frac{\partial u}{\partial y}+w\frac{\partial u}{\partial z}=\frac{\mu}{\rho}(\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}+\frac{\partial^2 u}{\partial z^2})\ -\frac{1}{\rho}\frac{\partial p}{\partial x}$$

$$\frac{\partial v}{\partial t}+u\frac{\partial v}{\partial x}+v\frac{\partial v}{\partial y}+w\frac{\partial v}{\partial z}=\frac{\mu}{\rho}(\frac{\partial^2 v}{\partial x^2}+\frac{\partial^2 v}{\partial y^2}+\frac{\partial^2 v}{\partial z^2})\ -\frac{1}{\rho}\frac{\partial p}{\partial y}$$

$$\frac{\partial w}{\partial t}+u\frac{\partial w}{\partial x}+v\frac{\partial w}{\partial y}+w\frac{\partial w}{\partial z}=\frac{\mu}{\rho}(\frac{\partial^2 w}{\partial x^2}+\frac{\partial^2 w}{\partial y^2}+\frac{\partial^2 w}{\partial z^2})\ -\frac{1}{\rho}\frac{\partial p}{\partial z}$$

Energy Equation

$$\frac{\partial T}{\partial t}+u\frac{\partial T}{\partial x}+v\frac{\partial T}{\partial y}+w\frac{\partial T}{\partial z}=\frac{\lambda}{\rho C_p}\left(\frac{\partial^2 T}{\partial x^2}+\frac{\partial^2 T}{\partial y^2}+\frac{\partial^2 T}{\partial z^2}\right)+S_T$$

Turbulence Equations

$$\frac{\partial(\rho k)}{\partial t}+\frac{\partial}{\partial x_i}(\rho k u_i)=\frac{\partial}{\partial x_j}\left(\alpha_k \mu_{eff}\frac{\partial k}{\partial x_j}\right)+G_k+G_b-\rho\varepsilon-Y_M+S_k$$

$$\frac{\partial(\rho\varepsilon)}{\partial t}+\frac{\partial(\rho\varepsilon u_i)}{\partial x_i}=\frac{\partial}{\partial x_j}\left(\alpha_\varepsilon \mu_{eff}\frac{\partial \varepsilon}{\partial x_j}\right)+C_{1\varepsilon}\frac{\varepsilon}{k}(G_k+C_{3\varepsilon}G_b)-C_{2\varepsilon}\rho\frac{\varepsilon^2}{k}-R_\varepsilon+S_\varepsilon$$

Figure 2.1 The Navier–Stokes equations for incompressible flow (Tu 2008).

Where u, v, w are the local velocity components; t is the time; ρ is the fluid density; p is pressure; Γ is the general diffusion coefficient; T is the fluid temperature; λ is the thermal conductivity; C_p is the thermal capacity; S_T is the internal thermal source; $S_\emptyset$ is the source term; G_k represents the generation of turbulent kinetic energy due to the mean velocity gradients, and G_b is the generation of turbulent kinetic energy due to buoyancy.μ_{eff} is the effective viscosity, and μ_t is the turbulent viscosity. The quantities α_k and α_ε are the inverse effective Prandtl numbers Pr for k and ε, respectively, $\alpha_k = \alpha_\varepsilon \approx 1.393$. The model constants values $C_{1\varepsilon} = 0.0845$, $C_{2\varepsilon} = 1.42$ and $C_{3\varepsilon} = 1.68$ are derived based on the analytically of the RNG theory. Historically, the first two-dimensional (2D) computational methods were developed in the 1930s using conformal transformations of the flow about a cylinder to the flow about an airfoil. One of the earliest types of calculations resembling modern CFD is those by Lewis Fry Richardson (Richardson 1922), in the sense that these calculations used finite differences and divided the physical space into cells. After the 1950s, the computer power available paced development of three-dimensional methods. The first paper with three-dimensional model was published by John Hess and A.M.O. Smith of Douglas Aircraft in 1967 (Hess and Smith 1967). More recently, some commercial CFD software packages have made the application of this numerical method much easier and more reliable.

Typically, in the field of heat and mass transfer in indoor environment, Yuan et al. 1999 published their experimental measurement data and computational simulation results of room airflow in 1999. In their experiment test, they built a well-insulated chamber divided into two parts, a test chamber and a climate chamber (shown in Figure 2.2), and each chamber has a separate heating, ventilation and air conditioning (HVAC) system. They used a flow visualization system to observe airflow patterns, a hot-sphere anemometer system to measure air velocity and a thermocouple system to measure surface and air temperature.

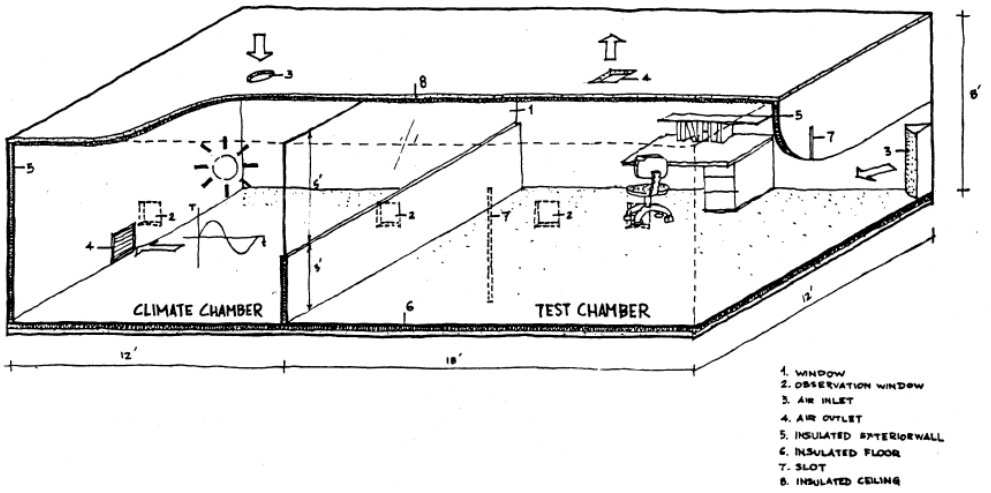

Figure 2.2 Sketch of the test facility in literature (Yuan et al. 1999).

For comparison, Yuan et al. also conducted a computational model using a commercial CFD code. The RNG k-ε turbulent model was used and the residual target for convergence criteria was set to be 10^{-3}. By limitation of computer performance in the 1990s, the mesh was coarse with only 48*44*24 grids. But they still got good enough results compared with their experimental data (see Figure 2.3 and Figure 2.4).

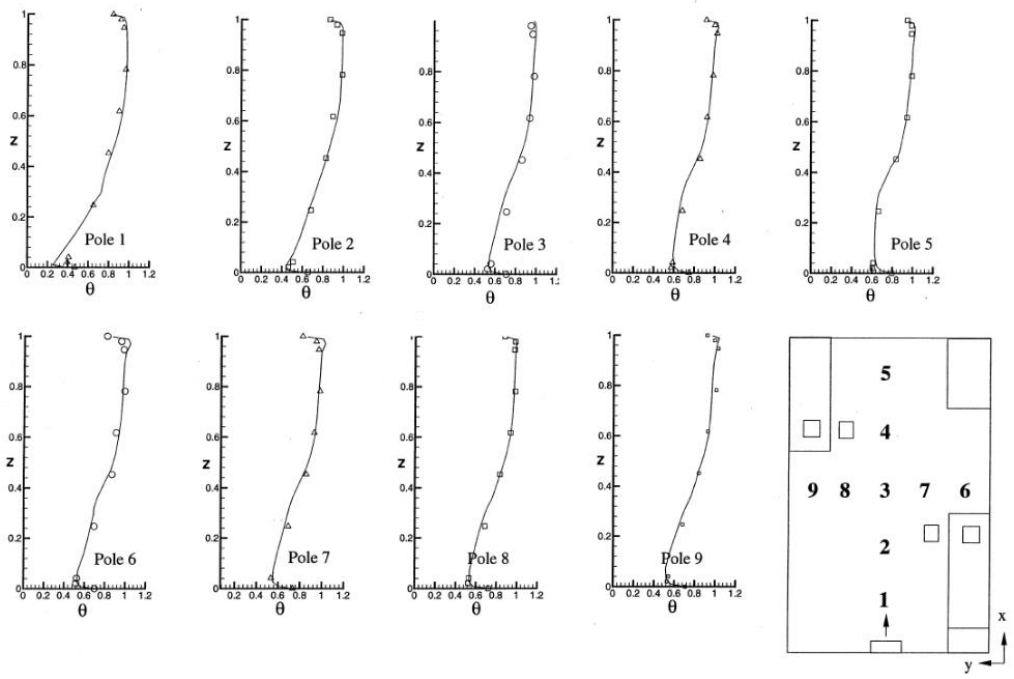

Figure 2.3 Comparison of temperature profile in a small office (Yuan et al. 1999).

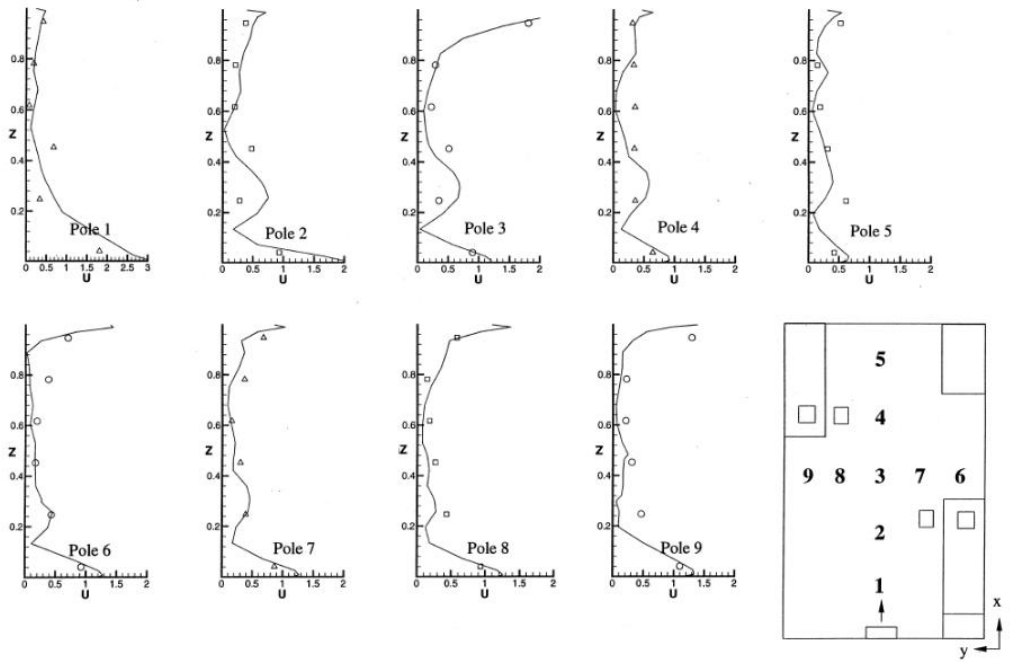

Figure 2.4 Comparison of velocity profile in a small office (Yuan et al. 1999).

After yuan et al.'s study, a lot of researchers used their experimental data to validate CFD simulation models. In addition, there are other researchers using CFD to help their study works. Zhang Lin et al. studied

the effect of internal partitions on the performance (thermal comfort and indoor air quality) of under floor air supply ventilation in a typical office environment using CFD numerical modeling (Lin et al. 2009), and Qingyan Chen et al. applied both experimental and CFD method to study ventilation performance prediction for buildings (Chen et al. 2010). All of these studies indicate that the CFD is a reliable tool for fluid flow analysis.

From the literature, the CFD can be used not only for prediction, but also can be used for optimization (Turner and Awbi 2015, Wang et al. 2013, Santos et al. 2012, Lee et al. 2011). However, the CFD optimization requires lots of professional experience to determine the potential optimal design area, otherwise the optimization process would be great uncertainly or blindly and cost a lot of time. With the help of some evolutionary algorithms, the optimization process can be guided and the designers would find the interesting area very quickly, even though without previous experience.

2.1.2 **Surrogate methods of CFD**

In comparison to the traditional design cycle, CFD simulations offer a faster and more economical way for engineers to carry out parametric studies, leading towards a more desirable system design. Albeit a near-optimum solution could be obtained, the parametric analysis using CFD technique is inherently discrete in the design space with pre-selected design variable values. The best optimum solution could be "hidden" in the discretized domain (Stavrakakis et al. 2011). Furthermore, the accuracy of a near-optimum solution depends largely on the "resolution" of the test matrix where significant computational costs are required. To enhance the accuracy of the near-optimum solution with practical computational time and resource, Artificial neural network (ANN) (Varol et al. 2007, Zhou and Haghighat 2009a,b, Stavrakakis et al. 2011, Lu and Xue 2014) or other surrogate techniques (Li et al. 2012, Li et al. 2013) are employed as an alternative approach to approximate the nonlinearity and complex behaviour of the multidimensional systems. One of the first studies using CFD-ANN coupled approach to assess the effect of architectural-designs on the thermal comfort can be found in (Krauss et

al. 1997). In general, numerical results predicted by CFD modelling were adopted to establish a database for training the ANN or surrogate models. These CFD-trained ANN or surrogate models then captured the relationship between design parameters and objective function. Optimization procedures using gradient methods (Gyulai et al. 2007, Stavrakakis et al. 2011, Welle et al. 2011) or evolutionary optimization algorithms (Luh and Lin 2011, Li et al. 2013, Afrand et al. 2015, Zhai et al. 2014) were then performed using the trained models for allocating the near-optimum solution within the continuous design space. Although significant computational time and resource can be reduced, a considerably large amount of CFD results are still required for constructing a reliable database for the training of ANN or surrogate models.

On the other hand, most of the HVAC system design normally involves multi-objective considerations. Design indices such as predicted mean vote (PMV), percentage dissatisfied of draft (PD), age of air, CO2 concentration and energy cost are commonly considered in literatures. Especially, in terms of indoor thermal comfort evaluation, substantial research works have done by Ricciardi's group (Buratti and Ricciardi 2009, Buratti et al. 2013, Nematchoua et al. 2014, Ricciardi and Buratti 2015, Buratti et al. 2016, Ricciardi et al. 2016). In most previous works, a single objective function was constructed by aggregating several design indices using pre-defined weighting factors (Laverge and Janssens 2013, Li et al. 2013). One particular disadvantages of this method is that the optimal solution could be sensitive to the values of the weighting factors. In other words, different values of weighting factors could result in substantially different solutions. The weighting factors must be therefore chosen carefully based on subjective factors such as engineering or expert judgements. Furthermore, the optimization procedure gives only one near optimal solution where there is no flexibility for the designer to strike a balance or "trade-off" of the conflicting parameters. For example, a lower indoor temperature may be preferable in summer which is in conflict with the goal to minimize energy consumption.

2.1.3 **Assessment of indoor thermal comfort**

According to the American Society of Heating Refrigeration and Air Conditioning Engineers (ASHRAE), the definition of human thermal comfort was expressed as the state of people's mind that expresses satisfaction with the surrounding environment (Ashrae 2004). The most important factors affecting the thermal comfort depend on four environmental variables, which are: the air temperature, the mean radiant temperature, the relative air velocity, and the relative humidity.

In recent year, the evaluation of the thermal comfort inside transportation cabins has been studied by many researchers. Currently, there are several indices that have been used for evaluating thermal comfort by international standards. One of these indices is the Predicted Mean Value (PMV) index (Gilani et al. 2015), which predicts the response of the thermal vote of a large group of people exposed to the same environment. Even though the PMV presents the empirical fit to human sensation, it has been developed based on the mathematical formulation proposed by Fanger (Fanger 1972). In the Fanger's mathematical model, it applies an energy balance for a human using the different energy exchange mechanisms, and experimentally derived physiological parameters. Table 2.1 describes the seven-point thermal sensation scale was used by Fanger based on the PMV index value.

Table 2.1 Thermal sensation scale used by (Fanger 1972).

Sign	Value	Description
+	3	Hot
+	2	Warm
+	1	Slightly warm
	0	Neutral
-	1	Slightly cool
-	2	Cool
-	3	Cold

Another proposed index is the Predicted Percentage Dissatisfied (PPD), which calculates a prediction of the number of thermally dissatisfied people (Gilani et al. 2015). The PPD is used for predicting the number of

people that feel uncomfortable warm or cold. It is a quantitative prediction of the percentage of thermally dissatisfied people based on PMV. Figure 2.5 illustrates the relationship between PMV and PPD. As can see, when the PMV index is near 0, the PPD value is also near 0 correspondingly. With the PMV changes to +3 (hot) or -3 (cold), the PPD values also exponential growth into 100%.

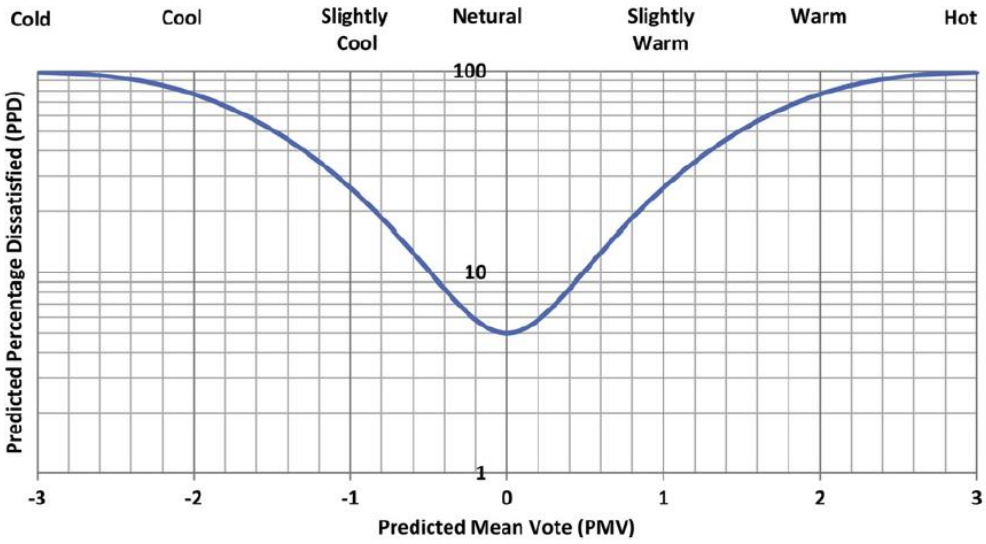

Figure 2.5 The Predicted Percentage of Dissatisfied (PPD) persons as a function of the predicted mean vote (PMV) index (Gilani et al. 2015).

According to ASHRAE (Ashrae 2004), draft which is unwanted local cooling due to air movement can cause occupant dissatisfaction in indoor spaces. The sensation of the draft depends on the speed of air, the fluctuating of air speed (turbulence), the air temperature and clothing conditions. Sensitivity to draft is most likely to be feel on the skin where is not covered by clothing, especially the neck, head, shoulders, ankle, feet and legs. Draft could increase with increasing air speed and turbulence intensity, meanwhile could decrease with increasing air temperature (Liu et al. 2017). Even though both PMV and PPD have considering the temperature, relative velocity and humidity, the influence of turbulence intensity is not included in the formula of neither PMV nor PPD. Thus, the percentage of dissatisfied (PD) due to draft model was applied for the thermal comfort analysis (Gilani et al. 2015).

Current thermal comfort standards, such as ASHRAE Standard 55-2004 (Ashrae 2004) and ISO 7730 (Iso 2005), defines a metric which quantifies draft in terms of what percentage of occupants will be dissatisfied with a space due to annoyance by the local draft conditions. The model was proposed by (Fanger 1972), which developed by curve-fitting experimental data utilizing a simple empirical model of human skin heat transfer. It expresses the percentage dissatisfied (PD) with the draft as a function of convective heat loss. The model relates air speed, temperature, and turbulence intensity to the percentage dissatisfied with air movement around people. According to ASHRAE Standard 55-2004 (Ashrae 2004), the PD should be < 20% for a comfort human occupancy environment.

For evaluating the PD value in a specific environment, the measurement locations are to be defined with respect to the occupant (Ashrae 2004). A minimum of three spatial locations are required to evaluate an occupant's environment: head area, waist area and ankle area. For a seated occupant, which is most common in HST cabin environment, the assessment levels are recommended as 1.1, 0.6 and 0.1 m above the floor.

2.1.4 **Assessment of indoor air quality**

During the travel, passengers are long time stay inside the HSR cabin. To maintain the good air quality inside the cabin is a critical issue. Air contaminants inside the relative closed cabin environment would downgrade indoor air quality (IAQ) and affect passengers' health and comfort (Wolkoff 1991, Franklin 2007). According to the existing studies (Kelly and Fussell 2015, Austin et al. 2002), the indoor air quality could be affected by a wide range of contaminants, such as smoke, volatile organic compounds (VOCs), biological particles, pollen, virus etc. As classified by previous researcher (Austin et al. 2002), the current found indoor contaminants can be mainly summarised into two categories: gaseous contaminants and particulate contaminants (which includes biological contaminants). Among the currently found indoor contaminants, some contaminates such as the virus can directly affect passengers' heathy, while some contaminants such as body odour are nontoxic but could

cause unpleasant sensory and mental distractions (Wolkoff 2013). In this section, the airborne contaminates were discussed from both gaseous contaminates and particle contaminates.

Gaseous contaminants are the smallest and most common type of cabin airborne pollution, which include hundreds of volatile organic compounds (VOCs) as well as many inorganic gaseous contaminants. Compared with the other category of indoor air contaminants, gaseous contaminants have distinctive properties related to their small sizes. They are existing in the form of individual molecules like oxygen and nitrogen in the air environment. Most gaseous molecules have a size between $3.4 - 20 \times 10^{-10}$ meters. The molecules of gaseous contaminants have fast thermal motion speed in random directions. The dynamic movement of indoor gaseous contaminants are governed by convection and diffusion laws.

Considering the composed of gaseous contaminants, two sub-categories can be further classified: inorganic gaseous contaminants and volatile organic compounds (VOCs). The inorganic compounds are defined as those molecules that lack carbon atom while organic compounds contain carbon bonds in which at least one carbon atom is covalently linked to an atom of another type (commonly hydrogen, oxygen, or nitrogen). However, in a normal indoor environment such as HSR cabin environment, the VOCs are the predominant gaseous contaminants. This is because the inorganic gaseous contaminants such as CO and H_2S are rare to be detected. Thus, the gaseous contaminants in this study would mainly focus on the VOCs.

From a previous study (Mirkhani et al. 2012), there are more than five hundred kinds of gaseous contaminants in the indoor environment. However, the total concentration of these gaseous contaminants in the indoor environment is usually very small. In a normal indoor environment, the concentration of VOCs is less than 1 milligram per cubic metre. Even the concentration of VOCs in the indoor environment is very small, there are many sources can release VOCs, such as the material (carpet, composite wood, upholstery fabrics, vinyl floors et al.), appliances

and personal care products (cooked food, fuel oil, Air fresheners et al.) and people. Among these VOCs released sources, the new materials may release VOCs, but the emission rate will decline very quickly over time. For old materials, the emissions of VOCs could almost negligible. The appliances and personal care products only release VOCs when in use. But for passengers, the VOCs release rate is continuous during all the travel. Therefore, in the HST cabin environment, the VOCs are mainly released from passengers.

For our human beings, the formation of VOCs (body odour) is mainly caused by skin glands excretions and bacterial activity (Lundström and Olsson 2010). Table 2.2, present the type of VOCs compounds in different places of the human body.

Table 2.2 VOCs compounds in different places of human body (Lundström and Olsson 2010)

Part of body	Examples of Compounds
Human skin	Lactic acid, aliphaticfatty acids, butanal, 3-methylbutanal, 2-methylbutanal, pentanal, hexanal, heptanal, octanal, phenylacetaldehyde, nonanal, decanal, andundecanal.
Foot	Acetic acid, butyric acid, isobutyric acid, isovaleric acid, propionic acid, valeric acid, and isocaproic acid
Human hair and scalp	Alkanes, alkenes, alcohols, aldehydes, ketones, acids, and 3' -lactone
Human axillary sweat and sebum	Esters (ethyl-2-methylpropanoate and ethylbutanoate), ketones(1-hexen-3-one and 1-octen-3-one) and, particularly, aldehydes[(Z)-4-heptenal, octanal, (E)-2-octenal, methional, (Z)-2-nonenal, (E,Z)-2,6-nonadienal, (E,Z)-2,4-nonadienal, (E,E)-2,4-decadienal, and 4 methoxybenzaldehyde]
Breath (mouth air)	Acetone, isoprene, ammonia, ethanol, acetaldehyde, ethylene, trimethyl amine, RSCs (carbonyl sulfide, hydrogen sulfide, methane thiol, and dimethyl sulfide), Saturated VOCs up to C14 + unsaturated C7-C10 + ketomonocarboxylic acids (C6-C10) + aldehydes
Urine	2-butanone, 2-pentanone, 4-heptanone, dimethyl disulfide, alkyl furans, pyrrole, carvone, benzaldehyde, p-cresol, phenol, trimethyl amine, 3-hydrobutyric acid and acetone

According to the study by Batterman and Peng (2010), human beings release VOCs in the form of body odor at a rate of 14.8 mg per hour per person. Considering the cabin environments are usually very crowded and relative sealed, this VOCs released rate could be a significant quantity for a long travel period. The high concentration of body odour might

cause health and comfort issues thus needs to be monitored and controlled (Wolkoff 2013, Franklin 2007).

Due to small sizes of molecule level, the transport behaviours of VOCs are quite different from the other type contaminants in a ventilated indoor environment. Compared with the surrounding air, the VOCs molecules have the same physics laws and mathematical equations. Since this reason, the molecules of VOCs can easily reach human's lung via respiratory airways. Meanwhile, the VOCs may have a chemical reaction with cells on the mucous membrane thus cause the sense of smell. The transport of VOCs in the indoor environment is influenced by both convections along with carrying airflow and interaction with surrounding air molecules.

Particulate contaminants are an inanimate tiny grain of mass such as tobacco smoke, oil smoke and dust while biological contaminants include virus, bacteria, mould, pollen etc. Unlike gaseous contaminants, the sizes of particulate and biological contaminants vary from 0.001 μm to 1000 μm, which has nearly six orders of magnitude difference (Howie 1990, Escombe et al. 2007). Relative size chart of common air contaminants has been plotted in Figure 2.6. Under such a large diverse range of size distribution, human's nasal filtration system is difficult to filter out all the harmful contaminants.

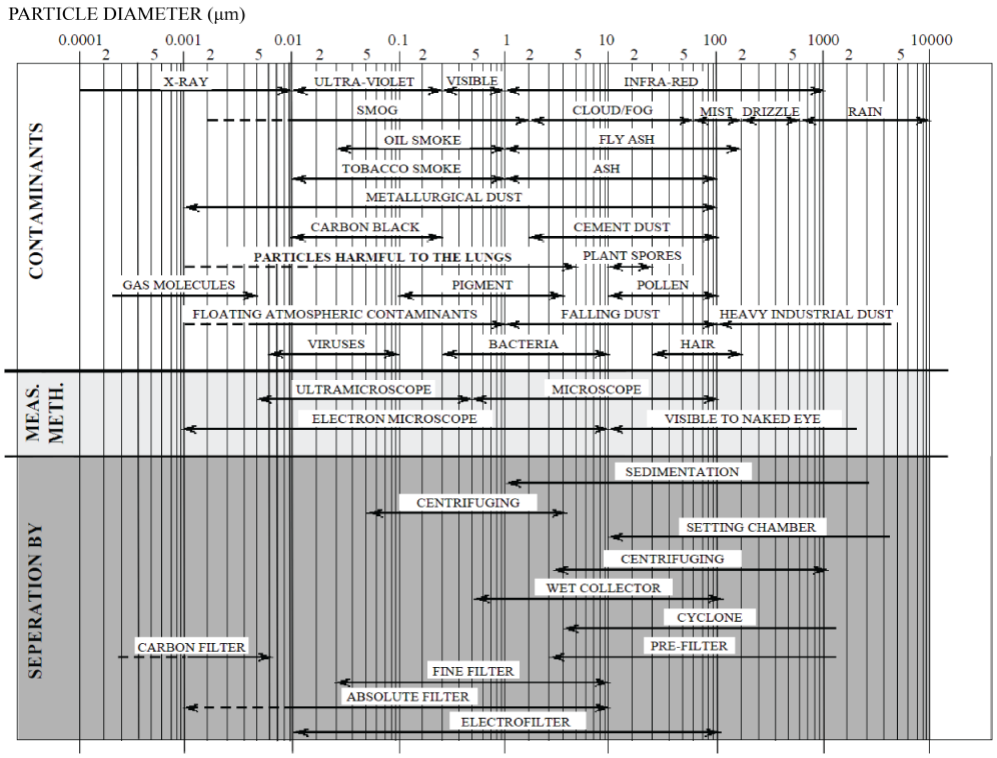

Figure 2.6 Relative size chart of common air contaminants (Howie 1990).

To identify the critical size range of harmful particulate contaminants, a number of existing studies carefully investigated the particle deposition rate in human's nasal cavity (Hsu and Chuang 2012, Kelly et al. 2004). The deposition efficiency of the inhaled articles measured have been summarised in Figure 2.7 (Kelly et al. 2004, Hsu and Chuang 2012). It can be noticed from Figure 2.7 that 80% of the particles smaller than 1 nm and larger than 10 μm have been filtered be human's nasal system. However, for the particle between100 nm and 5 μm, they have very low deposition efficiency when passing through the nasal system. These range of particles were also identified as PM2.5 (Chen et al. 2016), which are more dangerous than other particles due to their extremely low deposition rate in human's nasal cavity. Meanwhile, a few harmful contaminants are showing in this range such as viruses, bacteria, etc.

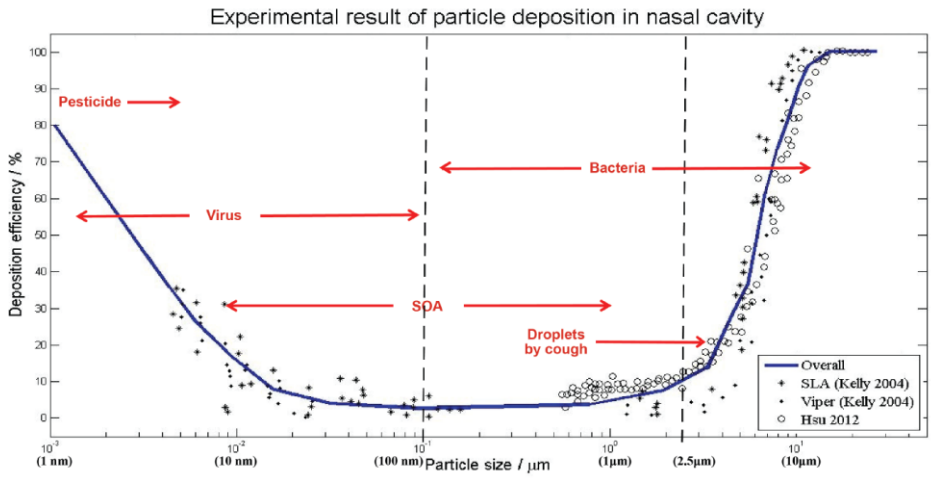

Figure 2.7 Particle deposition in nasal cavity (Kelly et al. 2004, Hsu and Chuang 2012).

According to previous study (Yan et al. 2014), the infectious virus released through coughing or sneezing of passengers is the main reason of disease transmission in public transport. Based on the experiment by Chao et al. (2009), the mean diameter of contaminants from coughing was 13.5 μm with average release speed of 11.7 m/s. Within half a second, most sputum droplets would quickly evaporate and become droplet nuclei with an average diameter of 3.5 microns (Redrow et al. 2011). The size evaporated droplet is in the range of PM 2.5 which has low deposit rate in human's nasal cavity. The perniciousness of the contaminants from coughing has attracted many attentions, especially after previous global outbreaks of diseases including SARS and H1N1. Gupta et al. (2011) numerically and experimentally investigated the distribution of contaminants released from coughing and breathing. It was found that the contaminants have similar transport characteristics.

Compared with the gaseous contaminates, particulate contaminants are much larger. Thus, the mathematical equations for gaseous contaminants cannot be applied for particulate contaminants. Li et al. (2014b) experimentally measured both gaseous and particulate contaminants transport behaviour. They found that the gaseous contaminants were primarily affected by the airflow, while the particle contaminants could suspend in the air or be carried by indoor airflow, which was affected by more factors such as the diameter of the particle and the release speed.

2.2 Evolutionary algorithms development

In artificial intelligence, an evolutionary algorithm (EA) is a subset of evolutionary computation, a generic population-based metaheuristic optimization algorithm, which uses mechanisms inspired by biological evolution, such as reproduction, mutation, recombination, and selection. Started from last century, a variety of EAs have been developed. Among them, Genetic Algorithm (GA) and Particle Swarm Optimization (PSO) are the most popular and widely used.

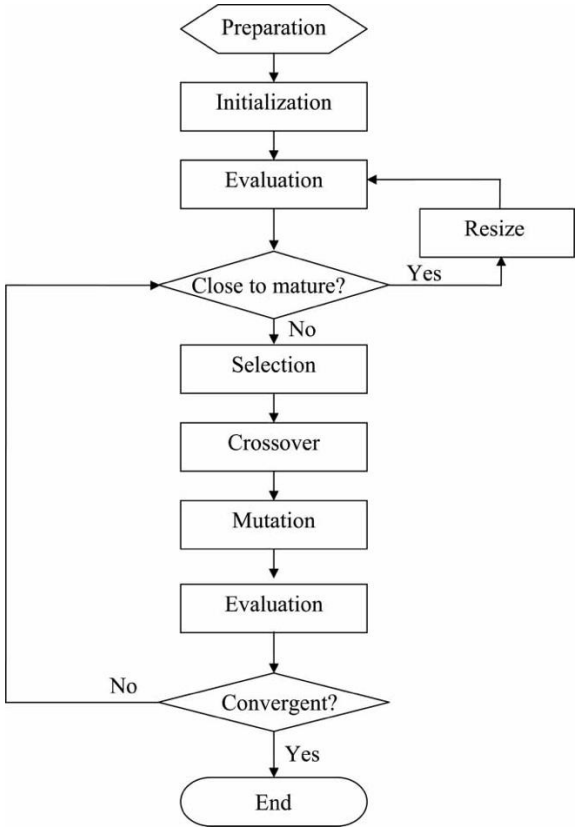

Figure 2.8 A typical flow chart of GA.

Genetic Algorithm is a search heuristic that mimics the process of natural selection. Some early pioneers including Fraser, Burnell (Fraser et al. 1970), Crosby (Crosby and L 1973) and Hans-Joachim Bremermann made great contribution to the modern GA, providing essential elements of GA like recombination, mutation, and selection. Figure 2.8 shows a typical

work flow of GA. In the process of GA, different crossover operators (single-point crossover, k-point crossover, uniform crossover, partially matched crossover, etc.) and/or mutation operators (bit string mutation, flip bit, boundary mutation, Gaussian mutation, shrink mutation, etc.) could be applied to solve specific problems. Although compared to classic gradient-based optimization, the heuristic-based searched techniques such as GA have some advantages in solving non-convex optimization problems, there are still limitations when GA is used to solve real-world engineering problems, which naturally involve multiple conflicting objectives. An example of real-world multiple conflicting objectives is shown in Figure 2.9 (Deb 2001). It clearly demonstrates the conflicting factors when buying a vehicle between price and comfort. One must choose a trade-off balance, which means more comfort costs more money. In addition, there exist multiple solutions when considering multiple objectives, which are suitable to different users.

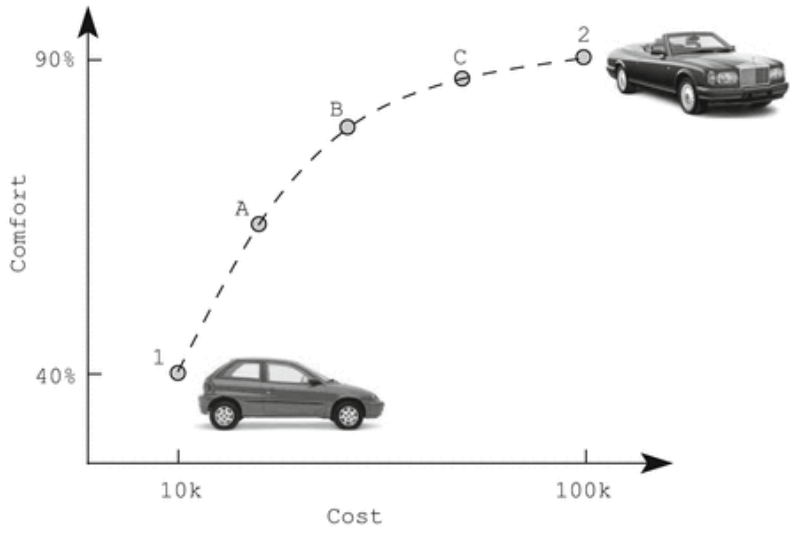

Figure 2.9 An example of multiple conflicting objectives when choosing a car.

In order to handle multi-objective optimization problems (MOPs), in early 1990s, Fonseca et al. 1993, Horn et al. 1994 and Srinivas et al. 1994 proposed a number of methods which could find multiple Pareto-optimal solutions in one single simulation run for solving MOPs. Among them, there are two popular methods which have been widely adopted by lots of

later researchers: niching approach (Horn et al. 1994) and nondominated sorting method (Srinivas and Deb 1994). Especially, in literature (Srinivas and Deb 1994), a new concept was suggested – nondominated sorting. Based on nondominated sorting method, a new population diversity-preservation mechanism (NSGA) instead of niching method was proposed in (Srinivas and Deb 1994). In 2002, Deb et al. improved the NSGA in three aspects (computational complexity, elitism and parameter-less diversity-preservation) and called the new algorithm NSGA-II. This is a significant contribution to multi-objective evolutionary algorithms (MOEAs) and their paper has been cited over 14,000 times.

Particle Swarm Optimization is another popular EA. Originally introduced by Kennedy and Eberhart (Kennedy 2001), the PSO has been developed by a large number of researchers. Maurice Clerc and James Kennedy studied the stability and convergence of PSO (Clerc and Kennedy 2002). In order to solve discrete optimization problem, Kennedy et al. introduced a binary version of PSO (Kennedy and Eberhart 1997). Following Kennedy's work, Khanesar et al. proposed a novel binary PSO which suggested that the velocity of a particle is its probability to change its state from its previous state to its complement value, rather than the probability of change to 1, making the original binary PSO find better results. Inspired by NSGA-II (Deb et al. 2002), Li made a contribution to multi-objective PSO by introducing nondominated sorting method into PSO and proposed a NSPSO in his paper (Li 2003b), and tested the algorithm using some classic benchmark test problems.

Rania et al compared the performance of PSO with GA in (Hassan et al. 2005) and concluded that PSO has the same effectiveness (finding the true global optimal solution) as the GA but with significantly better computational efficiency (fewer function evaluations) by implementing statistical analysis and formal hypothesis testing. It indicates that PSO converges faster than GA. Therefore in my PhD research, I will mainly focus on the application of NSPSO to engineering optimization design, in particular, heat and transfer system.

Before we can apply the EAs, a data set must be generated for searching. In traditional engineering design, a large amounts of historic data from industry should be collected and this method is extremely expensive in both economic and time. Recently, we can adopt CFD to generate the required database. Therefore, a new technology called CFD-EA coupling method has attracted much attention of researchers. This new method can be applied in many fields, including ventilation system design (Lin et al. 2005, Lin et al. 2009, Zhou and Haghighat 2009a), fire detection (Stavrakakis and Markatos 2009), truss structures design (Luh and Lin 2011), shape optimization (Peri et al. 2001), heat transfer (De Bellis and Catalano 2012, Charnay et al. 2001), composites material manufacturing (Santos et al. 2012, Wang et al. 2012), etc.

The application of CFD simulation technology makes the engineering optimization design more time efficient and economic compared with experimental design. Particularly, in the field of HVAC system optimization design, a lot of research has been done using numerical methods, which combine CFD and EAs. A review will be given as follows.

In 2006, K.F.Fong et al. studied energy management optimization in a HVAC system using simulation-EP (evolutionary programming) coupling technology. They built a mathematical model of energy consumption in a subway station with the help of TRNSYS and EP optimization was conducted to search for optimal design parameters. Their results show that the simulation-EP coupling method is feasible for handling discrete, non-linear and highly constrained nature of the HVAC problems. Zhou, L. and F. Haghighat proposed a CFD-ANN-GA optimization method for ventilation system design in (Zhou and Haghighat 2009a,b). They used a validated CFD model to get discrete airflow profiles with responding to different combinations of inlet boundary conditions. Artificial Neural Network (ANN) was then adopted to generate continuous response surface of design objectives based on the acquired data and finally, a GA optimizer searched for the optimal design parameter. Their results indicate that the combination of CFD, ANN and GA is a good way for solving engineering optimization design problem. In 2013, Li et al. improved Zhou's work by introducing a POD model instead of the ANN to

get the response surface in a faster way and they got similar optimal results. In addition, the CFD-EA coupling method was also adopted in other research papers related to HVAC system optimization design (Stavrakakis et al. 2011, Zhai et al. 2014).

2.3 Machine learning techniques

Machine learning, seen as a subset of artificial intelligence (AI), is an application of AI that provides systems with the ability to automatically learn and improve from experience without being explicitly programmed. The focus of machine learning techniques is to develop a computer program that can access data and use it to learn for themselves. Machine learning algorithms build a mathematical model of sample data, known as "training data", in order to make predictions or decisions without being explicitly programmed to perform the task (Koza et al. 1996, Bishop 2006). Machine learning algorithms are used in a wide variety of applications, such as email filtering, and computer vision, where it is infeasible to develop an algorithm of specific instructions for performing the task.

According to their approach, the type of data they input and output, and the type of task or problem that they are intended to solve, the machine learning algorithms are often categorized as supervised and semi-supervised learning, unsupervised learning, and reinforcement learning. Supervised learning algorithms build a mathematical model of a set of data that contains both the inputs and the desired outputs (Russell and Norvig 2016). The data is known as training data, and consists of a set of training examples. Each training example has one or more inputs and a desired output, also known as a supervisory signal. In the case of semi-supervised learning algorithms, some of the training examples are missing the desired output. In the mathematical model, each training example is represented by an array or vector, and the training data by a matrix. Through iterative optimization of an objective function, supervised learning algorithms learn a function that can be used to predict the output associated with new inputs (Elkan 2001).

Unsupervised learning algorithms take a set of data that contains only inputs, and find structure in the data, like grouping or clustering of data points. The algorithms, therefore, learn from test data that has not been labeled, classified or categorized. Instead of responding to feedback, unsupervised learning algorithms identify commonalities in the data and react based on the presence or absence of such commonalities in each new piece of data (Tucker 2004). Reinforcement learning is an area of machine learning concerned with how software agents ought to take actions in an environment so as to maximize some notion of cumulative reward. Due to its generality, the field is studied in many other disciplines, such as game theory, control theory, operations research, information theory, simulation-based optimization, multi-agent systems, swarm intelligence, statistics and genetic algorithms (Alpaydin 2014).

Although pioneering machine learning research was conducted in the 1950s using simple algorithms, it did not attract much attentions until 2010s when AI beats human in Go game. Because of new computing technologies, deep learning has become feasible, which leads to machine learning becoming integral to many widely used software services and applications, such as image recognition, object detection, human face recognition, chat bot, etc.

Among the deep learning algorithms, the most popular one is Convolutional Neural Networks (CNN), which has been widely used in image recognition and natural language processing. The CNN is a class of deep artificial neural networks firstly introduced by (Fukushima 1980), which has been successfully applied to analysing two-dimensional visual images, such as hand-written character recognition (Lecun 1989, Lecun and Bengio 1995, Niu and Suen 2012). A typical structure of CNN consists of one input and one output layer with multiple hidden layers in between as shown in Figure 2.10. The hidden layers of a CNN typically consist of a series of convolutional layers, subsampling (pooling) layers and fully connected layers (see Figure 2.10). The convolutional layers apply convolution filters to the original input to extract the feature maps which is passed to the next layer. Following the convolutional layer, there may be a pooling layer which combines the outputs of neuron clusters at one

layer into a single neuron. There are several methods of pooling, including max pooling, average pooling and linear combination pooling. By applying convolutional and pooling, the CNN significantly reduces the number of neurons, allowing the network to be deeper with fewer parameters (Aghdam and Heravi 2017), and making the network practical for training with limited computational resources. It is worth noting that there can be multiple convolutional layers and pooling layers in a CNN and more layers makes the network deeper. Finally, the fully connected layer receives the information from last hidden layer and turns it into high-level reasoning output.

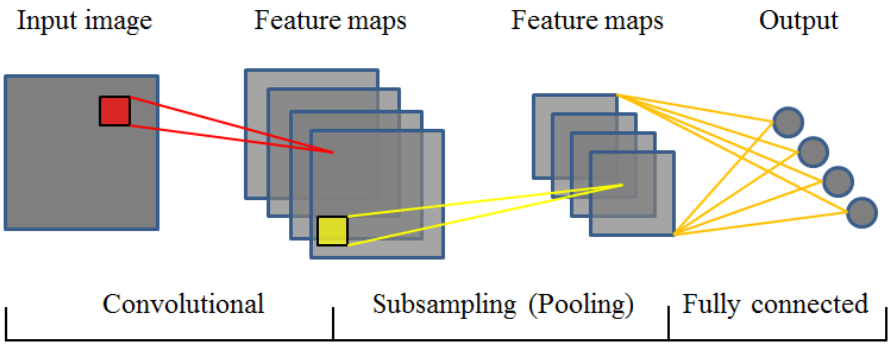

Figure 2.10 A typical CNN structure

Since the first successful application of LeNet on hand-written digits recognition (Lecun et al. 1990), the CNN has shown great potential on solving different problems including image classification (Ciresan et al. 2011), face recognition (Parkhi et al. 2015), natural language processing (Kumar et al. 2016) and object tracking (Hong et al. 2015).

Chapter 3

Multi-objective optimization of HVAC system in an office room using NSPSO

The main findings of this chapter have been published in:

- **Li, N.,** Cheung, S. C., Li, X. & Tu, J. (2017). Multi-objective optimization of HVAC system using NSPSO and Kriging algorithms—A case study. *Building Simulation,* 10(5), 769-781. ***IF = 1.673, Q1***

- **Li, N.,** Cheung, S. C., Li, X. & Tu, J. (2015). Multi-objective optimization of thermal comfort and energy consumption in a typical office room using CFD and NSM-PSO. *The 21st International Congress on Modelling and Simulation (MODSIM 2015)*, Gold Coast, Australia, November 2015.

In modern building design, engineers are constantly facing challenging to find an optimal design to maintain a high level of thermal comfort and indoor air quality for occupants while minimizing the system energy consumption. Over the past decades, several algorithms have been proposed and developed for optimizing the heating, ventilation and air conditioning (HVAC) system for indoor environment. Nevertheless, the majority of these optimization algorithms are focused on single objective optimization procedures and require a large training sample for surrogate modelling. For multi-objective HVAC design problems, previous studies introduced an arbitrary weighting factor to combine all design objectives into one single objective function. The near-optimal solutions were however sensitive to the chosen value of the weighting factor. Aiming to develop a multi-objective optimization platform with minimal computational cost, this paper presents a nondominated sorting-based particle swarm optimization (NSPSO) algorithm together with the Kriging method to perform optimization for the HVAC system design of a typical

office room. In addition, an adaptive sampling procedure is also proposed to enable the optimization platform to adjust the sampling point and resolution in constructing the training sample. The significant computational cost could be reduced without sacrificing the accuracy of the optimal solution. The proposed methods are applied and assessed in a typical HVAC system and the results indicate that compared to traditional methods, the presented approach can handle multi-objective optimization in a ventilation system with up to 46.6% saving of computational time.

3.1 Introduction

In the modern indoor environment design, driven by the higher expectation of occupants and soaring energy cost, indoor thermal comfort and energy efficiency are the two main concerns in the heating, ventilation and air conditioning (HVAC) systems. To achieve a higher system performance, computational fluid dynamics (CFD) simulation tools such as ANSYS Fluent, StarCCM+ and OpenFOAM have been widely adopted to analyze air distribution and flow characteristics and their relationship in thermal comfort and energy consumption (Ravikumar and Prakash 2009, Cardinale et al. 2010, Hiyama et al. 2010, Kochetov et al. 2015, Gangisetti et al. 2016).

In comparison to the traditional design cycle, CFD simulations offer a faster and more economical way for engineers to carry out parametric studies, leading towards a more desirable system design. According to our previous works (Tu 2008), CFD has been proved to be a reliable prediction tool which has been widely adopted in industrial applications and academia research. Albeit a near-optimum solution could be obtained, the parametric analysis using CFD technique is inherently discrete in the design space with pre-selected design variable values. The best optimum solution could be “hidden” in the discretized domain (Stavrakakis et al. 2011). Furthermore, the accuracy of a near-optimum solution depends largely on the “resolution” of the test matrix where significant computational costs are required. To enhance the accuracy of the near-optimum solution with practical computational time and

resource, Artificial neural network (ANN) (Varol et al. 2007, Zhou and Haghighat 2009a,b, Stavrakakis et al. 2011, Lu and Xue 2014) or other surrogate techniques (Li et al. 2012, Li et al. 2013) are employed as an alternative approach to approximate the nonlinearity and complex behaviour of the multidimensional systems. One of the first studies using CFD-ANN coupled approach to assess the effect of architectural-designs on the thermal comfort can be found in (Krauss et al. 1997). In general, numerical results predicted by CFD modelling were adopted to establish a database for training the ANN or surrogate models. These CFD-trained ANN or surrogate models then captured the relationship between design parameters and objective function. Optimization procedures using gradient methods (Gyulai et al. 2007, Stavrakakis et al. 2011, Welle et al. 2011) or evolutionary optimization algorithms (Luh and Lin 2011, Li et al. 2013, Afrand et al. 2015, Zhai et al. 2014) were then performed using the trained models for allocating the near-optimum solution within the continuous design space. Although significant computational time and resource can be reduced, a considerably large amount of CFD results are still required for constructing a reliable database for the training of ANN or surrogate models.

On the other hand, most of the HVAC system design normally involves multi-objective considerations. Design indices such as predicted mean vote (PMV), percentage dissatisfied of draft (PD), age of air, CO2 concentration and energy cost are commonly considered in literatures. Especially, in terms of indoor thermal comfort evaluation, substantial research works have done by Ricciardi's group (Buratti and Ricciardi 2009, Buratti et al. 2013, Nematchoua et al. 2014, Ricciardi and Buratti 2015, Buratti et al. 2016, Ricciardi et al. 2016). In most previous works, a single objective function was constructed by aggregating several design indices using pre-defined weighting factors (Laverge and Janssens 2013, Li et al. 2013). One particular disadvantages of this method is that the optimal solution could be sensitive to the values of the weighting factors. In other words, different values of weighting factors could result in substantially different solutions. The weighting factors must be therefore chosen carefully based on subjective factors such as engineering or expert judgements. Furthermore, the optimization procedure gives only one near

optimal solution where there is no flexibility for the designer to strike a balance or "trade-off" of the conflicting parameters. For example, a lower indoor temperature may be preferable in summer which is in conflict with the goal to minimize energy consumption.

As an attempt to overcome the aforementioned shortcoming, in this study, we propose the use of a nondominated sorting-based particle swarm optimization (NSPSO) algorithm to achieve multi-objective optimization without having to use any weighting factors. This population-based algorithm, as an improved technique of the basic particle swarm optimization (PSO), is capable to obtain a set of nondominated solutions (i.e. approximated Pareto front solutions); providing the engineers a set of optimal solutions where the most appropriate design solution based on professional judgment or end-user desire can be chosen (Carrese et al. 2011). Furthermore, to minimize the computational requirement for constructing a reliable training database, Kriging or Gaussian process regression together with adaptive sampling technique is also adopted to dynamically allocate additional CFD simulation dataset where there is a higher likelihood of having a near optimal solution. A case study is used to demonstrate the feasibility of the proposed optimization approach in a real-world HVAC application.

3.2 Multi-objective optimization methods

3.2.1 Basic principles of particle swarm optimization

Traditional mathematical programming methods for solving both single- and multi-objective optimization problems have been successfully adopted in many science and engineering problems (Martínez and Coello 2013). Nevertheless, it is also well known that these methods have difficulty in solving non-convex and multimodal problems (Deb 2001). In contrast, population-based stochastic optimization methods such as Evolutionary Algorithms (EAs) have the advantage of not requiring gradient information in the optimization process which could provide remedy for this class of problems (Martínez and Coello 2013). Among

existing population-based stochastic optimization methods, Particle Swarm Optimization (PSO) has proven to be faster in convergence in comparison with standard EAs (Hassan et al. 2005). PSO was first introduced by (Kennedy 2001) based on the inspiration drawn from observations of the social behaviours of insects including learning from previous experience and communicating with successful individuals. In PSO, each particle has its own position and velocity, which are represented by *xi* and *vi*, respectively. The position and velocity of the particle are updated according to the following equations:

$$\begin{aligned} v_i(\mathrm{t}+1) &= \omega v_i(\mathrm{t}) + c_1\varphi_1(p_i - x_i(\mathrm{t})) + c_2\varphi_2(p_g - x_i(\mathrm{t})) \\ x_i(\mathrm{t}+1) &= x_i(\mathrm{t}) + v_i(\mathrm{t}+1) \end{aligned} \tag{3.1}$$

where p_i and p_g represent the personal best position and global best position, respectively, and the c_1 and c_2 are two uniform random numbers within the range [0, 1]. The φ_1 and φ_2 are two constants which are usually set to 2. The parameter ω decreases with iterations within the range [0.4, 1.2]. To avoid going out of the search space, both the position and velocity are limited within boundaries, [x_{min}, x_{max}] and [v_{min}, v_{max}], respectively. Nevertheless, it is worth noting that the original PSO can only provide solutions for single-objective optimization problems.

3.2.2 **Nondominated sorting based PSO**

Inspired by the works done by (Deb et al. 2002), (Li 2003a) proposed a Nondominated Sorting Method to extend the original PSO to multi-objective optimization problems (MOP) - namely Nondominated Sorting based Particle Swarm Optimization (NSPSO). In NSPSO, the updating equations for particle position and velocity remain unchanged, but the selection of the personal best and global best has been re-designed. Two main mechanisms are used to determine the global best among the population – 1) nondominated sorting for identifying different fronts, and 2) crowding distance computed for particles within each front to encourage solution diversity. These kinds of information are used to select suitable leaders (i.e. global best) at each iteration to guide the

particles moving towards the Pareto-optimal Front while still maintaining a good distribution of solutions along the Pareto-front.

a) *Nondominated Sorting*

Figure 3.1 shows an example of the nondominated sorting process. Considering 2 objectives (i.e. *f1* and *f2*) to be optimized in the process, the entire population (i.e. the 10 particles that labelled as 1 to 10) is sorted into different levels of fronts according to the domination comparisons between particles. The particles in same front are nondominated with each other. As depicted in Figure 3.1, Front 1 is the highest-level nondominated front because all particles in it are not dominated by any other particles in the entire population. The main goal of nondominated sorting is to classify the whole population into different levels of nondominated fronts. The global best particle (leader of the population) can be randomly selected from the highest-level front. This kind of selection process will push the whole population towards the true Pareto Front. More information regarding nondominated sorting can be found in (Li 2003a).

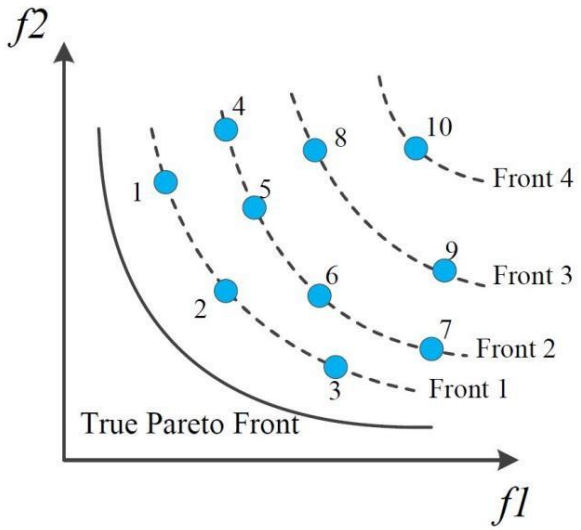

Figure 3.1 An example of nondominated sorting process in NSPSO.

b) Crowding distance

Unlike in single-objective optimization, maintaining the diversity in a set of solutions is vital in a multi-objective optimization (Deb et al. 2002, Li 2003a). Throughout the optimization process, the leader must be selected properly to avoid local optimal aggregation of the whole population. In NSPSO, computing the crowding distance values among particles in the highest-level of nondominated front is used to select leaders that are both good and far apart from each other. Inspired by (Deb et al. 2002), we introduced a new way to calculate the crowding distance. Figure 3.2 Crowding distances among individuals in the highest-level nondominated front. shows an example of the crowding distance among particles. For each particle, the crowding distance is defined as the following:

$$D_n = \begin{cases} \infty & n = 1, N \\ d_{n-1} + d_n & 1 < n < N \end{cases} \tag{3.2}$$

The particle with a higher crowding distance value will have a high probability to be selected as the leader. Consequently, particles in the top front are likely to maintain a good level of population diversity.

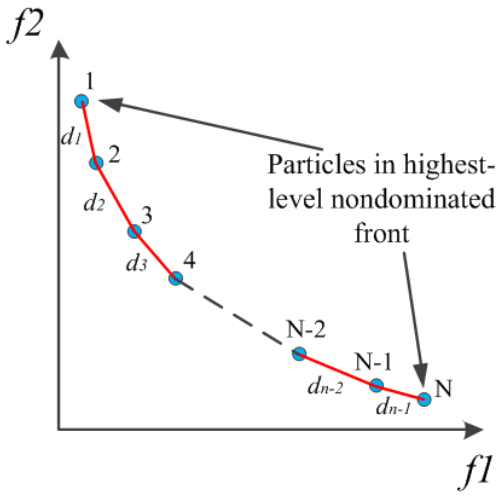

Figure 3.2 Crowding distances among individuals in the highest-level nondominated front.

Compared with other MOEAs, such as NSGAII and decomposition-based MOEAs, NSPSO has the advantages on running speed and convergence. As

the main goal of this study is to save computational cost, NSPSO was selected to conduct multi-objective optimization.

3.2.3 **Surrogate modelling**

Evolutionary optimization algorithms are efficient in obtaining a representation of the Pareto front for the MOP (Carrese et al. 2011). However, generating all the elements in an objective array by simulations could still be computationally demanding. Previous studies have adopted surrogate modelling techniques (e.g. ANN (Varol et al. 2007, Lu and Xue 2014) and Support Vector Machine (SVM) (Zhao 2009)) which is trained by computational simulated samples to achieve substantial saving in computational time and resources. However, to ensure its accuracy, ANN and SVM require considerably large training samples which pose significant burden on the computational cost. Alternatively, other than the ANN and SVM algorithms, the Kriging method has aroused much attention due to its capability in achieving high prediction accuracy with relatively small training sample size. Aiming to minimize the computational cost for large training samples, Kriging method is therefore adopted and incorporated with NSPSO in the present study. A brief introduction to the Kriging method is presented in this section. More detailed derivation and formulation of the Kriging method can be found in (Forrester et al. 2008) and the references therein. The Kriging technique provides the best linear unbiased estimator of the unobserved fields based on the sampled data (Journel and Huijbregts 1978). The basic idea of Kriging is to predict the value of a function at a given point by computing a weighted average of the known values of the function in the neighborhood of the point, which is expressed as:

$$\hat{z}(\mathbf{x}_L) = \sum_{\alpha=1}^{N} \lambda_\alpha z(\mathbf{x}_\alpha) \tag{3.3}$$

where $\hat{z}(\mathbf{x}_L)$ represents a local estimation at the data location $\mathbf{x}_L$, $z(\mathbf{x}_\alpha)$ is the sampled value at the data location $\mathbf{x}_\alpha$ and λ_α represents the weighting

coefficient which can be calculated by minimizing the estimation variance:

$$\min E\left\{\left[z(\mathbf{x}_L)-\hat{z}(\mathbf{x}_L)\right]^2\right\}= \\ C(0)-2\sum_{\alpha}\lambda_\alpha C(\left\|\mathbf{x}_\alpha-\mathbf{x}_L\right\|)+\sum_{\alpha}\sum_{\beta}\lambda_\alpha\lambda_\beta C(\left\|\mathbf{x}_\alpha-\mathbf{x}_\beta\right\|) \tag{3.4}$$

subjects to the unbiased condition:

$$E[z(\mathbf{x}_L)-\hat{z}(\mathbf{x}_L)]=0 \tag{3.5}$$

and the normalization condition:

$$\sum_{\alpha=1}^{N}\lambda_\alpha=1 \tag{3.6}$$

The weighting coefficient λ_α in Equation (3.3) can be solved using a quasi-Newton optimization method or other similar algorithm (Gano et al. 2006). Finally, the prediction value at the unobserved location can be given by Equation (3.3).

3.3 Case description

3.3.1 CFD modelling and validation

To assess the feasibility and performance of the aforementioned approach, this investigation focuses on a practical HVAC design optimization case study where the air quality, thermal comfort and energy consumption of a typical office room are optimized against the supply air velocity and supply air temperature (Zhou and Haghighat 2009a, Li et al. 2013). The case study is constructed with a reference to a full-scale experimental measurements reported by (Yuan et al. 1999). Figure 3.3 shows a three-dimensional representation of the typical office room. According to the experimental setup, the outside temperature is 26.7°C, and room temperature is maintained between 23.3°C - 26°C using displacement ventilation system. The supply cooled air is discharged one

side at the low level of the room (i.e. label 1 in the figure) and the return air leaves the room from the exhaust at the center of the ceiling (i.e. label 2). Two heat sources mincing two office workers were placed in the room (i.e. label 3 and 4). Computers and lightings were also dissipating heat to the room. In this study, all simulations are assumed to be steady state where ideal gas law were adopted for the air properties. More details of boundary conditions have been tabulated in Table 3.1. Based on the previous work by Yuan et al. (1999), experimental measurements were carried out within an environmental chamber where solar radiation contribution was neglected. To validate our CFD with the experimental data, no solar radiation is incorporated in the simulation.

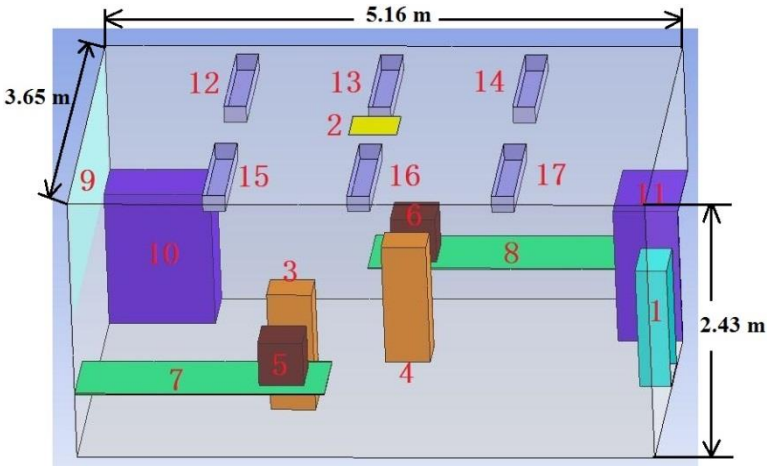

Figure 3.3 The geometry layout of the typical office room.

Table 3.1 The boundary conditions adopted in the CFD simulations.

Number	Name	Boundary details	Comments
1	Air-conditioning	Normal speed & Static temperature	Controlled variables
2	Exhaust	Average static pressure	0 [Pa]
3,4	Occupant	Temperature	37 [°C]
5,6	Desktop	Heat flux	108.5[W/m^2]
7,8	Table	Adiabatic	-------
9	Partition window	Heat transfer coefficient	3.7[W/(m^{2}K)]
10,11	Furniture	Adiabatic	-------
12-17	Light	Heat flux	34[W/m^2]

	Room wall	Heat transfer coefficient	0.19[W/(m^{2}K)]

A CFD model of the office room was built in ANSYS Workbench, which contains in total 1,043,811 nodes and 2,849,852 elements. To ensure the validity of the CFD simulation, predictions of the CFD model were first validated against the full-scale experimental data reported by (Yuan et al. 1999). In the experiment, a hot-sphere anemometer system was used for air velocity, velocity fluctuation measurements and a thermocouple system was used to measure surface and air temperatures. Figure 3.4 shows the comparisons between the measured and predicted air temperature and velocity along the vertical line at the center of the office room where supply air temperature and velocity were maintained at 17°C and 0.09m s^{-1} respectively.

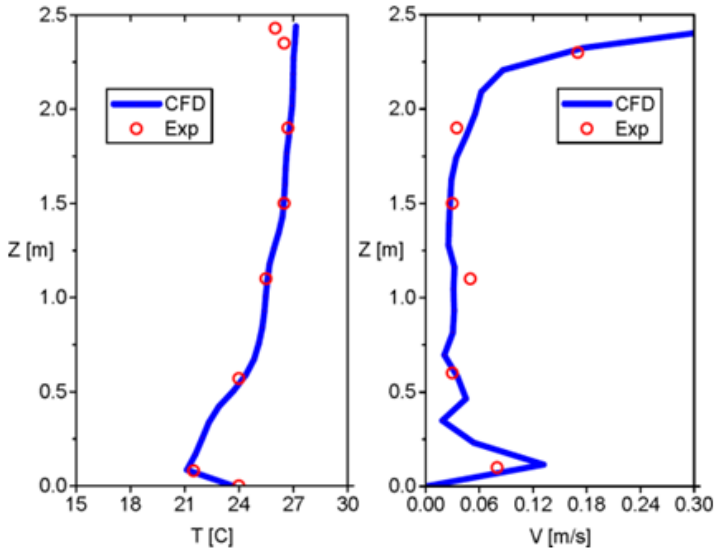

Figure 3.4 Comparisons between the CFD simulation results and experimental data.

As depicted, the predicted temperature variation was successfully captured by the CFD model and compared well agreed with the measurements. Similarly, the velocity profile was also in good agreement with the experiment. These encouraging results show the reliability of the CFD predictions and its capability for providing sampling data for the design optimization procedures.

3.3.2 **Optimization objectives**

In the presented study, the Predicted Mean Vote, CO_2 concentration, and energy consumption were adopted to quantitatively assess the performance of the HVAC system in terms of thermal comfort, air quality and energy efficiency respectively. The definition of the Predicted Mean Vote, CO_2 concentration, and energy consumption are briefly discussed in the following sections.

a) *Predicted mean vote*

The predicted mean vote (PMV) is a thermal comfort evaluation index which was first introduced by (Fanger 1972). It is used to assess indoor thermal comfort based on heat balance and a set of experimental data collected from a given controlled climate chamber. The index represents the mean subjective satisfaction with the indoor thermal environment with a number between -3 (cold) and +3 (hot). The zero value is defined as the ideal representation of thermal neutrality. The PMV index is evaluated based on an empirical equation which is correlated to the local air temperature, mean radiant temperature, relative humidity, air speed, metabolic rate, and clothing insulation (Fanger 1972). In this study, we assume that the occupants are seated in quiet position (i.e. metabolic rate of 1.0 met) with a summer clothing (i.e. 0.2 clo), and we evaluated the average PMV based on the predicted field information obtained from CFD simulations.

b) *CO_2 Concentration*

To assess the air quality within the space, the concentration of CO_2 emitted by occupants throughout the office room was also resolved in the CFD simulation. In the simulation, the CO_2 is emitted from the occupants with the emission rate 0.87L min^{-1}. Similar to the average PMV, the average CO_2 concentration was extracted from the predicted CFD field information.

c) *Energy consumption*

Following the previous study (Zhou and Haghighat 2009a, Li et al. 2013), the energy consumption of the air-conditioning system is divided into two parts: ventilation fan power and the cooling or heating load. Energy consumption in the two parts is determined as follow:

$$\begin{aligned} E_{fan} &= \frac{P \cdot V_{air}}{\eta_{fan}} \\ E_{cooling/heating} &= m_{supply} c_p (T_{return} - T_{supply}) \\ &\quad + m_{outdoor} (h_{outdoor} - h_{return}) \\ E_{total} &= E_{fan} + E_{cooling/heating} \end{aligned} \tag{3.7}$$

where P is air pressure difference of the fan and V is volume flow rate of supply air (m3 s-1), m represents the mass flow rate of the air (kg s-1), c_p is the specific heat capacity of air, T represents temperature, h is the specific enthalpy of air (J kg^{-1}) which is related to air temperature and relative humidity. Similarly, we can get energy costs from the CFD-Post package.

3.4 Optimization results and analysis

3.4.1 Multi-objective optimization platform

As mentioned earlier, the aim of this case study is to optimize the value of |PMV|, CO_2 concentration and energy consumption against a set of control design parameters (i.e. supply air velocity and temperature). However, these three indices are in conflict with each other, which means there does not exist an optimal solution where all indices are at the minimal value. Weighting factors were used to get trade-off solutions in previous research (Zhou and Haghighat 2009a, Li et al. 2013). In order to overcome the drawbacks of the weighting methods, we proposed a multi-objective optimization platform by integrating the CFD modelling technique, Kriging method and NSPSO algorithm to consider the three objectives simultaneously and obtain the corresponding Pareto Front of without any weighting factor. A schematic of the methodology is depicted in Figure 3.5.

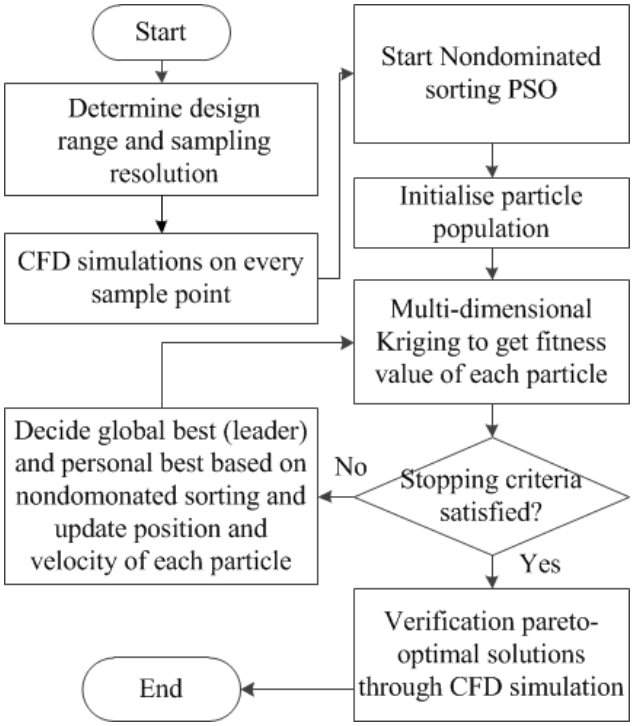

Figure 3.5 CFD-based multi-objective optimization system framework.

As depicted in the flowchart, the CFD simulation technique is adopted to establish the sample data as the input of the Kriging method. Following the previous studies (Zhou and Haghighat 2009a, Li et al. 2013), a total of 25 CFD simulations with different combinations of controlled variables have been carried out (see also Figure 3.6). Similar to the validation study, all simulations were carried out using the ANSYS CFX 14.5 with the identical mesh resolution and boundary conditions (expect the supply air temperature and velocity). Based on the simulated results, local air velocity, temperature and associated parameters were extracted for evaluating the corresponding PMV, CO_2 concentration and energy consumption value. All the obtained values (i.e. a total of 25 set of data) were then used to construct the sample data for the Kriging method. In terms of thermal comfort, we would like the PMV value to be as close to 0 as possible. At the same time, we also would like to minimize CO_2 concentration and energy consumption. In order to describe the conflicting relationships among these three objectives, we listed three groups of typical values in Table 3.2. The first row in Table 3.2 shows the point where the PMV is the closest to 0 while both the CO_2 concentration

and the energy consumption are quite large. Similarly, the second and the third row show the points where the CO_2 concentration and the energy consumption are minimum, respectively, while the other two objectives are quite large. Therefore, these three objectives are conflicting with each other. We cannot find a point where all the objectives are at their minimum value. As we mentioned before, traditional methods using weighting factors for solving MOP are inefficient. In this paper we developed a multi-objective optimization platform based on NSPSO to solve MOP efficiently. The details and results will be described in the followings.

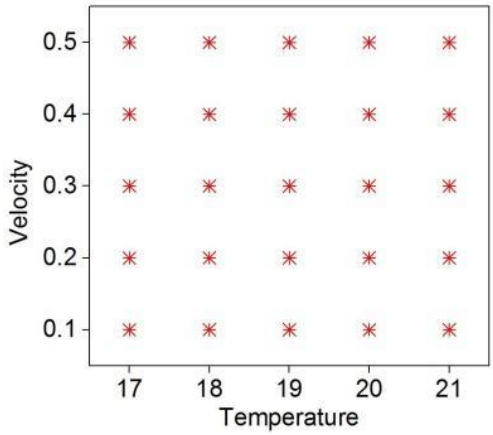

Figure 3.6 Definitions of inlet boundary conditions in CFD simulations.

Table 3.2 Three typical groups of values minimizing PMV, CO2, Energy, respectively.

T(°C)	V(m/s)	PMV	CO_2	Energy
19.5	0.19	-0.063	0.078	863.6
21	0.5	-0.964	0.028	1671.8
21	0.1	0.327	0.083	334.4

In order to study the accuracy of the Kriging prediction, 16 more CFD simulations were added, which are located in the center of each sampling grid. We compared the values of the three objectives exported from the CFD-Post with the values predicted by the fully Kriging surrogate model (i.e. the Kriging model using all the 25 CFD sample data defined in Figure 3.6). The percentage errors between the CFD results and the Kriging prediction are shown in Figure 3.7. From Figure 3.7, we can see that the

maximum errors of prediction for PMV and CO_2 are less than 5.2% and the maximum error of prediction for Energy is less than 0.6%, which indicates the Kriging prediction has achieved a good accuracy.

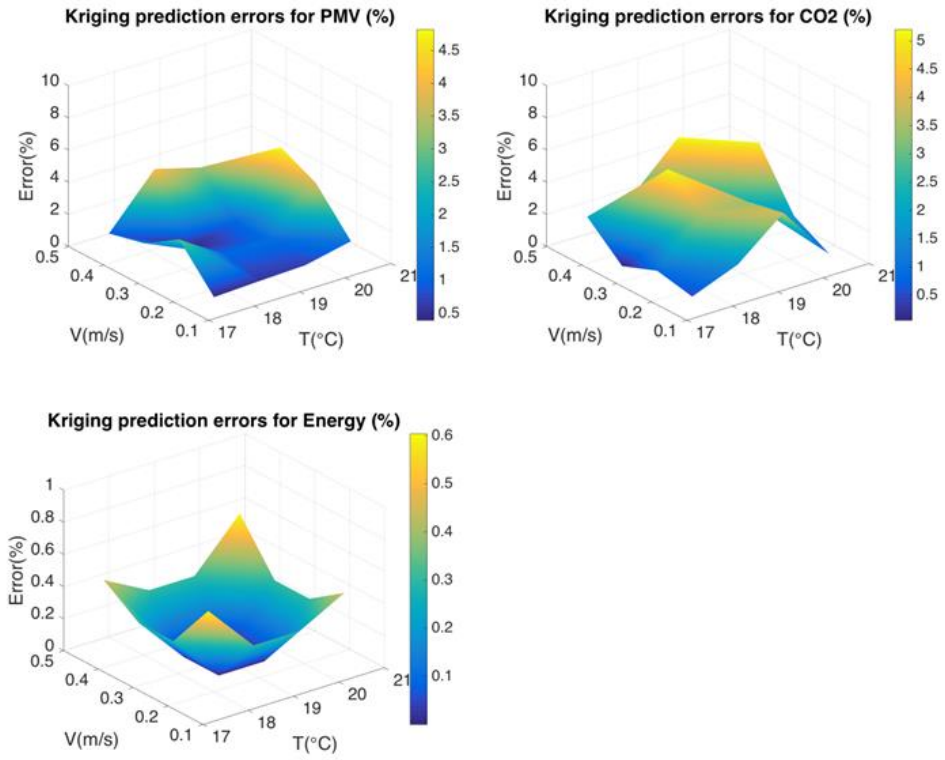

Figure 3.7 Percentage errors of Kriging prediction for the three objectives - PMV, CO2 and Energy, respectively.

Afterward, a three-objective optimizer based on the NSPSO algorithm was implemented using the MATLAB R2013b. In the optimization procedure, the Kriging surrogate method was used to calculate the fitness values for each particle in the population. The initial size of the swarm population was 200 and the maximum iteration number was set to be 100. Figure 3.8 shows the trade-off solutions given by the NSPSO algorithm. The blue dots represent the finally reserved particles in the objective space, which constitute a set of solutions approximating the Pareto-Front. In order to compare our results with the results given by weighted method introduced in (Li et al. 2013), the previous results are also plotted in Figure 3.8 (shown as red stars). Each red star in Figure 3.8 represents a solution given by a set of fixed weighting factors. From the comparison,

we can notice that the solutions given by previous method (red stars) are only a small subset of the Pareto front given by our approach. It is also noted that our approach does not depend on any weighting factor, and after only one simulation run, a set of trade-off solutions can be found in the objective space, providing the designers with the choices of a range of trade-off solutions.

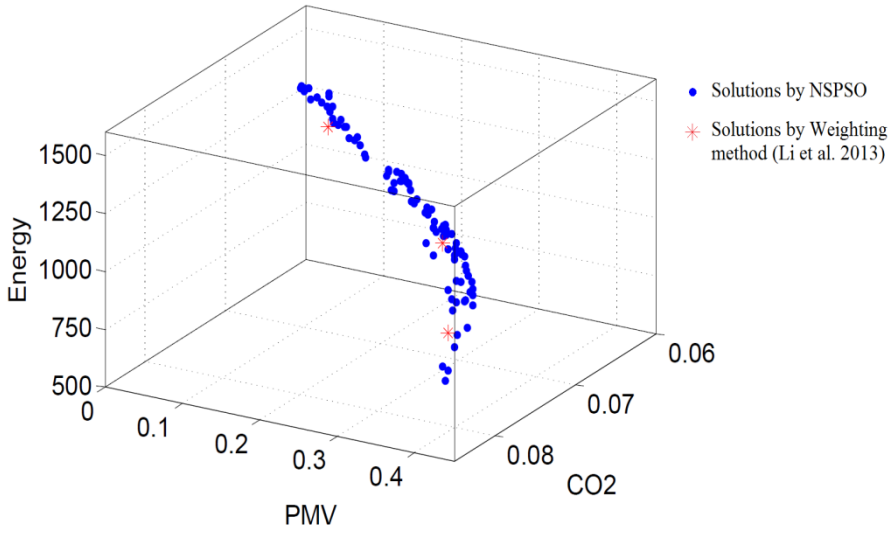

Figure 3.8 Comparison of solutions given by NSPSO and solutions given by Weighting method in 3D objective space.

3.4.2 **Adaptive sampling for Kriging**

The above section has clearly demonstrated the capacity of the proposed multi-objective optimization platform. By replacing the CFD simulations with the Kriging method, the optimization process could reduce significantly the computational time. Nevertheless, in constructing the sample data, CFD simulations are uniformly distributed within the design range (see Figure 3.6). The main disadvantage of uniform sampling is its high cost, because the sampling density must be uniformly high everywhere in order to meet the sampling requirements in some particular area. Nevertheless, in most practical cases, the final optimal solutions are normally concentrated in a certain region within the design space. Therefore, substantial computational time could be wasted in constructing the sample data for some virtually redundant samples which are far away from the optimal solution. To strike a balance between

accuracy and computational cost, the sampling points should be strategically placed in the region where has a higher likelihood of getting optimal solutions rather than uniformly distributed throughout the design space. To achieve this, we introduced an adaptive sampling procedure to determine sampling point. Adaptive sampling designs, also known as response-adaptive designs, are ones where the accruing data (i.e., the observations) are used to adjust the experiment as it is being run (Hardwick and Stout 2016). The adaptive sampling method is then adopted to govern the construction of the sample data where sampling points are allocated based on the likelihood of having optimal solution within the region. A flowchart showing the procedures in constructing the sample data for Kriging and its integration with the multi-objective optimization platform is also shown in Figure 3.9.

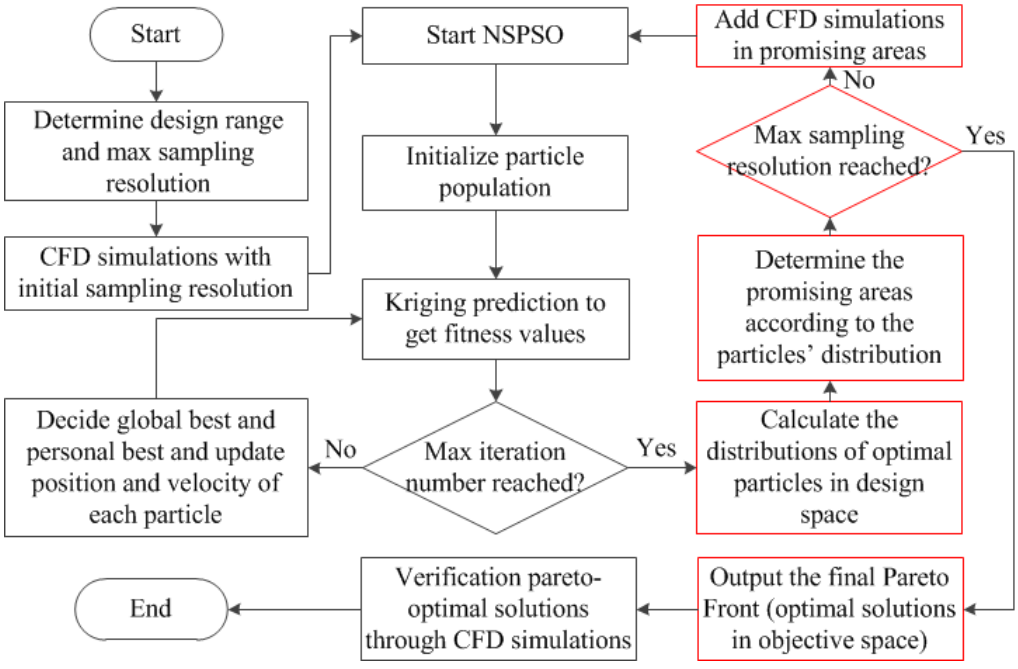

Figure 3.9 Framework of the CFD-based multi-objective optimization approach with adaptive sampling procedure.

To assess the potential computational saving, the adaptive sampling procedure will be applied to the same HVAC design optimization case study and compared against the results from uniform sample data (i.e. as presented in previous section). To make a comparison with the previous

result, instead of locating the sampling point in any point within the design space, the adaptive sampling method is applied to determine the next sampling location based on halving method. This is to ensure the sampling locations are consistent with the uniform sampling method. To initialize the process, four sampling locations (red stars in Figure 3.13) were firstly specified at all corners of the design space (i.e. corner at the minimum and the maximum value of supply air temperature and velocity, which are (17, 0.1), (17, 0.5), (21, 0.1), (21, 0.5)). Based on the four sampling locations, Kriging method was then applied to evaluate the response surface. The NSPSO algorithm was then performed to evaluate the distribution of optimal particles throughout the design space.

After the CFD simulations are finished, we generate the response surfaces using Kriging prediction. Then we run the NSPSO for 100 times and calculate the distribution density of optimal particles in design space. The distribution density contour is shown in Figure 3.10(a) and the column charts in Figure 3.10(b) and Figure 3.10(c) indicate the projections of the density contour in temperature plane and velocity plane, respectively. As mentioned before, in order to make a comparison with the case described in Section 3.4.1, we want the sampling locations in this adaptive case would be a subset of the locations described in Figure 3.6. It is easy to reach this goal by using halving method. Therefore, in Figure 3.10(b) and Figure 3.10(c), the red line in the middle cuts the plane into two sides and for each side, reserving or removing depends on the distribution density of the optimal solutions in the areas. For example, in Figure 3.10(b), the right side (i.e. T > 19°C) should be reserved for adding more sampling points rather than the left side (i.e. T < 19°C), because there located much more optimal solutions (98%) on the right segment than the optimal solutions on the left segment (only 2%). Differently, in Figure 3.10(c), both sides are required to add more sampling points, because the results show that there are almost as many optimal solutions on both sides (57% vs 43%). The promising area after first halving process has been determined which is shown in Figure 3.12(a) (green shadow) and accordingly, the green points in Figure 3.13 are inserted to be simulated in next iteration. The first iteration of process described in Figure 3.9 has been finished. The next iteration process is almost same except with 4

more CFD sample data used in the Kriging prediction. The distribution density contour and having analysis are shown in Figure 3.11. The acquired promising area in the second iteration process is indicated in Figure 3.12(b) (red shadow) and the blue points in Figure 3.13 are added, accordingly. Since the minimum sampling resolution has been reached, we stop adding more sampling points. Figure 3.13 illustrates the final CFD sampling locations.

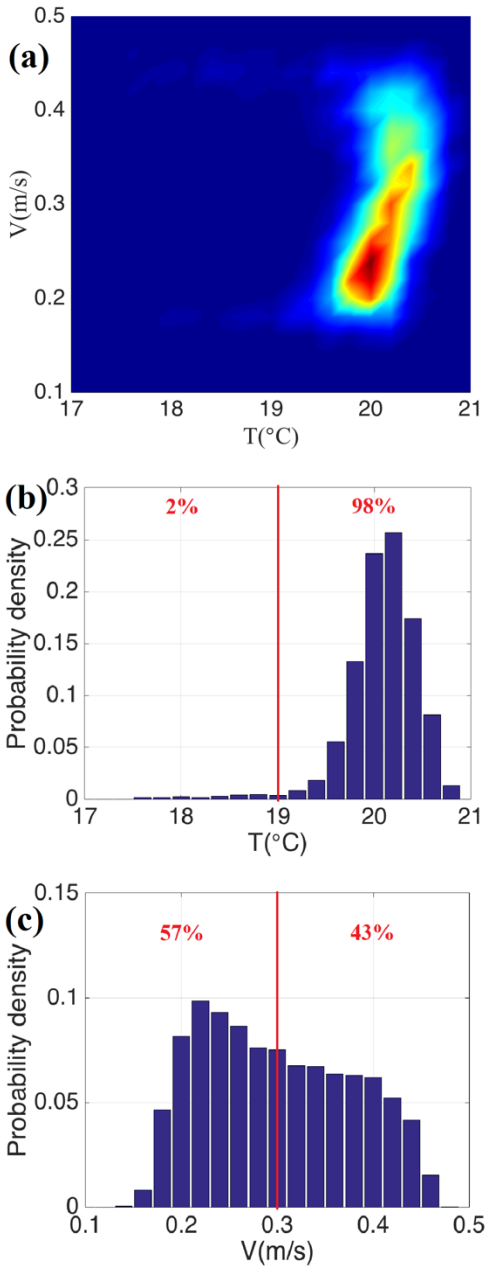

Figure 3.10 Probability density of optimal solutions in the design space (initial 4 CFD sampling locations). (a) Contour of probability density in 2D design space. (b) Projected probability density on the temperature design space. (c) Projected

probability density on the velocity design space.

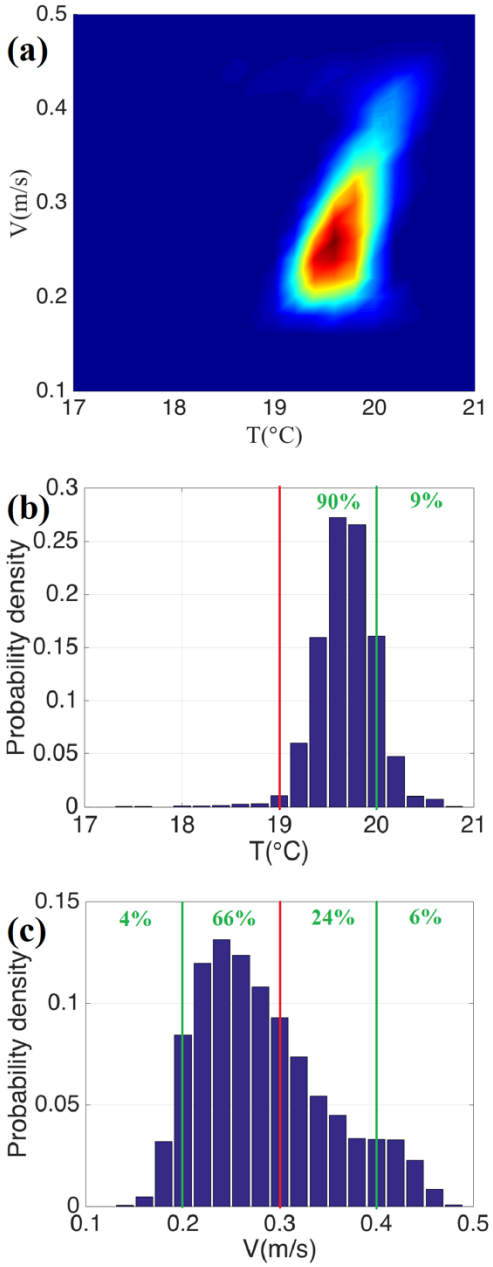

Figure 3.11 Probability density of optimal solutions in design space (initial 4 CFD sampling locations + 4 new adding sampling locations). (a) Contour of probability density in 2D design space. (b) Projected probability density on the temperature design space. (c) Projected probability density on the velocity design space.

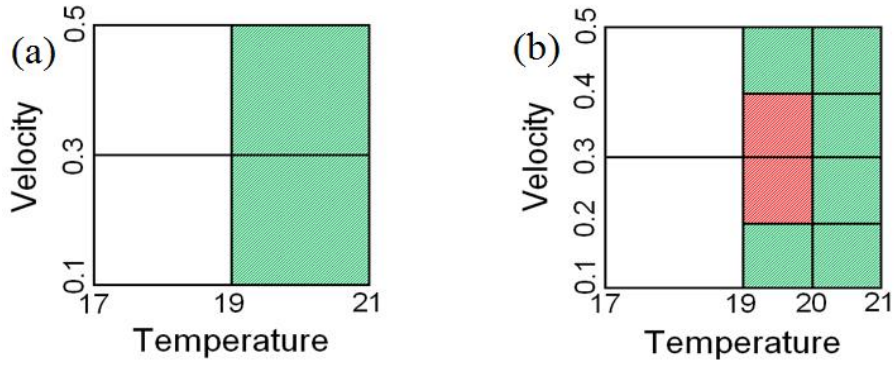

Figure 3.12 Halving process in design space. (a) Halving process in first iteration of adaptive sampling procedure cutting the whole design space into 4 parts. (b) Halving process in second iteration of adaptive sampling procedure cutting the promising areas (green shaded in (a)) into 4 parts.

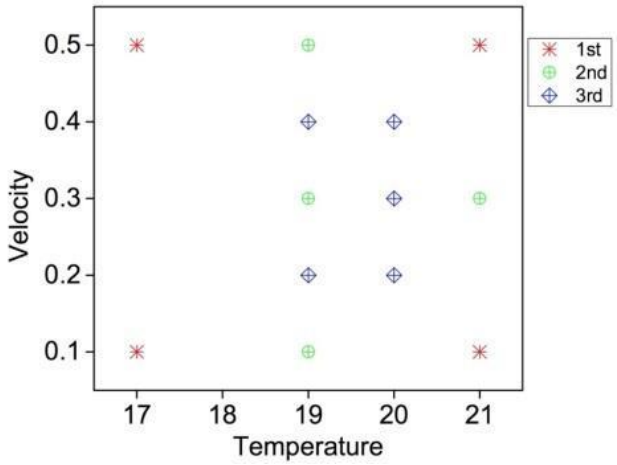

Figure 3.13 Final CFD Sampling locations in the adaptive sampling case.

With the adaptive sampling procedure, one can notice that the final sample data only constructed with 13 sampling points. Therefore, only 13 CFD simulations were required in constructing the sample database. Table 3.3 shows a comparison of the required total computational time (i.e. the total of CFD simulation and the optimization computational time) for both cases. The table clearly shows that the adaptive sampling procedure could reduce up to 46.6% of the total computational time.

Table 3.3 Comparisons of CPU time consumptions (500 particles in the optimization process).

Approach	**Procedure**	**CPU Time (s)**	**Iterations**	**Total CPU Time (s)**
Traditional	CFD simulation	7.904E+04	25	1.976E+06
	Single objective PSO	1.168	~ 500	~ 5.840E+02
NSPSO + Kriging	CFD simulation	7.904E+04	13	1.028E+06
	NSPSO	1.372E+02	200	2.744E+04
Total Saved CPU Time	9.211E+05(s) (**46.6%**)			

Although a significant saving has been achieved, it is essential to verify the accuracy of the optimal results with less sample data. Figure 3.14 shows the comparison of the predicted Pareto Front with the uniform sample data and the adaptive sampling procedure. As depicted, the blue dots are the solutions obtained from the uniform sample data; while solutions for adaptive sampling are represented in red dots. The figure clearly shows a good agreement of the Pareto Front obtained from both cases; demonstrating that the NSPSO algorithm together with the Kriging method are capable to maintain the accuracy of the optimal solution while achieving significant saving with less sample data.

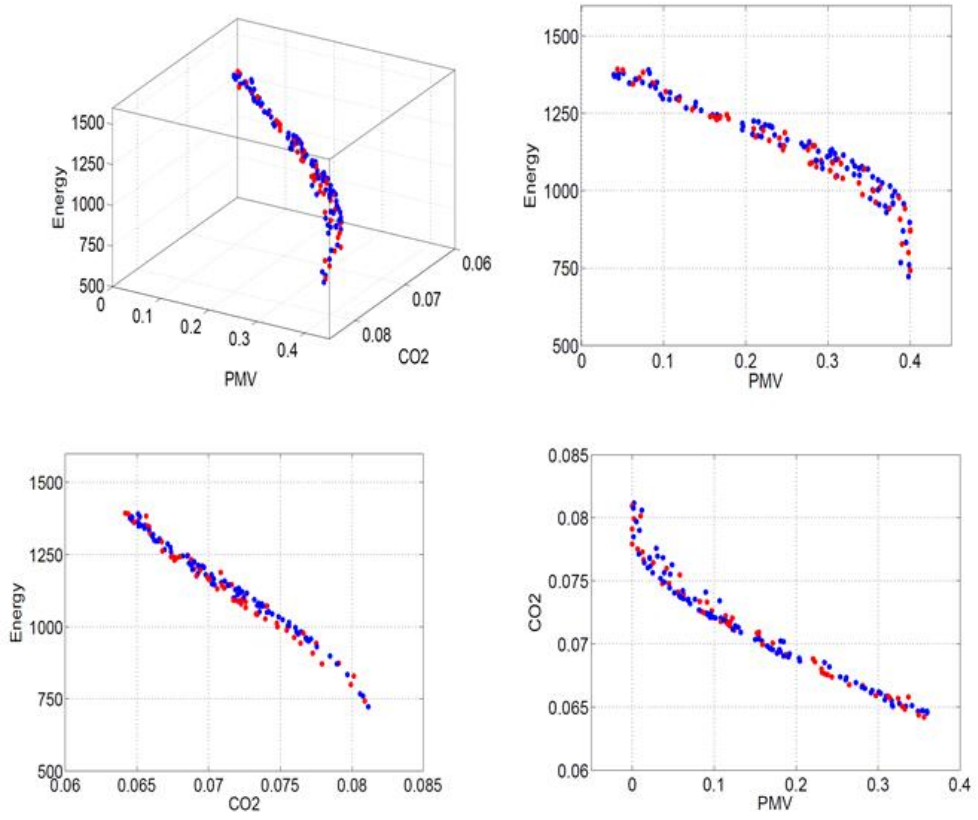

Figure 3.14 Comparisons of Pareto Fronts between using traditional CFD sampling and using adaptive CFD sampling (blue - traditional, red - adaptive).

A closer examination on the Kriging method in both cases is also presented here. Figure 3.15 shows the comparisons of the prediction design objectives (i.e. PMV, CO_2 concentrations and energy consumption) based on the uniform sampling data and the 13 sample data from adaptive sampling. As depicted, prediction differences between both cases are represented in the contour plot; while the probability distribution of the optimal solutions were also presented in while contour lines. As depicted, majority of difference between both cases are located on the left half of the design space where most of the sampling points were eliminated by adaptive sampling procedure. This clearly exemplifies that the adaptive sampling procedure has strategically allocated the sampling points where the optimal solutions are highly likely to be found. While relative error could be higher in other region, the lack of sampling data did not affect the accuracy of the final optimal solutions.

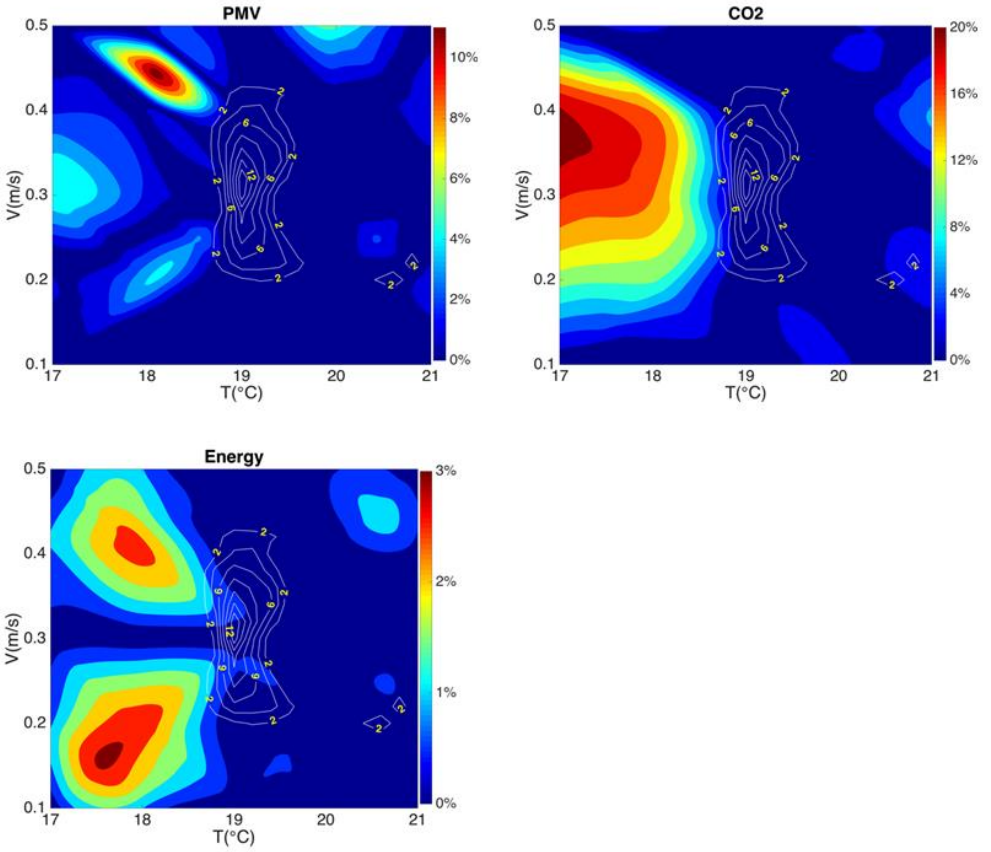

Figure 3.15 Contours of Kriging prediction differences between using 13 CFD samples and 25 CFD samples (PMV, CO2, Energy, respectively).

3.5 Conclusions

A multi-objective optimization platform has been proposed and developed by incorporating the Nondominated Sorting based Particle Swarm Optimization (NSPSO) algorithm with the Kriging method. To remedy the drawback of some previous studies, the NSPSO algorithm removes the necessity of using weighting factors in constructing the objective function and obtains the corresponding trade-off solutions (i.e. Pareto Front) for the given objective space. With the visualization of solutions in objective space, designers could easily pick up the most appropriate design solution according to their own judgments and preferences, rather than being struggled to decide the value of weighting factor in advance.

Special attention is also taken to minimize the computational cost where considerably large training sample based on Computational Fluid Dynamics (CFD) simulations are usually required for the surrogate modelling. The Kriging method where the best linear unbiased value of the unobserved fields is estimated based on the known sampled data is adopted in the present study. One particular advantage of the Kriging method is its capability in achieving high prediction accuracy with relatively small training sample size. Predictions from the Kriging method are compared and assessed with the CFD predictions. The comparison has shown that the Kriging method provides excellent accuracy in prediction with the maximum error of 5.12%. In addition, with the proposed adaptive sampling procedure, further reduction of computational cost could be realized. Based on the given case study, the optimization platform achieves a saving of 46.6% of CPU time without sacrificing the accuracy of the optimal solution.

In this study, for assessing the performance of the proposed optimization algorithm, a benchmark case study that has been validated by many researchers (Yuan et al. 1999, Zhou and Haghighat 2009a,b, Li et al. 2012, Li et al. 2013) was selected. To further examine the capacity of the algorithm, research work is currently carrying out to adopt and assess the proposed method with a more complex optimization case in related to the HVAC design of an airliner cabin occupied with passengers (Yan et al. 2014).

Chapter 4

Multi-fidelity surrogate algorithm for fire origin determination

The main findings of this chapter have been published in:

- **Li, N.,** Lee, E.W.M., Cheung, S. C. & Tu, J. (2019). Multi-fidelity surrogate algorithm for fire origin determination in compartment fires. *Engineering with Computers*. ***IF = 1.951, Q1***

With the significant growth of the world population, our cities are becoming more and more crowding. In this situation, any fire occurring would cause severe consequences, including property damage and human injuries or even deaths. In assessing the fire cause, the fire origin determination is a crucial step identifying origin of fire outbreak and the sequential fire and smoke propagation. Traditionally, fire investigators relied upon the visible fire damages at the fire scene to determine the location of fire originated based on their own professional experience. The fire origin determination process is however subject to the expert interpretation inherently embedded in the qualitative analyses. Aiming to develop an alternative methodology assisting the fire investigation, we proposed a new Multi-fidelity Kriging algorithm to quantitatively determine the fire origin based on the soot deposition patterns predicted by the numerical simulations. Advantage of the Multi-fidelity Kriging is its capacity in maintaining a reliable accurate prediction with very limited computational requirement in simulations. The algorithm is tested against a total of 41 different fire origins in a single compartment (i.e. 5 m width × 5 m length × 4 m height) with only one doorway for ventilation. The test results demonstrated that the Multi-fidelity Kriging algorithm is capable to predict the fire origin based on the simulated soot deposition pattern while posing significant saving to the computational cost by

correcting low-fidelity samples based on knowledge extract on high-fidelity simulation results.

4.1 Introduction

Fire origin determination is a challenging task in forensic sciences owing to the complex transient fire behaviour and its inherent nature in destroying evidences. One of the key hypothesis of the fire investigator is the correct identification of the fire origin (Association 2013). For most of the fire scenarios, the determination of fire origin is crucial for an accurate and reliable fire cause assessment. Today, with the development of computational technology, the application of fire simulation for fire scene reconstruction has gained measured success in the past decades. One of the representative studies is the fire scene reconstruction by fire modelling for the Station Nightclub fire occurred in Rhode Island of the United States on the 20th February 2003 (Bryner et al. 2007, Galea et al. 2008). The reconstructed fire scene revealed that the burning of the convoluted polyurethane foam caused the fire grew drastically within 90s after the ignition; generating large amount of smoke and hot gases. Another successful example is the reconstruction of the arson fire scene occurred in a 10 storeys hotel in Taoyuan city of Taiwan (Shen et al. 2008). The arson fire was lit with gasoline poured on the hallway carpet which accelerated the fire growth and spreading rate; restricting the evacuation of occupants and resulting 5 deaths and 10 injuries. Recently, a number of successful fire reconstruction studies have been carried for various serious fire cases using the latest fire modelling technique (Chi 2013, Yuen et al. 2014, Jahn et al. 2015). These studies clearly exemplified the feasibility of using computational fluid dynamics (CFD) to capture and the realistically reconstruct the fire scenarios. Nonetheless, fire reconstruction studies are efficient in forward engineering processes where fire consequences are predicted based on a known fire size and location. Fire origin determinations are unfortunately reserve engineering processes where fire consequences or damages are known except the fire size or locations. As each CFD simulation requires considerable computational time and resources, it is still impractical

relying CFD solely for fire investigation; especially for fire origin determination where a huge amount of simulations are required to support the reversed engineering process.

Several surrogate methods (also referred as meta methods) have been proposed in previous literatures as the fast alternatives to CFD simulation (Varol et al. 2007, Zhou and Haghighat 2009a,b, Stavrakakis et al. 2011, Li et al. 2012, Li et al. 2013). One of these methods is artificial neural network (ANN). Research works done by pioneers (Okayama 1991, Milke and Mcavoy 1995, Lee et al. 2000, Li et al. 2001) have proved that ANN is an efficient way to be applied in specific fire predictions. Afterwards, in recent years, more researchers have put their efforts on improving the performance of ANN for some particular scenarios (Lee et al. 2004, Yuen et al. 2006, Lo et al. 2009, Maeda et al. 2009, Erdem 2010, Safi and Bouroumi 2013). Among them, a good example is from Lee et al.'s work (Lee et al. 2004), which proposed a novel ANN fire model that is capable to process fire data with embedded noise. One of the main drawbacks is that ANN requires large amounts of data for model training to ensure reliable prediction accuracy. Previous literature proved that complex models like ANNs tend to be overfitting the training data due to the nature of the model, without sufficient training data, the model could yield inaccurate predictions for unseen data (Atiya and Ji 1997). In addition, for constructing the training data, high resolution mesh were adopted in the CFD simulations posing significant burden on the computational resources and time. Reducing the grid number (i.e. adopting a coarse mesh) could certainly save the computational time, while the accuracy and reliability would be sacrificed.

The fire investigator / forensic scientist is responsible to interpret and present the significance of the evidence to the justice system (Taroni et al. 2006). Without reliable testimony of witnesses or video records, the investigators are required to determine the fire origin based on the post-fire observations and expert interpretation of the scene evidence. Unfortunately, many of the interpretation and data analysis processes are implicit and subject to investigator bias. A survey of 586 fire investigators revealed that only 50% had an undergraduate qualification or higher, of

which only 18% were related to science or engineering (Tinsley and Gorbett 2013). The recent criminal arson case resulted in the execution of Cameron Todd Willingham in America exemplifies the irreversible consequence due to the failure of the investigator "to acknowledge or apply the contemporaneous understanding of the limitations of fire indicators" (Beyler 2009). Apart from providing evidence to the legal system, fire investigation also plays a significant role in identifying potential fire hazard due to improper designs or faulty devices and products.

In this paper, at the first step in developing an alternative methodology for assisting the fire investigation, we proposed a new Multi-fidelity Kriging algorithm to quantitatively determine the fire origin based on the CFD numerical predictions. Unlike the traditional surrogate methods, the Multi-fidelity Kriging technique utilizes a few of high-fidelity CFD predictions and large numbers of low-fidelity CFD predictions to guarantee the prediction accuracy but limited computational time and resources. To demonstrate the concept of the algorithm, the scope of this paper focuses on a relatively small fire size corresponding to pre-flashover fire scenario.

4.2 Methodology

4.2.1 **Kriging Technique**

As demonstrated in previous studies (Yeoh et al. 2002a, Yeoh et al. 2002b, Lee et al. 2004), it is certainly possible to use numerical model predicting the fire consequence of several fire scenarios and matching it with the fire scene details to identify the plausible fire origin. As discussed above, unfortunately, re-constructing the fire scenarios using numerical fire model could be time consuming and computationally exhaustive. Surrogate techniques (also referred as Meta models in literatures) have been proposed as a fast alternative to numerical fire model, such as ANN (Lee et al. 2004, Yuen et al. 2006, Lo et al. 2009, Maeda et al. 2009, Erdem 2010, Safi and Bouroumi 2013) and Support Vector Machine (SVM) (Ko et

al. 2009, Zhao 2009, Zhao et al. 2011, Yang et al. 2012). However, to ensure accuracy, both ANN and SVM require large amounts of training data which still put a significant challenging to the computational cost. Alternatively, other than ANN and SVM, the Kriging method has aroused much attention due to its capability in achieving high prediction accuracy with relatively small training sample size. Kriging is one of the methods of interpolation, deriving from regionalized variable theory (Cressie 1990). It gives predictions according to the best linear unbiased estimator of the distribution of the process based on the spatial analysis on the measured data (Stein 2012). Different from complex neural networks, Kriging does not rely on large data set for training. In fact, it has the advantage of providing accurate predictions with limited measured data. Detailed formulations of Kriging have been given elsewhere in (Handcock and Stein 1993, Van Beers and Kleijnen 2004, Sivia and Skilling 2006, Wikle and Berliner 2007, Forrester et al. 2008). In this section, a brief introduction to Kriging is presented as followings.

Kriging was originally developed in geostatistics by Danie G. Krige (Krige 1951), to estimate the most likely distribution of gold based on samples from a few boreholes. It can also be used to predict a Gaussian process governed by prior covariance in other engineering fields. The basic idea of Kriging is to predict the value of a function at a given point by computing a weighted average of the known values of the function in the neighbourhood of the point, which is expressed as:

$$\hat{z}_o = \sum_{i=1}^{n} \lambda_i z_i \tag{4.1}$$

where $\hat{z}_o$ represents a local estimation at the data location (x_o, y_o), z_i is the known sampled value at the data location (x_i, y_i) and λ_i represents the weighting coefficient which can be calculated by minimizing the estimation variance:

$$\min_{\lambda_i} Var(\hat{z}_o - z_o) \tag{4.2}$$

subjects to the unbiased condition:

$$E(z_o - \hat{z}_o) = 0 \tag{4.3}$$

and the normalization condition:

$$\sum_{i=1}^{n} \lambda_i = 1 \tag{4.4}$$

The weighting coefficient λ_i in Equation (4.1) can be solved using a quasi-Newton optimization method or other similar algorithm (Gano et al. 2006). Finally, the prediction value at the unobserved location can be given by Equation (4.1).

4.2.2 **Multi-fidelity Kriging**

Based on the formulation, Kriging poses great saving of computational resources without going through the tedious simulation process as required in CFD technique. Nevertheless, similar to other methods, it still requires training data extracted from CFD simulation results. For each simulation result, to ensure the accuracy, it is well known that high mesh resolution is crucial to minimize the discretization and truncation errors in the numerical procedures. The high mesh resolution impose significant computational burden in constructing the training data. To remedy this problem yet retaining the same level of prediction accuracy, we will introduce a Multi-fidelity Kriging method where the training data could be extracted from a small sample of high quality mesh while majority of the rest are constructed by computationally efficient low resolution mesh.

Multi-fidelity Kriging, also known as Multivariate Kriging or Co-Kriging, was originally developed for mineral explorations where measurements of different ores are available (De Baar et al. 2015). It calculates estimates or predictions for a poorly sampled variable (low-fidelity data) with help of a well-sampled variable (high-fidelity data). A full derivation of Multi-fidelity Kriging can be found in (Forrester et al. 2007). Here we assume that we have two sets of data yield to Gaussian process:

$$Z_1(s) = \mu_1 + \varepsilon_1(s)$$
$$Z_2(s) = \mu_2 + \varepsilon_2(s) \tag{4.5}$$

where Z_1 represents the high-fidelity process and Z_2 represents the low-fidelity process, ε_1 and ε_2 are the random errors contained in the high-fidelity process and low-fidelity process, respectively. There exists autocorrelation for each process and cross-correlation between them. Multi-fidelity Kriging attempts to predict the high-fidelity process $Z_1(s_0)$ using information (autocorrelation and cross-correlation) in the covariate Z_2 to make a better prediction, as the following:

$$Z_1 = \rho Z_2 + Z_D \tag{4.6}$$

where ρ is regression coefficient and Z_D is a new process representing the difference between the high-fidelity process and the low-fidelity process, which will correct the low-fidelity data. The whole procedure of the proposed method in this paper is shown in Figure 4.1.

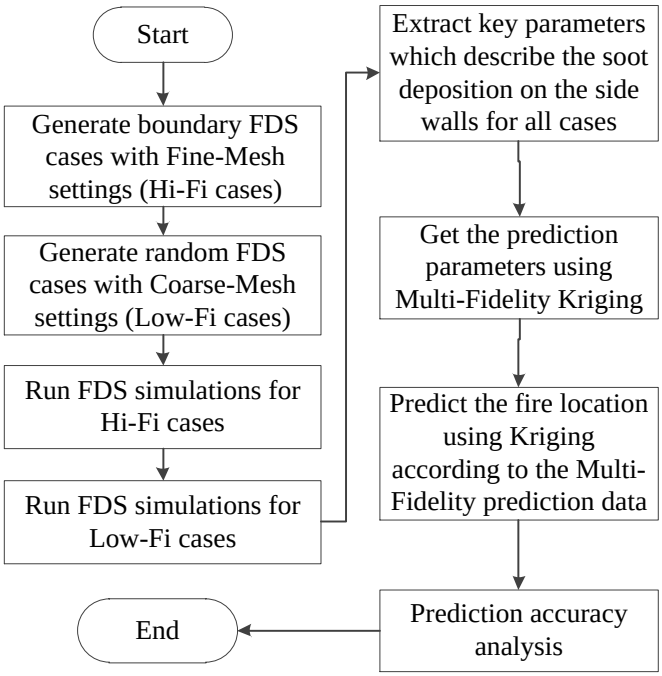

Figure 4.1 Framework of the proposed fire prediction method based on Multi-fidelity Kriging.

4.2.3 **Soot deposition model**

In most indoor fire cases, the fire materials do not burn cleanly due to organics contained in furniture, and thus generate large amounts of soot. Some of these soots are deposited on the walls and blacken the surfaces (see Figure 4.2 as an example). Since soot are mostly fine particles (i.e. range from 0.1 - 10 micron) which suspend in the hot smoke layer before deposition, soot deposition pattern could uniquely reflect the flow and distribution of smoke layer of each fire case and contains rich information of the fire behaviour. Riahi et al. (Riahi et al. 2013) investigated the smoke or soot deposition from the hot smoke layer onto wall surface using various fuels (i.e. PMMA, PP and gasoline). Their studies have shown that soot deposition pattern on the wall can be quantitatively measured using digital photographs and digital image analysis. Their measured data also validated against the predictions from a thermophoretic smoke deposition model with encouraged agreements. They concluded that the flame height and fire size could be calculated based on the clean zone area (i.e. area without soot deposition) on the wall. Their studies clearly ascertained that soot deposition pattern could be one of the effective evidences collected from post-fire investigate or predicted from numerical simulation for estimating the fire size and identifying the fire origin in fire investigation. In this study, soot deposition patterns were predicted based on numerical model using CFD technique. A brief description of the soot deposition model is summarized below.

Figure 4.2 An example of soot deposition patterns on the wall after fire is extinguished.

The smoke deposition can be driven by several factors including Brownian diffusion, thermophoresis, sedimentation, inertial impaction and turbulent diffusion. Previous studies (Talbot et al. 1980, Ciro et al. 2006, Cohan 2010, Hartman et al. 2012, Riahi et al. 2013) have shown that the thermophoretic soot deposition model is capable to give accurate prediction of smoke deposition process on walls during fire, by comparing the numerical simulation results against the experimental results. The governing equation in thermophoretic model is:

$$v_T = \frac{K_{th}\mu_{air}}{\rho_{air}T}\nabla T \tag{4.7}$$

where v_T is deposition velocity of soot particles, μ and ρ are properties of air at a film temperature. T is the gas temperature and K_{th} is a constant value, 0.55. The temperature gradient ∇T dominates the thermophoretic process which can be calculated by the following equation:

$$\nabla T = \frac{h(T_{gas} - T_{wall})}{k_{air}} \tag{4.8}$$

where h is the heat transfer coefficient and k_{air} is the thermal conductivity of air. Thus the soot mass deposited on a surface is the integration of thermophoretic velocity and optical density with respect to time:

$$m = A\int v_T \left(\frac{OD}{\sigma_{s,g}} \right) \mathrm{dt} \tag{4.9}$$

where A represents the area of the surface, OD is the optical density, and $\sigma_{s,g}$ is the gas phase mass specific extinction coefficient. Since FDS 5.5.1 provides a soot deposition model which has been verified by experimental data in both Gottuk (Gottuk et al. 2008) and Cohan's works (Cohan 2010), in this paper, we use this embedded model to predict soot deposition patterns.

4.2.4 **Workflow in post-fire investigation**

To help the readers understand how the proposed method would be used to help fire investigators locate the fire origin, Figure 4.3 depicts the workflow of utilizing the proposed multi-fidelity Kriging technique in practical fire investigation. Firstly, the investigators need to measure the geometry dimensions of the compartment which was on fire and record the smoke patterns on the walls. The next step is to reconstruct a series of fire scenarios in FDS where the fire origins can vary among all possible locations. The simulation results together with the corresponding fire origins could be used to train a Kriging model to provide investigators with prediction fire origin according to measured smoke patterns. Finally, we can validate the prediction result by running one more FDS simulation using the predicted fire origin and comparing the soot patterns from the simulation with those from the fire fields.

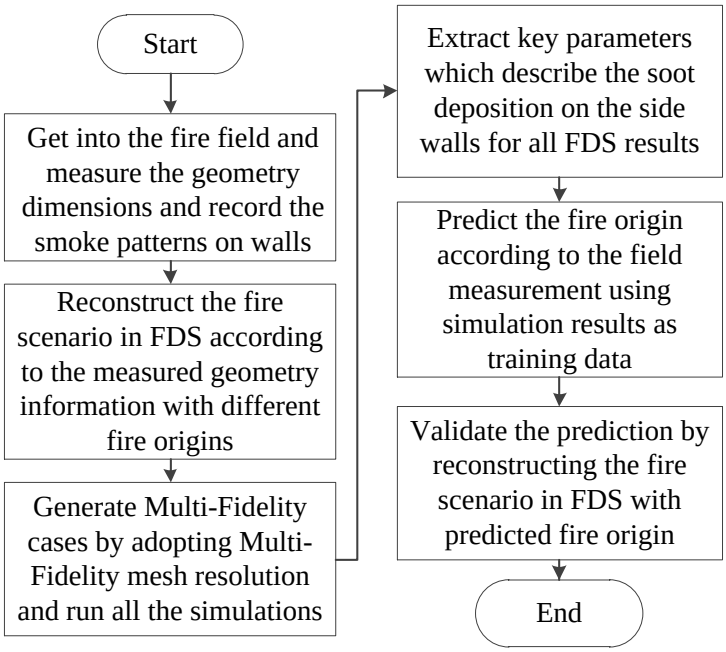

Figure 4.3 Workflow when the proposed method is used in post-fire investigation.

4.3 Preparation of Fire Scenario and Training data

4.3.1 Fire Scenario in a single compartment

To assess the feasibility of the multi-fidelity Kriging for fire investigation, this study focuses on a typical single compartment fire case. The typical single compartment (i.e. 5 m width × 5 m length × 4 m height) is considered with a single two-metre-high doorway as the only ventilation opening installed at the centre of the front wall. Detailed information about the model is listed in Table 4.1. Aiming to develop a non-biased and accurate fire origin determination algorithm based on the post-fire soot deposition pattern (location and magnitude), a group of fire scenarios were established where all other variables (such as geometry, fuels, heat release rate and duration) are assumed to be constant while a raised fire bed with the heat release rate of 204.3kW could be varied at any location on the floor level. Numerical simulations were conducted to simulate each fire scenario using the Fire Dynamic Simulator (FDS, version 5.5.1). Together with other field information, soot position pattern on each wall

were simulated for each fire scenario using the embedded soot deposition model in FDS. After each simulation completed, contour of soot deposition patterns on the three vertical walls (except the doorway wall) were then extracted together with the known fire origin location as the input of the training data for the multi-fidelity Kriging. More details regarding the numerical simulations are described in next section.

Table 4.1 Details of the single compartment fire model in FDS.

Name	Length(m)	Width(m)	Height(m)	Comments
Room	5.0	5.0	4.0	----------
Extended Region	5.0	5.0	8.0	----------
Doorway	0.2	1.0	2.0	----------
Fire bed	1.0	1.0	0.1	Propane
Wall	---	---	---	Isothermal

4.3.2 **Numerical modelling details**

Figure 4.4 shows the isometric view of the geometry arrangement of the computational domain for the numerical simulation. To isolate the end effect from the boundaries, an extended region was imposed beyond the wall boundaries of the single compartment connecting to the doorway. All boundaries of the extended region were specified as opening boundary condition where static pressure equals to atmospheric pressure (i.e. 101.3 kpa). All wall boundaries of the compartment were assumed to be isothermal. The raised fire sources with the heat release rate of 204.3kW were adopted for all fire cases. Figure 4.5 illustrates the location of the fire source for a total 41 different fire origin scenarios. As depicted, the fire scenarios consist of 16 fire origins uniformly allocated alongside with the wall (i.e. 0.5 m away from the nearest wall coloured as red dots in the figure). Another 25 fire origins were randomly allocated within the compartment floor area (see the blue dots in the figure). All fire origins were assumed to be a propane gas burner with a raised fire bed of 0.1 m above the floor level. For all fire scenarios, numerical simulations were

carried out from the ignition of fire until 1400 seconds where the flow field and soot deposition pattern had become steady and fully developed.

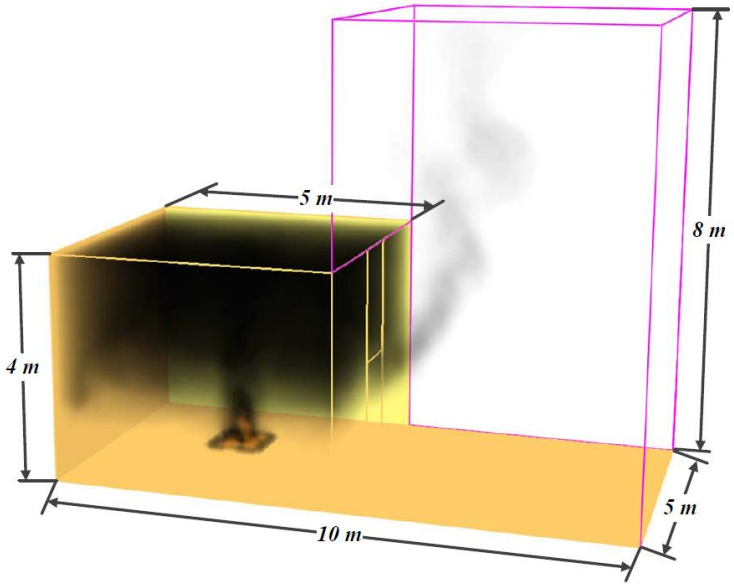

Figure 4.4 Model of single compartment fire in FDS.

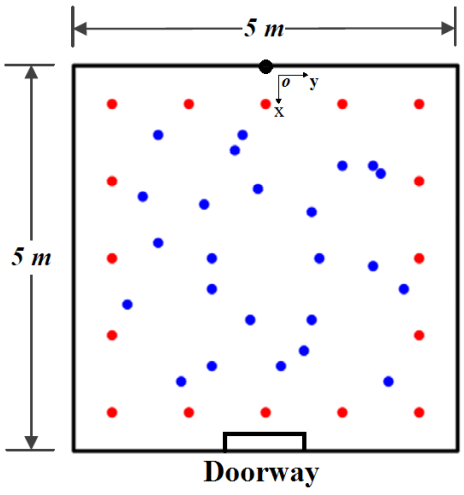

Figure 4.5 Sample locations of the fire. Red – with fine mesh setup. Blue – with coarse mesh setup.

In terms of the mesh setup, based on our previous works (Cheung and Yeoh 2009), to ensure the simulation accuracy, the grid size should range

between 4-16 times smaller than the characteristic length scale of a fire plume structure (Mcgrattan et al. 1998), which can be calculated as:

$$L^* = \left(\frac{Q}{\rho_{ref} T_{ref} C_p \sqrt{g}} \right)^{2/5} \tag{4.10}$$

where Q is the heat release rate of the fire. With respect to our design fire size, the characteristic length is estimated to be roughly 0.5m. As we mentioned before, in order to reduce computational cost, multi-fidelity cases are adopted. Therefore, we adopted 0.1m for the high-fidelity case and 0.2m for the low-fidelity case, which makes the total grid number in the high-fidelity case is 8 times of that in the low-fidelity case. In terms of the computational time, on a 2.1-GHz 8-core windows 7 machine, the high-fidelity case costs 24 hours to complete, while the low-fidelity case only consumes 1 hour to finish. Figure 4.6 shows a comparison of grid size between the high-fidelity case (fine mesh) and low-fidelity case (coarse mesh). The simulation results from these in total 41 cases will be used as multi-fidelity data to train the Multi-fidelity Kriging surrogate model.

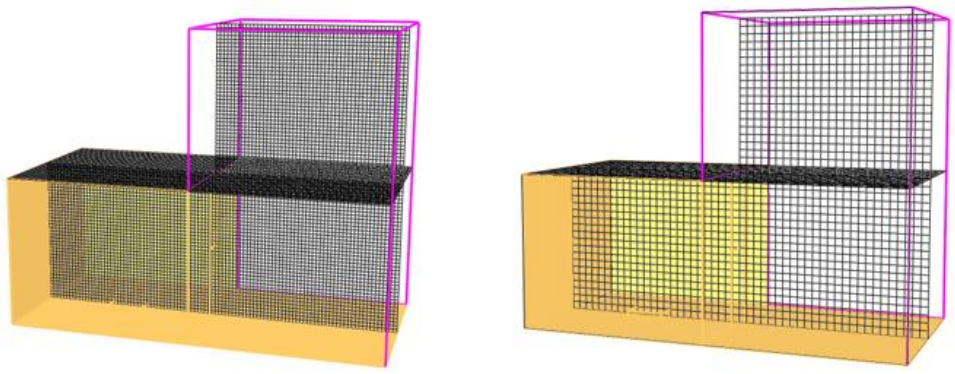

Figure 4.6 Comparison of grid size between high-fidelity case (50×50×40 grids) and low-fidelity case (25×25×20 grids).

The numerical model has been fully verified by a series of comprehensive tests in the authors' previous study (Cheung and Yeoh 2009), where the

numerical predictions of time-averaged velocity were compared and validated against experimental measurements. The purpose of validation is to verify the robustness of the mathematic models and numerical procedures in the simulation and its performance in capturing the turbulent air flow behaviours and heat transfer processes during fire. With the satisfactory agreement between numerical predictions and experimental data as shown in (Cheung and Yeoh 2009), the authors would claim that the adopted numerical model is valid and reliable.

4.3.3 **Soot deposition pattern parameterization**

As shown in Figure 4.2, after the fire is extinguished, the soot deposition patterns on wall surfaces clearly reflect the fire behaviour. We can simulate a similar soot deposition process in FDS using the embedded thermophoretic model. Figure 4.7 gives an example of soot deposition contours which is directly exported from FDS simulation results. However, the contours contain too much information and thus are too complicated to be surrogated using fast alternatives like Kriging. In addition, in reality, it is hard to measure the exact value of soot concentration while it is easy to get a boundary line of the soot pattern, as shown in Figure 4.2. Therefore, in this section, we will introduce a method which is capable to extract the boundary line from the soot deposition contour.

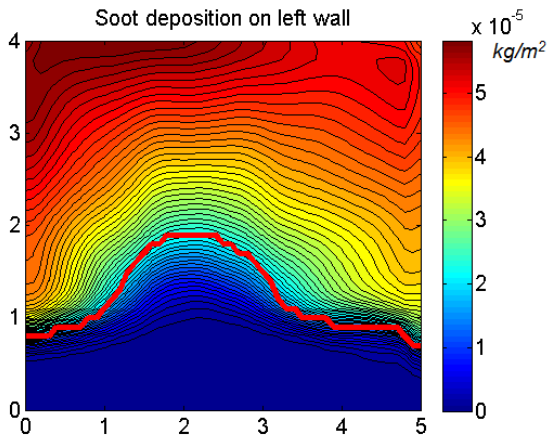

Figure 4.7 An example of extracting the boundary line of soot deposition. The

contour shows the soot deposition mass on the wall and the red line is the extracted boundary which can be used to describe the soot deposition pattern.

The approach developed by He et al. (He et al. 1998) is adopted to determine the boundary line. It is based on least-square techniques as shown the following equations:

$$\phi(y,H)=\begin{cases} p_l & for\ y<H \\ p_u & for\ y>H \end{cases} \tag{4.11}$$

where $p_l=\frac{1}{H}\int_0^H p(y)dy$ and $p_u=\frac{1}{H_r-H}\int_H^{H_r} p(y)dy$.

The height of the smoke layer above floor of a room is assumed to be H. The headroom is H_r. A function $\phi(y,H)$ separates the upper smoke layer and the lower clear layer by the following equation where p is taken to be the mass of soot deposited on wall. The least-square approach is to determine the height of the smoke layer, h, such that the squared error in equation (4.12) is minimized.

$$H=\min_h\left\{\frac{1}{H}\int_0^h\left[p(y)-p_l\right]^2dy+\frac{1}{H_r-H}\int_h^{H_r}\left[p(y)-p_u\right]^2dy\right\} \tag{4.12}$$

An example of such a process is shown in Figure 4.7. The soot deposition mass (kg/m2) is plotted as the contour in Figure 4.7 and the bold red line indicates the extracted boundary line representing the height of the smoke layer. After getting the smoke height, we then used 4 parameters (A,B,C,D) to describe that boundary line, where A is the mean height, B is the slope of the line, C and D are extracted from a second order fitting function of the line, see bellowing for the definition:

$$\begin{aligned} A &= mean(y) \\ B &= slope(y) \\ y &= Cx^2+Dx+\ldots \end{aligned} \tag{4.13}$$

In later prediction procedure, we used these extracted 4 parameters to describe the feature of the soot deposition on each wall.

4.4 Results and analysis

4.4.1 Performance validation of Multi-fidelity Kriging

The performance of the multi-fidelity Kriging is crucial in correlating the high fidelity CFD prediction with the low one without sacrificing the accuracy. Performance assessment for the multi-fidelity Kriging technique is firstly presented in this session. A one-dimensional wave function with low and high-fidelity attributes is adopted for the assessment which is given by the following equations:

$$\begin{aligned} f_{HF}(\mathrm{x}) &= 0.2\sin(3\pi\mathrm{x}) + 0.1*\cos(9\pi\mathrm{x}) + 0.3\cos(\pi\mathrm{x}) \\ f_{LF}(\mathrm{x}) &= 0.2\sin(3\pi\mathrm{x}) + 0.1*\cos(9\pi\mathrm{x}) \end{aligned} \tag{4.14}$$

Training samples were then extracted from the low and high fidelity function as an input for Kriging process. Figure 4.8(a) shows the true plot of f_{HF} (black solid line) and the Kriging prediction (red dash curve) with 5 samples, from which one can observe that there is a huge difference between the true values and the predicted values, i.e. in this case, the function has been under-sampled due to too small amounts of samples. One method to improve prediction accuracy is to add more samples (see as Figure 4.8(b) with 20 samples). However, these additional high-fidelity training samples would boost the computational cost significantly. Figure 4.8(c) shows the Multi-fidelity Kriging prediction of the same process f_{HF} but with only 5 high-fidelity training samples plus 20 training samples from the low-fidelity process f_{LF}. The comparison between Figure 4.8(b) and Figure 4.8(c) indicates the replacement of high-fidelity samples with low-fidelity samples does not affect the prediction accuracy but could help save lots of computational costs. Similarly, Figure 4.9 gives a 2D example, where the multi-fidelity process is defined by the following equations:

$$f_{HF}(x,y) = 0.2\sin(2\pi\, xy) + 0.1\cos[4\pi(x^2+y^2)] + 0.3\cos[\pi(x+y)]$$
$$f_{LF}(x,y) = 0.2\sin(2\pi\, xy) + 0.1\cos[4\pi(x^2+y^2)] \quad (4.15)$$

Figure 4.9(a) gives the true value for reference. Figure 4.9(b) and Figure 4.9(c) show the predictions by ordinary Kriging with 10 and 50 high-fidelity training samples, respectively and Figure 4.9(d) shows the Multi-fidelity Kriging results replacing 50 high-fidelity samples with 50 low-fidelity samples. From the comparisons of the contours in Figure 4.9, we can conclude that Multi-fidelity Kriging can give accurate predictions with a few of high-fidelity data plus a large amount of low-fidelity data, which helps relieve the computational burden significantly.

The above assessment clearly demonstrated the significant advantage of the Multi-fidelity Kriging in correlating the fine and coarse training data input and its capacity in producing high accuracy predictions with limit computational cost. On the other hand, it is worth noting that the Multi-fidelity Kriging works at best when there is a strong cross-correlation between the low and high fidelity data. In this study, the multi-fidelity data are extracted from the identical CFD package (i.e. FDS) with different mesh resolutions (i.e. high-fidelity data – fine mesh, low-fidelity data – coarse mesh). In other word, we could assume that predicted results were strongly cross-correlated with different level of discretization and truncation errors.

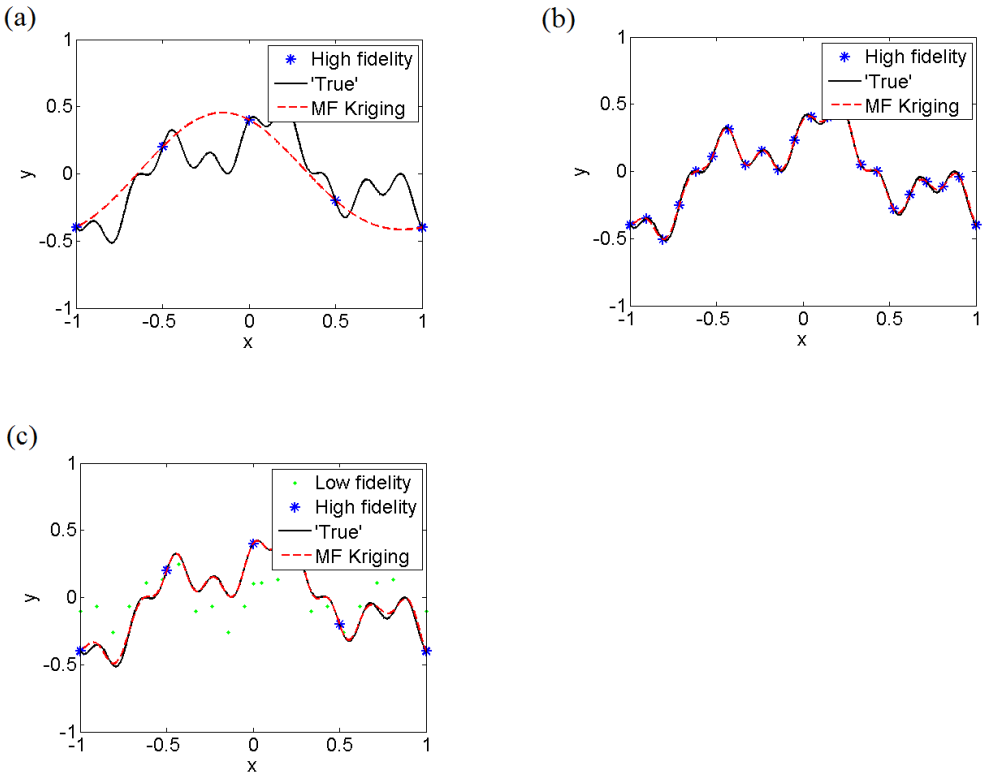

Figure 4.8 An example of 1-D Multi-fidelity Kriging prediction results. (a) The prediction using only a few of high-fidelity data is under sampled and there are huge differences between the prediction values and the true values. (b) The prediction using a large amount of high-fidelity data is well sampled. (c) The prediction using a few of high-fidelity data plus a large amount of low-fidelity data achieves good prediction accuracy, where the low-fidelity data are inaccurate but help to correct the prediction.

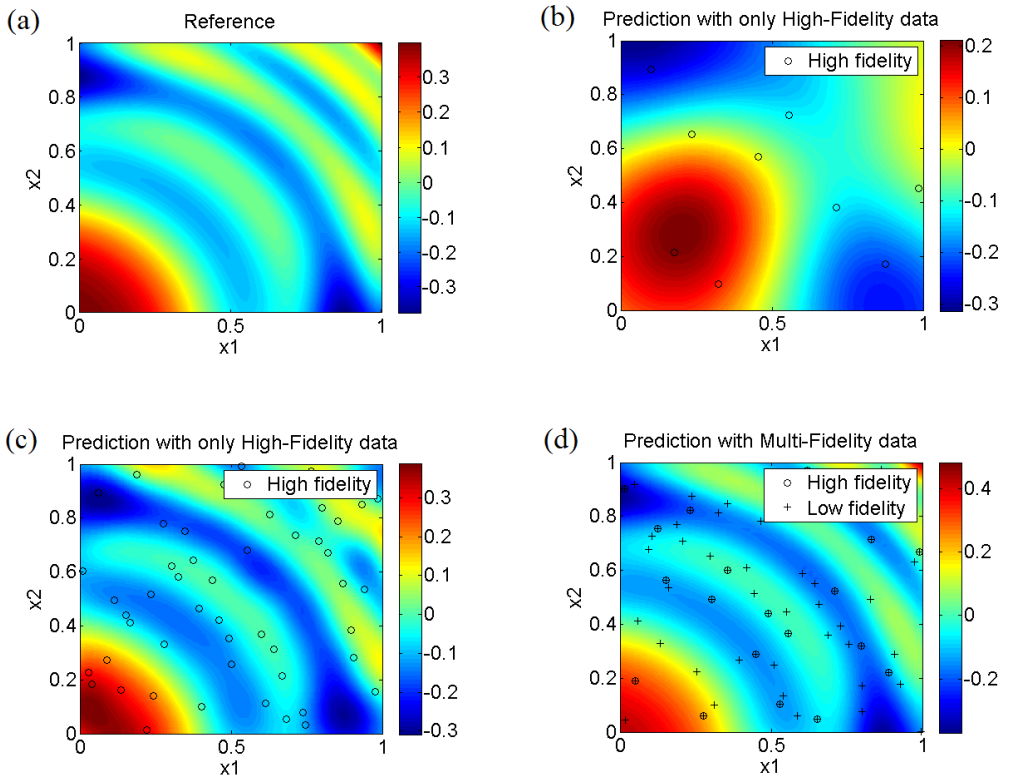

Figure 4.9 An example of 2-D Multi-fidelity Kriging prediction results. (a) Contours of true values for reference. (b) Contours of prediction values using only a few of high-fidelity, huge error observed. (c) Contours of prediction values using a large amount of high-fidelity, well sampled. (d) Contours of prediction values using a few of high-fidelity data plus a large amount of low-fidelity data.

4.4.2 **Fire induced flow pattern and corresponding soot deposition profiles**

Similar to the study by Gorbett et al. (Gorbett et al. 2017), numerical simulation were conducted using fire dynamics simulator (FDS) to develop a matrix of fire scenarios and its corresponding soot deposition. As mentioned before, analogue to the degree of damage used by Gorbett et al. (Gorbett et al. 2017), the soot deposition profiles based on variety of fire origin were predicted and extracted representing the fire consequence of each fire scenario.

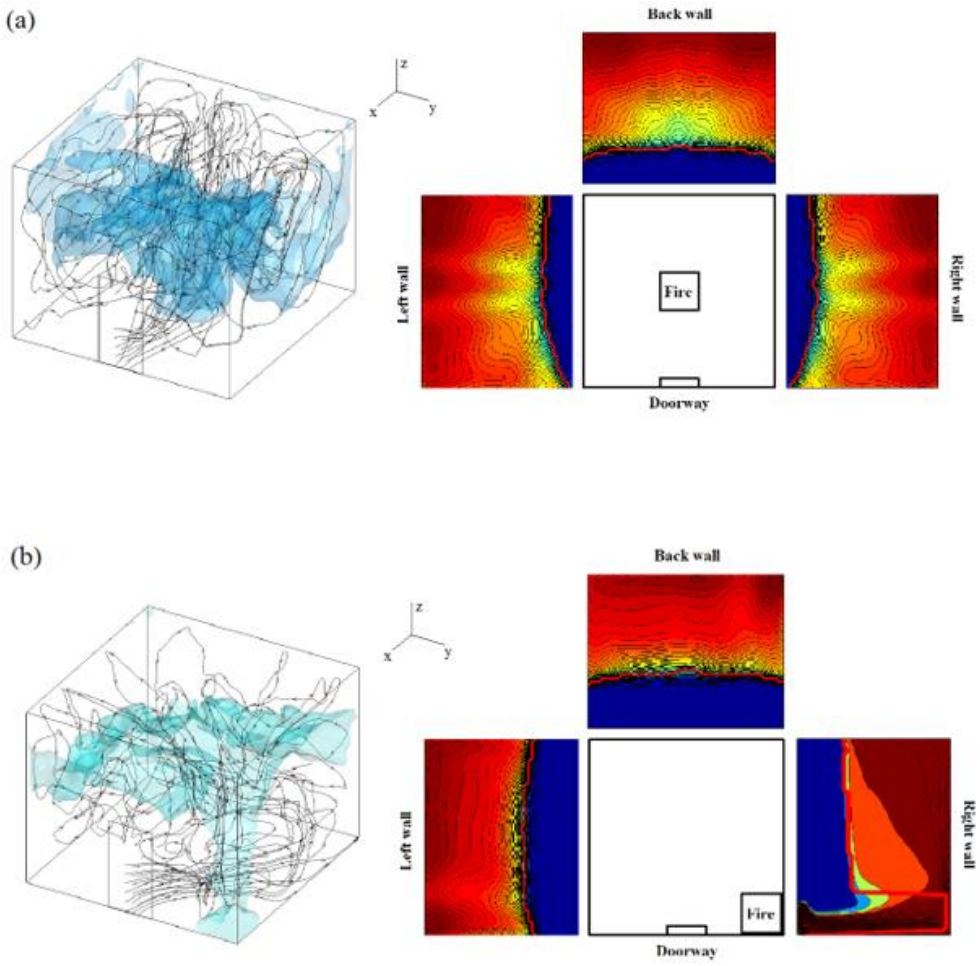

Figure 4.10 Comparisons of flow streamlines and soot deposition against different fire locations. (a) Fire is put in the centre. (b) Fire is put at the corner.

Figure 4.10 shows the predicted fire induced flow pattern and its corresponding soot deposition contour with respect to two different fire origin locations (i.e. central and corner fire). On the left of the figure, the isosurface of the hot smoke layer (i.e. at 378K) is depicted while the black streamlines trace the fire induced airflow pathway entering through the doorway, entraining into the fire and circulating at the upper layer of the room. Driven by the entrainment, different locations of the fire origin would result in different air flow pattern and hot layer distributions within the room; causing different soot deposition patterns on the walls. As shown in Figure 4.10(a), for the central fire case, cold air is entrained through the doorway towards the central fire marking roughly symmetric airflow distributions on both sides of the doorway (i.e. left and right) within the room. As shown in the contour, soot deposition pattern on both

walls are largely symmetric. The strong entrainment of air also push the fire and hot air towards the rear wall causing soot deposition level at the rear wall (see also in the figure). On the other hand, for the corner fire case as shown in Figure 4.10(b), the fire at the corner induces a strong fire plume and ceiling jet at the corner top level; causing higher concentrated soot deposition at the edge of right wall near the corner. The fire plume incur a ceiling jet carrying soot particles along the ceiling and causing soot deposition at the opposite corner (i.e. the rear and left walls). The figure clearly exemplifies that physical mechanism of fire induced air flow and its corresponding soot deposition pattern under different fire origin locations. The soot deposition pattern therefore could pose vital information for determining the fire origin locations. More importantly, such information could be obtained at the fire scene or predicted by validated numerical models.

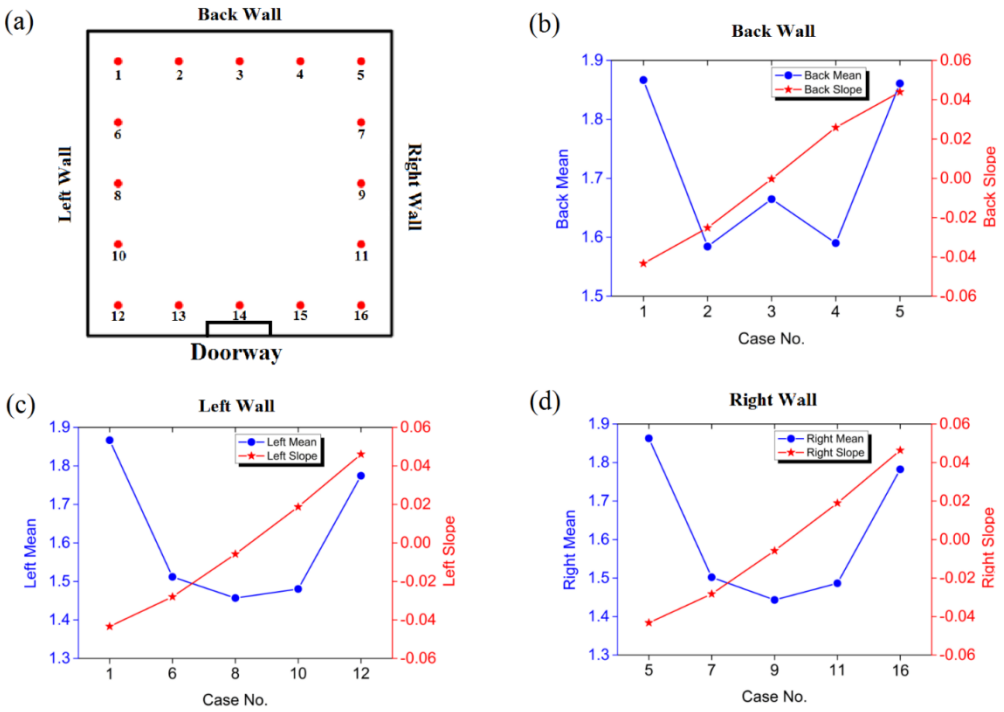

Figure 4.11 Comparisons of parameterized soot profiles among boundary cases.

Furthermore, Figure 4.11 shows the comparisons of parameterized soot profiles among boundary cases, indicating a strong relation between fire

origin and soot profile. Figure 4.11(a) labels all boundary cases with numbers. Figure 4.11(b) demonstrates the mean height (blue dots) and the slope (red stars) of the soot boundary line on the back wall in corresponding fire cases. As depicted in Figure 4.11(b), one could clearly observe that the mean smoke height appears almost symmetric with respect to the centre of the wall and the slope of the smoke layer increases with the corresponding fire locations. In case of the fire located on the left side of the back wall (i.e. location 1 and 2), a negative value of slope can be observed, representing the mean height of smoke layer on the left half wall is higher than that on the right half wall. An opposite trend can be observed while the fire is located on the right side (i.e. location 4 and 5), where the slope values turn to positive. In case 3, because the fire is located exactly in the centre, the slope of the smoke layer is zero, which is consistent with common knowledge. Regarding the mean smoke height (i.e. blue line), an extremely high level of smoke layer can be observed from case 1 and 5, which is because the smoke is driven by the fire plume at the corner flowing straight up towards the ceiling (see also in Figure 4.11(b)). Similar features can also be observed from Figure 4.11(c) and Figure 4.11(d) which demonstrate the soot profiles on the left wall and the right wall, respectively. It can be noted that the mean smoke heights in Figure 4.11(c and d) are not symmetric, which is caused by the impact of the airflow through the doorway. The above assessment clearly exemplified the relation between the soot deposition profiles and the locations of fire origin.

4.4.3 Prediction comparison between high and low-fidelity predictions

Although multi-fidelity Kriging technique could be adopted saving computational resource, the matter requires a strong correlation between the high and low fidelity training data. Figure 4.12 shows the comparisons of the soot deposition contours between the high-fidelity result (top) and the low-fidelity result (below). As depicted, one can observe that there are slight differences in the details of the contours (e.g. the yellow curve), however, in terms of the clean zone area (i.e. the blue parts of the

contour), both of the high-fidelity and low-fidelity predictions exhibit similar characteristic within the room. A closer examination on the soot deposition pattern on the three walls is shown in Figure 4.13. Following the discussion in Section 4.3.3, the boundaries of the clean zone area on the three walls were extracted (see the red curve in the figure). Furthermore, we compared the boundary lines extracted from the soot deposition contours using the method discussed in Section 4.3.3 and the results are shown in Figure 4.13. In the figure, the colorful contours show the soot deposition on side walls which are exported from FDS directly, and the red curves are the extracted lines. The comparisons indicate that the low-fidelity results are not 100% accurate but there is a strong positive cross-correlation between the high-fidelity results and the low-fidelity results, which is the basic requirement of the Multi-fidelity Kriging. More details about the cross-correlation will be discussed in next section.

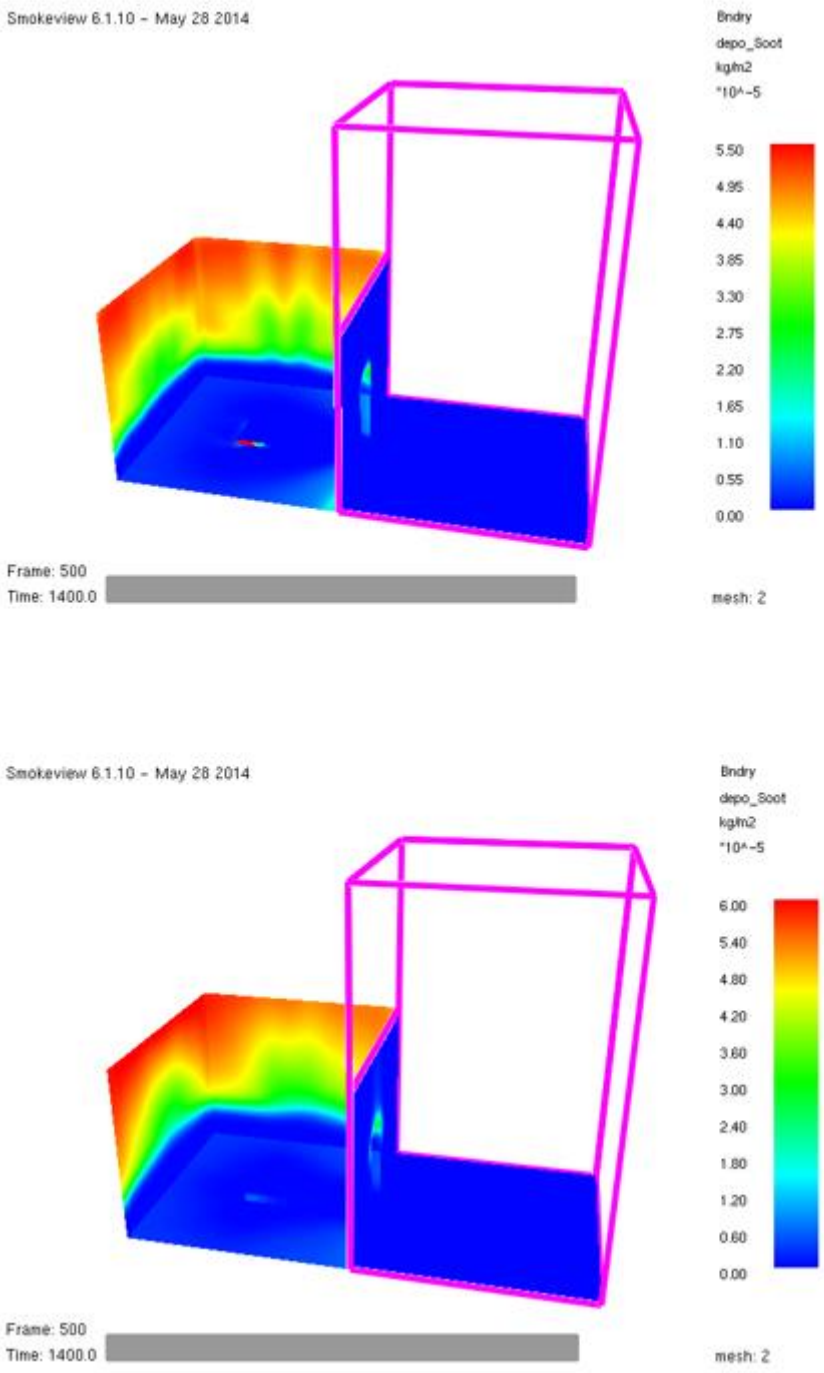

Figure 4.12 Contours of soot deposition on the room walls at 1400 seconds after burning of the fire which is put in the centre of the floor. Top: Results from high-fidelity case (fine mesh). Bottom: Results from low-fidelity case (coarse mesh).

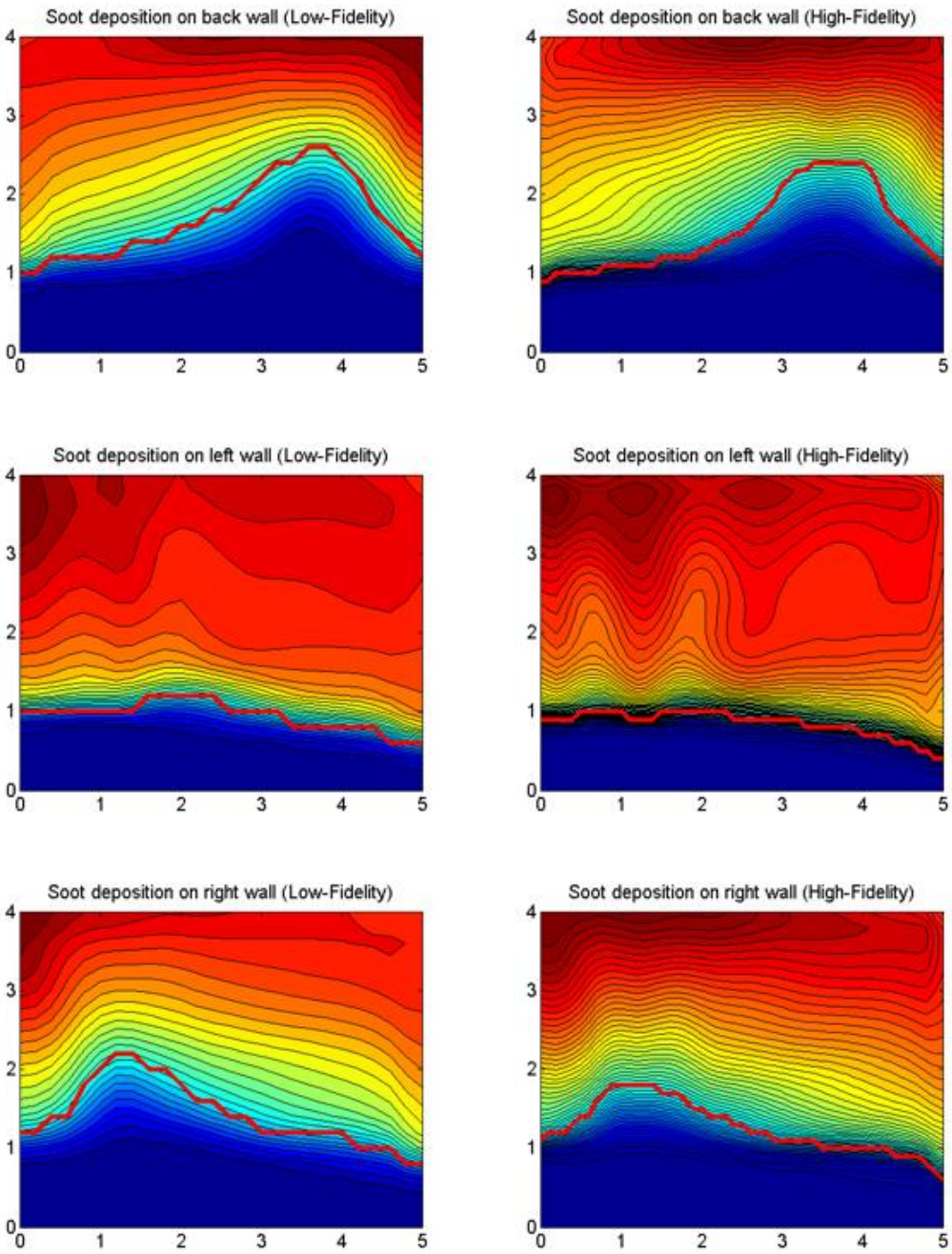

Figure 4.13 Comparisons of contours and extract lines of soot deposition on the walls at fire location (3.5, 1), in low-fidelity case and high-fidelity case.

4.4.4 **Prediction of fire origin**

A correlation tests for all the parameters extracted from the soot deposition pattern (section 4.3.3) between high-fidelity and low-fidelity cases were carried out according to (Lewis 1995). Table 4.2 shows the cross-correlation coefficients between the predicted the high-fidelity and

low-fidelity data. As shown in the table, all cross-correlation coefficients are of high value (i.e. greater 0.8) which clearly reflected that all parameters embedded a strong cross-correlation link between high-fidelity and low-fidelity data. In other word, both high-fidelity and low-fidelity data are reliable as a training database for the proposed Multi-fidelity Kriging algorithm.

Table 4.2 Cross-correlation coefficients between the high-fidelity data and the low-fidelity data.

Parameter	Left wall	Back wall	Right wall
A	0.8817	0.9590	0.8593
B	0.8589	0.9719	0.9120
C	0.8512	0.8270	0.8571
D	0.8467	0.8022	0.8130

As mentioned before, a series of fire scenarios with a total 41 different fire locations (i.e. 16 boundary cases and 25 random cases) within a single compartment were developed. The performance of the Multi-fidelity Kriging algorithm was assessed against the 41 fire scenarios database. In each prediction, one of the random cases was selected as the unseen data. The rest of the 40 cases (including 16 boundary cases and 24 random cases) were utilized for training the algorithm. All high-fidelity, multi-fidelity and low-fidelity data were utilized as the training sample (except the selected unseen case). Predictions were then made by the algorithm determining the fire origin location of the unseen case based on the unseen soot deposition pattern. Table 4.3 lists the absolute errors between the true location and the predicted fire origin location for all the 25 random cases. As expected, in general, predictions based on all high-fidelity database are in higher accuracy in comparison to those based on low-fidelity database. More importantly, predictions based on the multi-fidelity database retain almost identical accuracy compared to all high-fidelity database. A closer examination on the prediction accuracy can be found in Figure 4.14, which shows the cumulative probabilities calculated based on the normalized distribution of the prediction errors. As

depicted, in case one taking 95% as the confidence interval, the prediction errors of using high-fidelity training data, multi-fidelity training data and low-fidelity training data are 0.89m, 0.90m and 1.31m, respectively. By using only 16 high-fidelity simulation data, the algorithm could capture the correlation between the high-fidelity and low-fidelity simulation results and produce accurate predictions almost comparable to those made by all high-fidelity training samples. This clearly demonstrates the robustness of the algorithm. More importantly, without scarifying the accuracy, the algorithm poses significant computational saving in constructing the training samples. In this study, since only 16 out of 41 cases require high-fidelity simulation results, the multi-fidelity Kriging algorithm managed to reduce the associated computational time by 58%.

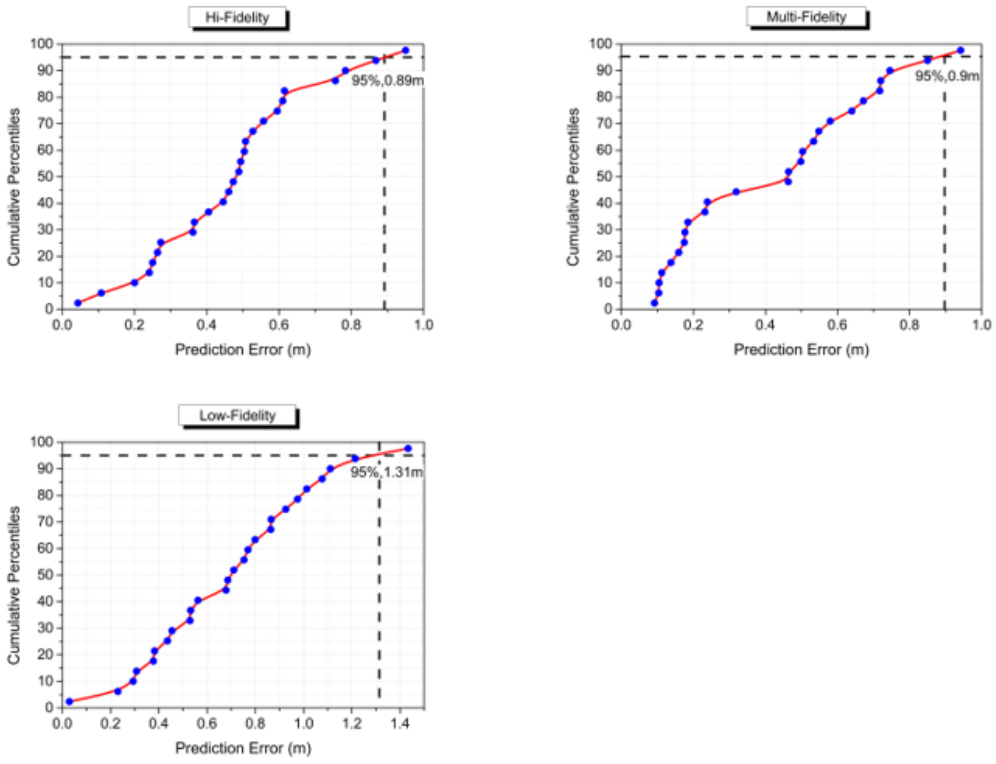

Figure 4.14 Cumulative probability plots of prediction error (high-fidelity, multi-fidelity and low-fidelity, respectively).

Table 4.3 Absolute errors between the true location and the predicted location.

True location		High-Fidelity			Multi-Fidelity			Low-Fidelity		
x	y	Predicted location		Error (m)	Predicted location		Error (m)	Predicted location		Error (m)
2.4	0.4	2.40	0.13	0.27	2.19	0.31	0.23	3.31	0.06	0.98
1.1	1.6	1.07	1.63	0.04	1.16	1.69	0.11	2.11	1.60	1.01
1.4	-1.6	1.85	-1.45	0.47	0.73	-1.86	0.72	1.40	-0.92	0.68
3.5	1.2	3.02	1.34	0.50	3.94	0.83	0.58	3.53	1.51	0.31
4.1	-1.6	4.30	-0.76	0.87	4.14	-0.88	0.72	3.30	-0.69	1.21
1.8	0	1.99	0.07	0.20	1.98	0.05	0.18	2.25	0.01	0.45
3.2	0	3.59	0.40	0.56	3.87	0.00	0.67	3.21	0.03	0.03
4	1.1	4.24	0.53	0.62	4.25	1.49	0.46	3.67	0.91	0.38
3.9	-0.1	3.17	0.51	0.95	2.96	-0.03	0.94	3.23	0.04	0.69
4.3	-0.4	4.54	-0.94	0.60	3.56	0.02	0.85	2.94	0.05	1.43
0.7	-0.6	0.77	-0.95	0.36	1.14	-0.84	0.50	1.39	-1.47	1.11
0.9	0.8	1.41	0.68	0.53	1.16	1.22	0.50	1.46	0.82	0.56
1.1	0.2	1.32	0.09	0.24	1.35	0.39	0.32	1.13	-0.18	0.38
1.7	0.7	1.45	0.74	0.25	1.75	0.87	0.17	2.53	0.46	0.87
1.8	-1.4	1.38	-1.55	0.45	2.28	-1.17	0.53	2.65	-1.55	0.86
1.8	-0.4	1.74	-1.18	0.78	1.81	-0.26	0.14	1.63	0.29	0.71
2.2	1.6	2.46	1.22	0.46	1.96	1.21	0.46	1.87	1.19	0.53
2.3	-0.8	2.07	-1.25	0.51	2.36	-1.34	0.55	2.65	-1.82	1.08
2.4	0.9	2.02	1.03	0.40	2.50	0.90	0.10	2.54	1.08	0.23
3	-1.2	3.27	-0.79	0.49	3.05	-1.43	0.24	3.43	-1.24	0.44
2.1	1.4	1.49	1.32	0.61	2.84	1.49	0.75	2.89	1.28	0.80
3.1	-0.8	2.74	-0.72	0.37	3.22	-0.67	0.18	3.85	-0.75	0.75
3.1	0.6	3.17	0.52	0.11	2.95	0.66	0.16	2.52	0.09	0.77
3.9	1.2	3.61	0.80	0.49	3.89	1.29	0.09	4.07	0.29	0.93
2.7	-1.4	2.47	-0.68	0.76	2.55	-0.78	0.64	2.27	-1.09	0.53

4.5 Conclusions

A novel Multi-fidelity Kriging algorithm for fire forensic investigations was proposed in this paper, which is capable to quantitatively determine the fire origin based on the soot deposition patterns that could be measured in the post-fire fire scene or predicted by means of numerical simulations. Aiming to develop an alternative methodology assisting the fire investigation, the proposed methodology does not involve expert

interpretations and professional judgments in process, which could eliminate the inevitable subjective bias and human error in the determination process. A series of fire scenarios (i.e. totally 41 cases) based on a single compartment fire situation was utilized for assessing the performance of the proposed Multi-fidelity Kriging algorithm. Soot deposition pattern on the compartment wall were predicted by the widely adopted fire CFD package - Fire Dynamics Simulator (FDS). The predicted soot deposition patterns on three vertical wall surfaces were extracted and parameterized into 4 parameters describing the boundary line of the pattern. Performance assessment shows the algorithm could capture the correlation between the high-fidelity and low-fidelity simulation results and produce accurate predictions almost comparable to those made by all high-fidelity training samples. Taking 95% as the confidence interval, the prediction errors of using high-fidelity and multi-fidelity training data are of 0.89m and 0.90m, respectively. More importantly, without scarifying the accuracy, the multi-fidelity Kriging algorithm managed to reduce the associated computational time for constructing the training sample by 58%.

As a preliminary validation of the proposed approach, the authors built a single compartment fire model in this paper. Predictions of fire origin in complex geometries (e.g. multi-compartments) is subject to on-going research work. In this paper, we used propane as the burning material in the CFD model, which is a relative lightly sooting fuel, where the level of soot concentration is 10^{-5} kg/m^2. Most of the building fires involve burning heavily sooting materials such as foamed furniture and natural or synthetic carpets. Using propone as the primary fuel in this paper representing a conservative assumption in terms of soot deposition on the wall. In practical building fires, it is expected to have a more rigorous soot deposition on the walls. In addition, more scenarios with post-flashover fire will be studied in the future.

Chapter 5

Multi-objective optimization design of HST cabin ventilation system

The main findings of this chapter have been published in:

- **Li, N.,** Yang, L., Li, X., Tu, J. & Cheung, S. C. (2019). Multi-objective optimization design of high-speed train cabin ventilation system using particle swarm optimization and Multi-fidelity Kriging. *Building and Environment*, 155, 161-174. ***IF = 4.539, Q1***

Maintaining a high level of thermal comfort and indoor air quality for occupants while minimizing the system energy consumption is crucial for long-haul High-speed train cabins. The traditional way of handling the multi-objective problem relies on the "trial and error" design which involves lengthy manual design parameter adjustment and performance evaluation based on on-site measurements or analytical and empirical models. To shorten design optimization process, a multi-objective optimization platform has been developed using the nondominated sorting-based particle swarm optimization (NSPSO) algorithm for searching the trade-off optimal design of the ventilation system in a fully occupied high-speed train (HST) cabin. A computational model of the HST cabin occupied by four full rows of passengers was constructed using ANSYS Fluent. To ensure the accuracy of the CFD model, high resolution computational thermal manikins were adopted to simulate the thermal and pollutant dispersion under influence of the passengers. Different combinations of ventilation operation parameters were evaluated against its performance in terms of thermal comfort, air quality and energy consumption. Furthermore, to reduce the computational cost of constructing the training sample, a Multi-fidelity Kriging technique is also proposed a surrogate method in replacing the time-consuming CFD simulations while maintaining an acceptable accuracy. The result

demonstrates that the presented approach is capable to perform a multi-objective optimization for indoor ventilation system design and yield accurate Pareto-front result with up to 35.61% saving of computational time.

5.1 Introduction

Driven by the rapid expansion of aviation in a global scale, a number of researchers have conducted studies on the impact of the ventilation system on the aircraft cabin environment via both experimental and numerical techniques (Chen et al. 2013, Liu et al. 2013, Li et al. 2015, Li et al. 2018a, Li et al. 2018b, You et al. 2018). Meanwhile, high-speed trains (HST) has emerged as an alternative fast public transportation around the world due to their huge transport capacity and high efficiency (Yang et al. 2018). Different from aircraft cabin, trains operate on the ground where the associated heat transfer between the cabin and outside ambient are significantly different from the aircraft cabin. In addition, the air infiltration and transient pressure change in train cabin when another train passing by or the train going through the tunnel have a unique impact on the cabin environment. Unlike the airliner cabin, an HST cabin is a semi-enclosed space with high occupant density where passengers are normally staying inside for many hours throughout the journey. The heating, ventilation and air conditioning (HVAC) systems are critical in controlling the cabin environment which has a direct impact on the passengers' thermal comfort and health. On the other hand, while providing a comfortable environment for passengers, air conditioning system consumes around 75% of all other non-propulsive energy consumption in the cabin (Liu et al. 2015). The investigation of optimal design of the ventilation system in HST cabin to ensure passengers' thermal comfort and health while minimizing energy consumption is of significant importance to the railway industry.

To obtain an optimal design, in the conventional "trial and error" design cycle, system design parameters such as supply airflow rate, air temperature and humidity are manually adjusted and evaluated based on

on-site measurements or analytical and empirical models (Liu et al. 2015). Over the past decades, with the rapid development of computer technology, computational fluid dynamics (CFD) techniques have been widely adopted to predict the indoor air environment aiming to shorten the time and reduce the cost of the lengthy HVAC system design cycles in buildings (Ravikumar and Prakash 2009, Cardinale et al. 2010, Hiyama et al. 2010, Kochetov et al. 2015, Limane et al. 2018, Tian et al. 2018, Cao et al. 2014, Pu et al. 2014, Wang and Zhai 2016, Yan et al. 2016) and long-haul transportation cabins (Liu et al. 2012, Li et al. 2016, Liu et al. 2013, Kwon et al. 2009, Zhang and Li 2012, Konstantinov and Wagner 2014,2015). Without any doubt, compared with the on-site measurement or experimental analysis, the CFD technique appears to be a cost and time effective alternative for design optimization. Nevertheless, due to the nature of trial and error design process, a large number of simulations is usually required for covering the entire design space; leading to a significant computational time and resource.

In order to reduce the computational time, some surrogate techniques such as artificial neural network (ANN) (Zhou and Haghighat 2009a,b, Acikgoz et al. 2017, Bre et al. 2018), Support Vector Machine (SVM) (Mousa et al. 2017), Kriging (Li et al. 2017) and Proper Orthogonal Decomposition (POD) (Li et al. 2013) are employed as a fast alternative approach replacing the CFD simulation to approximate the nonlinear and complex behaviour of the indoor airflow. On the other hand, aiming to automate the trial and error process, evolutionary algorithms (EA) have been proposed and coupled with the CFD to search the globally optimal solution (Luh and Lin 2011, Li et al. 2013, Zhai et al. 2014). Although the CFD-EA coupled approach significantly reduces the required number of CFD simulations to reach optimal solution, it still requires a substantial amount of CFD simulations for training the surrogate models to construct a reliable response space for EA.

Furthermore, the evaluation of indoor environment can be a complex process where multiple evaluation criteria are normally involved such as air velocity and temperature near the occupants, contaminant concentrations, percentage dissatisfied of draft, age of air, total energy consumption, etc. The system design is therefore a multi-objective optimization process where trade-off relations among those design indices are usually needed to be considered. Especially, in terms of indoor thermal comfort evaluation, lots of research works have been done by Ricciardi's group (Buratti and Ricciardi 2009, Buratti et al. 2013, Nematchoua et al. 2014, Ricciardi and Buratti 2015, Buratti et al. 2016, Ricciardi et al. 2016). In most of the previous works, in order to handle multi-objective problem, all the design objectives are aggregated into one single objective function through pre-defined biased weighting factors (Laverge and Janssens 2013, Li et al. 2013). One of the major disadvantages of this method is that the optimal design could be sensitive to the weighting factors, thus different values of the weights could result in substantially different solutions. Therefore, the values of these weighting factors are very much dependent on professional knowledge and expert judgements. In addition, the optimization procedure provides only one optimal solution per simulation run, offering the designers no flexibility in selecting alternative solutions for striking a trade-off of the conflicting criteria.

In attempting to address the above shortcomings, we propose a novel design scheme where a nondominated sorting-based particle swarm optimization (NSPSO) algorithm is utilized to achieve multi-objective optimization without using any biased weights. Furthermore, in order to minimize the computational cost in the design process while without sacrificing the accuracy, an improved surrogate method - Multi-fidelity Kriging algorithm is established and adopted in the present study. This improved Kriging algorithm is capable to accurately capture the system characteristics extract from CFD simulations yet requires very limited computational time for constructing the training database. The proposed design approach is tested and verified on a full-size computational HST

model which is assumed to be fully occupied by passengers. The rest part of this paper will describe the details of the design procedure and discussions of the validation results.

5.2 Methodology

5.2.1 Particle swarm optimization (PSO) and nondominated sorting-based PSO

As one of the the nature-inspired computational intelligence techniques, the particle swarm optimization (PSO) has been widely adopted in solving many engineering optimization problems, especially for those non-convex and multimodal engineering problems where traditional gradient-based mathematical programming methods have difficulties in finding the optimal solution (Deb 2001, Martínez and Coello 2013). Furthermore, previous study has also proved that the PSO has a faster convergence rate in comparison to other population-based stochastic optimization methods (e.g. genetic algorithms) (Hassan et al. 2005). As a brief history, the PSO was first introduced by (Kennedy 2001) based on the inspiration drawn from observations of the social behaviours of insects including learning from previous experience and communicating with successful individuals. In PSO, each particle has its own position and velocity, which are represented by *xi* and *vi*, respectively. The position and velocity of the particle are updated according to the following equations:

$$\begin{aligned} v_i(\mathrm{t}+1) &= \omega v_i(\mathrm{t}) + c_1\varphi_1(p_i - x_i(\mathrm{t})) + c_2\varphi_2(p_g - x_i(\mathrm{t})) \\ x_i(\mathrm{t}+1) &= x_i(\mathrm{t}) + v_i(\mathrm{t}+1) \end{aligned} \tag{5.1}$$

where *pi* and *pg* represent the personal best position and global best position, respectively, and the *c1* and *c2* are two uniform random numbers within the range [0, 1]. The $\varphi 1$ and $\varphi 2$ are two constants which are usually set to 2. The parameter ω decreases with iterations within the range [0.4, 1.2]. To avoid going out of the search space, both the position and velocity are limited within boundaries, [*xmin*, *xmax*] and [*vmin*, *vmax*],

respectively. Nevertheless, it is worth noting that the original PSO can only provide solutions for single-objective optimization problems.

Inspired by the works done by Deb (Deb et al. 2002), X. Li proposed a Nondominated Sorting Method to extend the original PSO to multi-objective optimization problems (MOP) - namely Nondominated Sorting based Particle Swarm Optimization (NSPSO) (Li 2003a). In NSPSO, the updating equations for particle position and velocity remain unchanged, but the selection of the personal best and global best has been re-designed. Two main mechanisms are used to determine the global best among the population – 1) nondominated sorting for identifying different fronts, and 2) crowding distance computed for particles within each front to encourage solution diversity. These kinds of information are used to select suitable leaders (i.e. global best) at each iteration to guide the particles moving towards the Pareto-optimal Front while still maintaining a good distribution of solutions along the Pareto-front.

a) *Nondominated Sorting*

Figure 5.1 shows an example of the nondominated sorting process. Considering 2 objectives (i.e. *f1* and *f2*) to be optimized in the process, the entire population (i.e. the 10 particles that labelled as 1 to 10) is sorted into different levels of fronts according to the domination comparisons between particles. The particles in same front are nondominated with each other. As depicted in Figure 5.1, Front 1 is the highest-level nondominated front because all particles in it are not dominated by any other particles in the entire population. The main goal of nondominated sorting is to classify the whole population into different levels of nondominated fronts. The global best particle (leader of the population) can be randomly selected from the highest-level front. This kind of selection process will push the whole population towards the true Pareto Front. More information regarding nondominated sorting can be found in (Li 2003a).

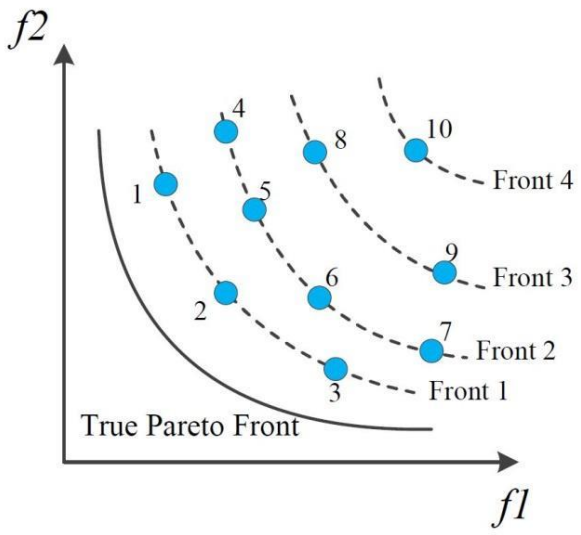

Figure 5.1 An example of nondominated sorting process in NSPSO.

b) Crowding distance

Different from a single-objective optimization, maintaining the diversity in a set of solutions is vital in a multi-objective optimization (Deb et al. 2002, Li 2003a). Throughout the optimization process, the leader must be selected properly to avoid local optimal aggregation of the whole population. In NSPSO, computing the crowding distance values among particles in the highest-level of nondominated front is used to select leaders that are both good and far apart from each other. Inspired by (Deb et al. 2002), we introduced a new way to calculate the crowding distance. Figure 5.2 shows an example of the crowding distance among particles. For each particle, the crowding distance is defined as the following:

$$D_n = \begin{cases} \infty & n = 1, N \\ d_{n-1} + d_n & 1 < n < N \end{cases} \tag{5.2}$$

The particle with a higher crowding distance value will have a high probability to be selected as the leader. Consequently, particles in the top front are likely to maintain a good level of population diversity.

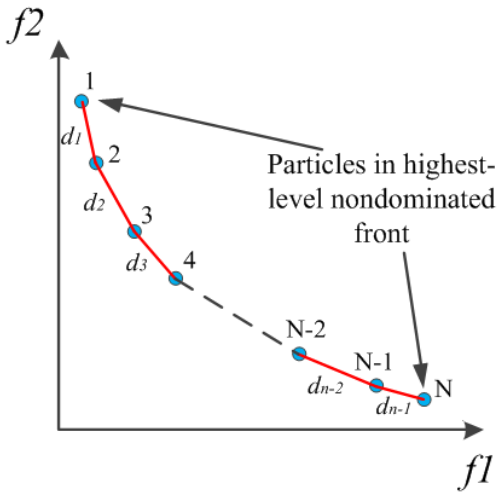

Figure 5.2 Crowding distances among individuals in the highest-level nondominated front.

Compared with other MOEAs, such as NSGAII and decomposition-based MOEAs, NSPSO has the advantages on running speed and convergence. As the main goal of this study is to save computational cost, NSPSO was selected to conduct multi-objective optimization.

5.2.2 **Kriging and Multi-fidelity Kriging**

As demonstrated in previous studies (Liu et al. 2013, Tian et al. 2018), it is certainly possible to use numerical model predicting the airflow behaviours inside buildings or vehicle cabins. However, as discussed above, the nature of CFD simulations is inherently time consuming and computationally exhaustive, making it impractical for design optimization. Surrogate techniques (also referred as Meta methods in literatures) have been adopted as a fast alternative to the CFD simulations, such as ANN (Zhou and Haghighat 2009a,b) and SVM (Mousa et al. 2017). However, to ensure accuracy, both ANN and SVM require large amounts of training data which still put a significant challenging to the computational cost. Nonetheless, to achieve a reliable and accurate prediction, both ANN and SVM require large amounts of training data generated from numerical simulations. The construction of training database remains a significant computational burden jeopardising the progress of the optimization process. Alternatively, other than ANN and

SVM, the Kriging method has aroused much attention due to its capability in achieving high prediction accuracy with relatively small training sample size (Li et al. 2017). Detailed formulations of Kriging have been given elsewhere in (Handcock and Stein 1993, Van Beers and Kleijnen 2004, Sivia and Skilling 2006, Wikle and Berliner 2007, Forrester et al. 2008). In this section, a brief introduction to Kriging is presented as followings.

Kriging was originally developed in geostatistics by Danie G. Krige (Krige 1951), to estimate the most likely distribution of gold based on samples from a few boreholes. It can also be used to predict a Gaussian process governed by prior covariance in other engineering fields. The basic idea of Kriging is to predict the value of a function at a given point by computing a weighted average of the known values of the function in the neighbourhood of the point, which is expressed as:

$$\hat{z}_o = \sum_{i=1}^{n} \lambda_i z_i \tag{5.1}$$

where $\hat{z}_o$ represents a local estimation at the data location (x_o, y_o), z_i is the known sampled value at the data location (x_i, y_i) and λ_i represents the weighting coefficient which can be calculated by minimizing the estimation variance:

$$\min_{\lambda_i} Var(\hat{z}_o - \mathrm{z}_o) \tag{5.2}$$

subjects to the unbiased condition:

$$E(\mathrm{z}_o - \hat{z}_o) = 0 \tag{5.3}$$

and the normalization condition:

$$\sum_{i=1}^{n} \lambda_i = 1 \tag{5.4}$$

The weighting coefficient λ_i in Equation (5.1) can be solved using a quasi-Newton optimization method or other similar algorithm (Gano et al. 2006). Finally, the prediction value at the unobserved location can be given by Equation (5.1).

Based on the formulation, Kriging poses great saving of computational resources without going through the tedious simulation process as required in CFD technique. Nevertheless, similar to other methods, it still requires training data extracted from CFD simulation results. For each simulation result, to ensure the accuracy, it is well known that high mesh resolution is crucial to minimize the discretization and truncation errors in the numerical procedures. The high mesh resolution imposes significant computational burden in constructing the training data. To remedy this problem yet retaining the same level of prediction accuracy, we will introduce a Multi-fidelity Kriging method where the training data could be extracted from a small sample of high quality mesh while majority of the rest are constructed by computationally efficient low resolution mesh.

Multi-fidelity Kriging, also known as Multivariate Kriging or Co-Kriging, was originally developed for mineral explorations where measurements of different ores are available (De Baar et al. 2015). It calculates estimates or predictions for a poorly sampled variable (low-fidelity data) with help of a well-sampled variable (high-fidelity data). A full derivation of Multi-fidelity Kriging can be found in (Forrester et al. 2007). Here we assume that we have two sets of data yield to Gaussian process:

$$\begin{aligned} Z_1(s) &= \mu_1 + \varepsilon_1(s) \\ Z_2(s) &= \mu_2 + \varepsilon_2(s) \end{aligned} \tag{5.5}$$

where Z_1 represents the high-fidelity process and Z_2 represents the low-fidelity process, ε_1 and ε_2 are the random errors contained in the high-fidelity process and low-fidelity process, respectively. There exists autocorrelation for each process and cross-correlation between them. Multi-fidelity Kriging attempts to predict the high-fidelity process $Z_1(s_0)$

using information (autocorrelation and cross-correlation) in the covariate Z_2 to make a better prediction, as the following:

$$Z_1 = \rho Z_2 + Z_D \tag{5.6}$$

where ρ is regression coefficient and Z_D is a new process representing the difference between the high-fidelity process and the low-fidelity process, which will correct the low-fidelity data.

5.2.3 **Workflow of the design procedure**

The whole procedure of the proposed method in this paper is shown in Figure 5.3. At the beginning, we need to determine the design range of the input variables and decide sampling locations. Then we need to build the computational model of the design case. In the meshing process, a mesh independent analysis is required to decide the proper grid size for the fine mesh case and coarse mesh case. Next we run the CFD simulations to get Multi-fidelity training results. In the next stage, the multi-objective optimization algorithm is used to search for global optimal solutions (Pareto Front), where, in each iteration, the Multi-fidelity Kriging is used to calculate the objective values of each particle according to the training samples from the CFD simulation. Finally, the provided optimal solutions are validated by running CFD simulations with the optimal input setups. The details of the CFD model and optimization results will be discussed in later sections.

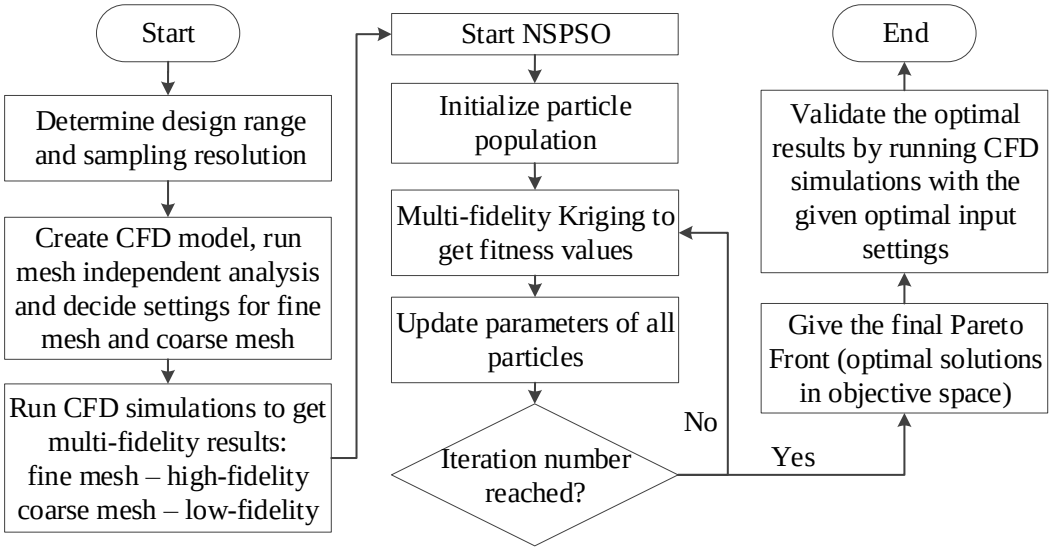

Figure 5.3 Framework of the CFD-based multi-objective optimization approach with Multi-fidelity Kriging.

5.3 Computational model description

5.3.1 **Model preparation and validation**

In this study, CRH2 model, a typical HST model in China (Wang et al. 2014), was chosen as the ventilation design environment. The CRH2 model cabin has the 3-2 seat arrangement (Figure 5.4). In order to save computational resources, instead of using a full HST cabin, four rows cabin section was built as the computational domain with the front and back plane being specified as translational periodic boundaries. The approximate periodicity of air flow field in train cabins has been verified in our previous studies (Yang et al. 2018). The HST is assumed to be operated in summer while the external ambient temperature is assumed to be 35℃. To consider both the heat conduction and heat convection from the outside to inside through the cabin envelope, an equivalent heat transfer coefficient of 2.2 W/(m^{2}K) was assigned for the material of cabin walls. The incident heat flux of solar radiation through the windows was 688.5W/m^2 (Yang et al. 2018). Other solid cabin walls including floor and seats were assumed to be adiabatic.

The cabin is assumed to be fully occupied by adult female passengers which are represented by the 3D-scanned manikin models. Figure 5.4 shows the detail of the computational domain. Radiative heat dissipation from the manikins surface is ignored and only convective heat transfer is considered in this study, which results in a total convective heat load of 30W (Yan et al. 2016) for each manikin. In addition, volatile organic compounds (VOCs) are assumed to be released from the surface of Passenger-S, who is seating in the second row of the second column (which is the second column). It should be noted that as a preliminary case study of the proposed approach, the authors assumed that only one passenger is releasing VOCs, to make the computational time feasible.

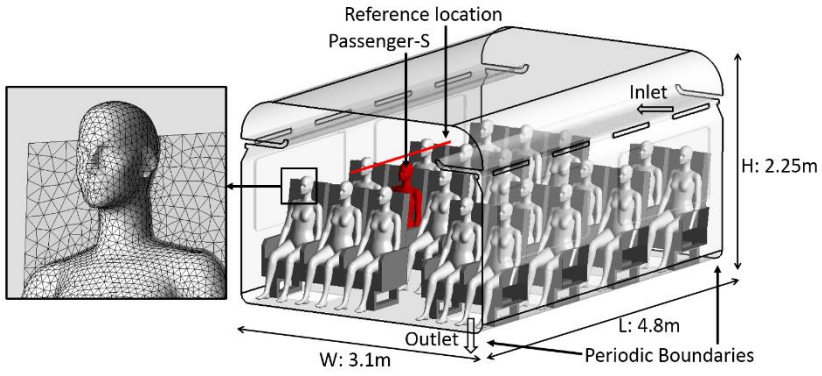

Figure 5.4 Cabin geometry and detailed meshing features.

The ventilation system of the HST cabin is assumed to be operating in compliance to the prevalent railway HVAC standard UIC 553-1 (Uic 2005). Air was supplied by inlet diffusers and exhausted by the outlets located under seats.

The commercial CFD software – ANSYS CFX 17.0 was employed as the numerical solver. The cabin airflow field was solved using the incompressible Navier-Stokes equations, where the generic transportation equation is given by:

$$\frac{\partial \emptyset}{\partial t} + \frac{\partial}{\partial x_i}\left(u_i \emptyset - \Gamma \frac{\partial \emptyset}{\partial x_i}\right) = S_\emptyset \qquad (5.7)$$

where $\emptyset$ is the general variable; u is the fluid velocity; Γ is diffusion coefficient; $S_{\emptyset}$ is the source term. The RNG k-epsilon turbulence model was adopted to model the air turbulence.

The above mathematic models and numerical procedures have been fully validated using a 7-row mock-up cabin (Figure 5.5). Figure 5.5a shows the inner look of the experimental cabin mock-up from literature (Li et al. 2018b), and the right image gives the corresponding geometry layout of the CFD model. Details of the experiment setup can be found in literature (Li et al. 2018b). The temperature and velocity profiles along three different vertical lines were compared against the airflow and temperature measurements in the mock cabin (Li et al. 2018b) (see Figure 5.5b). Furthermore, referring the study by Cao et al. (Cao et al. 2016), the statistical error of the PIV measurement for velocity fields range from 3% to 8%. Error bars representing an average 5.5% measurement error have also been included in Figure 5.5b. As depicted in Figure 5.5b, our numerical predictions have successfully captured the general trend and are in satisfactory agreement with the experimental data (Li et al. 2018b). One may notice that there are deviations between the predicted and measured results, especially for the velocity field. The averaged absolute prediction errors of temperature and velocity fields are also indicated in the figure. The maximum absolute errors are 0.79 °C and 0.074 m/s for the temperature and velocity field, respectively. The prediction error could be attributed to the simplification of the numerical model. As depicted in Figure 5.5a, the experimental measurements were conducted in a highly complex environment, involving many factors and geometrical details, such as the luggage geometry, the material difference, and non-uniform heat release rate of manikins. Such details in the physical mock-up model are unfortunately simplified in the computational model. The simplification could contribute to the minor deviations between the experimental data and the simulated results. Overall speaking, comparing our predictions with similar validation in literatures (Li et al. 2014a, Li et al. 2018b), our prediction accuracy is on the par with existing literatures predicting airflow and temperature in the indoor environment. Validating the predicted concentration field of VOCs is challenging due to the absence of experimental data in the literature.

Theoretically speaking, VOCs are gaseous contaminants where transportation in the air is subject to the convection and diffusion processes. In a convective-driven indoor space (e.g. HST train cabin), the transport of VOCs is predominantly controlled by the convective mechanism. Similar conclusion was also presented in a previous study by Zhuang et al (Zhuang et al. 2014). Therefore, with any better option, it is assumed that the VOCs dispersion is dominant by the convection current of the velocity field.

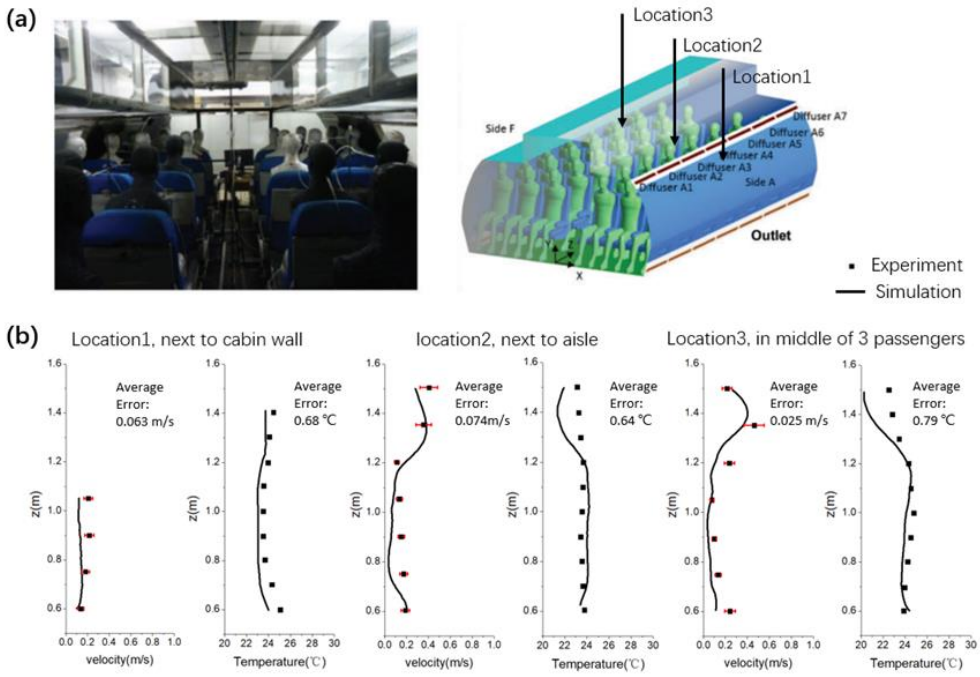

Figure 5.5 Validation of the computational model. (a) left: The inner look of the experimental cabin mock-up (Li et al. 2018b); right: The geometry layout of the CFD model. (b) Comparison of temperature and velocity between experimental and simulation results at different locations.

5.3.2 **Design objectives**

In order to assess the performance of the ventilation system inside a HST cabin, the Predicted Mean Vote (PMV), Contaminant concentration, and energy consumption were adopted as the quantitative indicators to evaluate thermal comfort, air quality and energy efficiency, respectively. The definitions of these design criteria are briefly introduced as follow.

a) *Predicted mean vote*

The predicted mean vote (PMV) is a thermal comfort evaluation index which was first introduced by (Fanger 1972). It is used to assess indoor thermal comfort based on heat balance and a set of experimental data collected from a given controlled climate chamber. The index represents the mean subjective satisfaction with the indoor thermal environment with a number between -3 (cold) and +3 (hot). The zero value is defined as the ideal representation of thermal neutrality. It can be calculated by the following equation, where the details can be found in (Fanger 1972).

$$\begin{aligned} PMV = {} & [0.303 * e^{-0.036M} + 0.028]\{M - W \\ & - 3.05 * 10^{-3}[5733 - 6.99(M - W) - P_a] - 0.42 \\ & [(M - W) - 58.15] - 1.7 * 10^{-5} M(5867 - P_a) \\ & - 0.0014M(34 - t_a) - 3.96 * 10^{-8} f_{cl}.[(t_{cl} + 273)^4 \\ & - (\overline{t_s} + 273)^4] - f_{cl} h_c (t_{cl} - t_a)\} \end{aligned} \tag{5.8}$$

The PMV index is evaluated based on an empirical equation which is correlated to the local air temperature, mean radiant temperature, relative humidity, air speed, metabolic rate, and clothing insulation (Fanger 1972). In this study, we assume that the occupants are seated in quiet position (i.e. metabolic rate of 1.0 met) with a summer clothing (i.e. 0.2 clo), and we evaluated the average PMV based on the predicted field information obtained from CFD simulations.

b) *Contaminant Concentration*

To assess the air quality within the space, the concentration of VOCs emitted by passengers throughout the cabin was also resolved in the CFD simulation. For presenting the VOCs contaminants emitted from human skin, acetaldehyde is chosen, as it has a relative high concentration level compared with other components in VOCs. The kinematic diffusivity of acetaldehyde is 1.6e-05 m2/s, and concentration measured from a normal human skin is 3.0e-5 kg/m-3. Due to the extremely low concentration of gaseous contaminants, the transportation process of VOCs was modeled using the drift-flux model:

$$\frac{\partial \varphi}{\partial t} + \frac{\partial}{\partial_{x_i}}(U\varphi) = \frac{\partial}{\partial_{x_i}}\left(D_\varphi \frac{\partial \varphi}{\partial_{x_i}}\right) + S_\varphi \tag{5.9}$$

where U is the fluid velocity in air-flow domain; φ is the concentration of contaminant; S_φ is a volumetric source term; D_φ is the kinematic diffusivity. Similar to the average PMV, the average VOCs concentration was extracted from the predicted CFD field information.

c) *Energy consumption*

Following the previous study (Zhou and Haghighat 2009a, Li et al. 2013), the energy consumption of the air-conditioning system is divided into two parts: ventilation fan power and the cooling or heating load. Energy consumption in the two parts is determined as follow:

$$\begin{aligned} E_{fan} &= \frac{P \cdot V_{air}}{\eta_{fan}} \\ E_{cooling/heating} &= m_{supply} c_p (T_{return} - T_{supply}) \\ &\quad + m_{outdoor} (h_{outdoor} - h_{return}) \\ E_{total} &= E_{fan} + E_{cooling/heating} \end{aligned} \tag{5.10}$$

where *P* is air pressure difference of the fan and *V* is volume flow rate of supply air (m3 s-1), m represents the mass flow rate of the air (kg s-1), c_p is the specific heat capacity of air, *T* represents temperature, *h* is the specific enthalpy of air (J kg^{-1}) which is related to air temperature and relative humidity. Similarly, we can get energy costs from the CFD-Post package.

5.4 Results and discussions

5.4.1 Design Optimization problem for the HST cabin ventilation

The design optimization of this study aims to allocate the best air ventilation conditions (i.e. supply air velocity and temperature setting) for the HST cabin of where multiple objectives (i.e. thermal comfort, air

quality and energy) are optimized. According to the UIC 553-1, inner temperature of the cabin should be set around 27 ℃ under the exterior environment condition set in this study. The supply air temperature and flow rate for optimization is evaluated based on the energy balance and the thermal load in the HST cabin. It is assumed that the thermal energy generated and absorbed inside a cabin is equal to the energy taken away by the airflow from inlet vents to outlet vents ($\dot{Q}_{air}$). The main sources of thermal load inside the cabin are the heat flux of human body ($\dot{Q}_{human}$) and thermal conduction of cabin wall ($\dot{Q}_{wall}$). The energy balance in the cabin can be expressed as follow:

$$\dot{Q}_{air} = \dot{Q}_{wall} + \dot{Q}_{human} \tag{5.11}$$

Based on the cabin energy balance, the fresh air supply rate at inlet diffusers is arranged from 20 to 24 m^3/h/person, the supply air temperature should be ranging from 295K to 299 K. The design ranges of the inlet temperature and mass flow rate should be 295 K to 299 K and 0.51 kg/s to 0.61 kg/s, respectively. In order to generate sufficient training samples for the Kriging prediction without significant computational burden, a total of 25 set of design parameters are firstly located and evenly distributed within the design space (see Figure 5.6). Sequentially, a total of 25 CFD simulations with different design parameters have been carried out with the identical fine mesh resolution and boundary conditions (excepts the supply air temperature and mass flow rate). All simulations were carried out in a computational domain consisting of a total of 5,951,997 nodal points. A mesh dependent study has been carried out and proven that the predicted results are mesh independent. More details are presented in the later session.

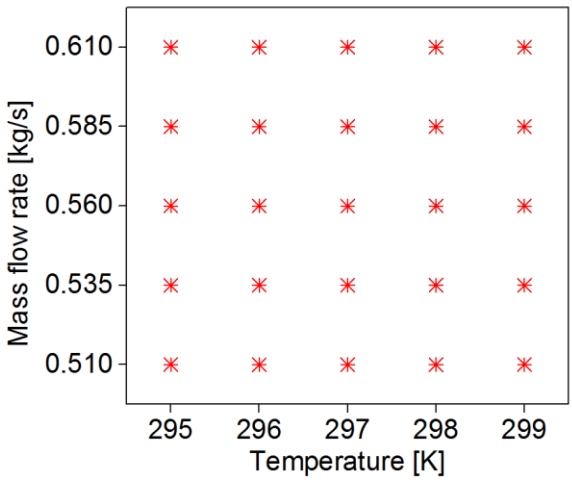

Figure 5.6 Definitions of inlet boundary conditions in CFD simulations.

After the simulations, the key air flow parameters (such as local air velocity, air temperature, contaminant concentration, etc.) were extracted from the CFD simulated result for evaluating the corresponding design objectives – PMV, contaminant concentration and energy consumption. Among the objectives, designers would like to create the most comfortable and cleanest environment and at the same time, to minimize the energy consumption of ventilation system. In reality, it is very common that these three objectives are conflicting with each other. In order to demonstrate the trade-off relationships, we listed three groups of typical results in Table 5.1, where the first row shows the global PMV minimum point (i.e. the most thermal comfortable) while both the contaminant concentration and the energy consumption are quite large. Similarly, the following rows give other situations where contaminant concentration and the energy consumption are the global optimal, respectively. As shown in the table, all three design objectives are conflicting each other where one could only optimize a trade-off solution. The CFD simulation results of these three cases are shown in Figure 5.7, Figure 5.8, Figure 5.9, respectively. In the figures, the left image gives the temperature distribution contour and velocity vectors on a section plane inside the cabin, and the right top indicates the PMV distributions around passengers' body surfaces. The right bottom one demonstrates the distribution of contaminant concentration near the VOCs release source (Passenger S). By comparing Figure 5.7 and Figure 5.8, it is clearly shown

that Case 1 has the better thermal comfort (i.e. Case 2 is too cold for the passengers) with the worse contaminant control. In contrast, in Case 2, contaminants are driven towards a higher elevation level of the cabin further away from breathing zone (see also in Figure 5.8). Based on the Table 5.1 and Figure 5.7, Figure 5.8, Figure 5.9, it can be observed that the design objectives are conflicting where a multi-objective optimization platform is required for obtaining the pareto-front of the ventilation design problem.

Table 5.1 Three typical groups of values minimizing PMV, Contaminant, Energy, respectively.

Case No.	T(°C)	V(kg/s)	PMV	Contaminant(kg/m3)	Energy(W)
1	24.45	0.5731	-1.5E-03	1.52E-07	1646.20
2	22.04	0.5938	-1.2E00	9.19E-08	2058.44
3	25.85	0.5100	1.5E-01	1.44E-07	1463.66

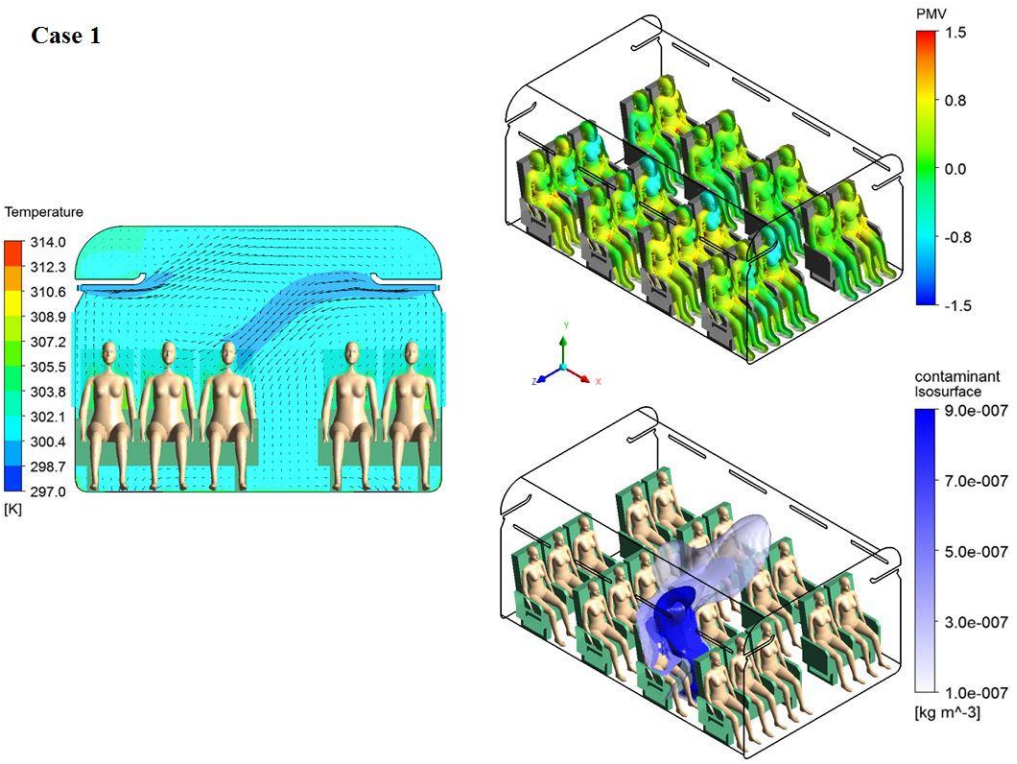

Figure 5.7 CFD simulation results of case 1 in Table 5.1.

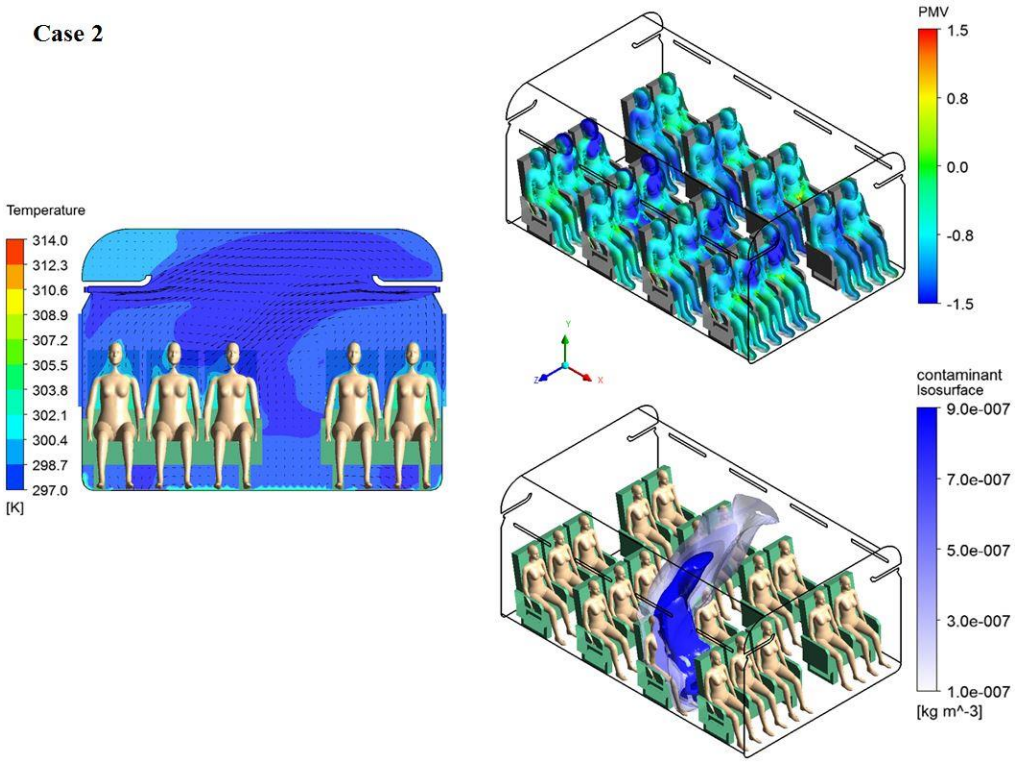

Figure 5.8 CFD simulation results of case 2 in Table 5.1.

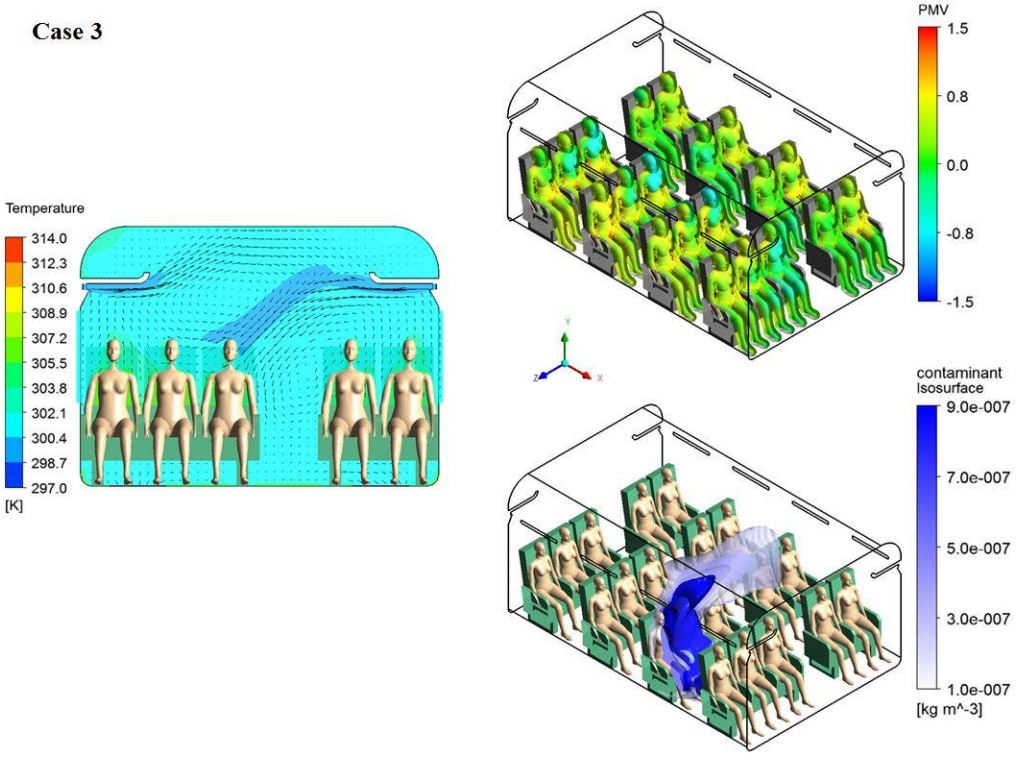

Figure 5.9 CFD simulation results of case 3 in Table 5.1.

5.4.2 **Validation of the High-fidelity Kriging Method**

To accelerate the optimization process, the High-fidelity Kriging algorithm is used as a fast alternative to the lengthy CFD simulations. In order to verify the accuracy of the Kriging prediction, a total of 16 additional ventilation conditions were simulated by the CFD model (i.e. the 16 green points within the design space in Figure 5.10) were carried for constructing a reliable respond surface of the objective values within the design space.

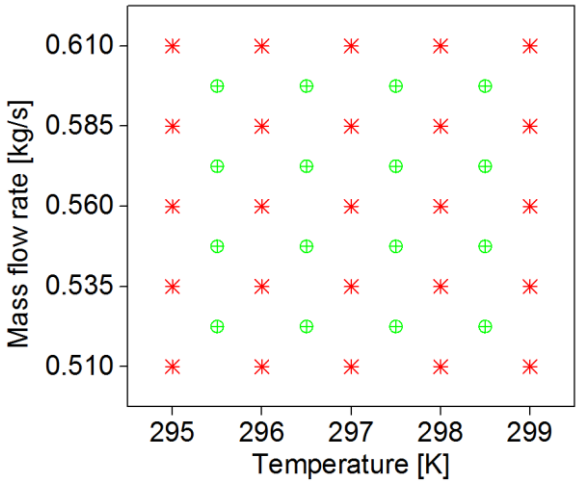

Figure 5.10 The additional CFD simulations and its corresponding ventilation conditions within the design space (i.e. green points).

The simulations conducted at red points in Figure 5.10 are used to train the Kriging algorithm. After the training, the Kriging method is used to predict the three objective values under the 16 different ventilation conditions (i.e. green points) and validated against the CFD predictions. The prediction error of the Kriging in comparison to the corresponding CFD results are shown in Figure 5.11. From the figure, one can observe that the maximum error is maintained below 5% for all design objectives. Especially, for energy prediction, the Kriging gives a prediction error below 0.6%, clearly indicating the robustness of Kriging surrogate

method in capturing the response surface of the objective value with respect to different ventilation conditions.

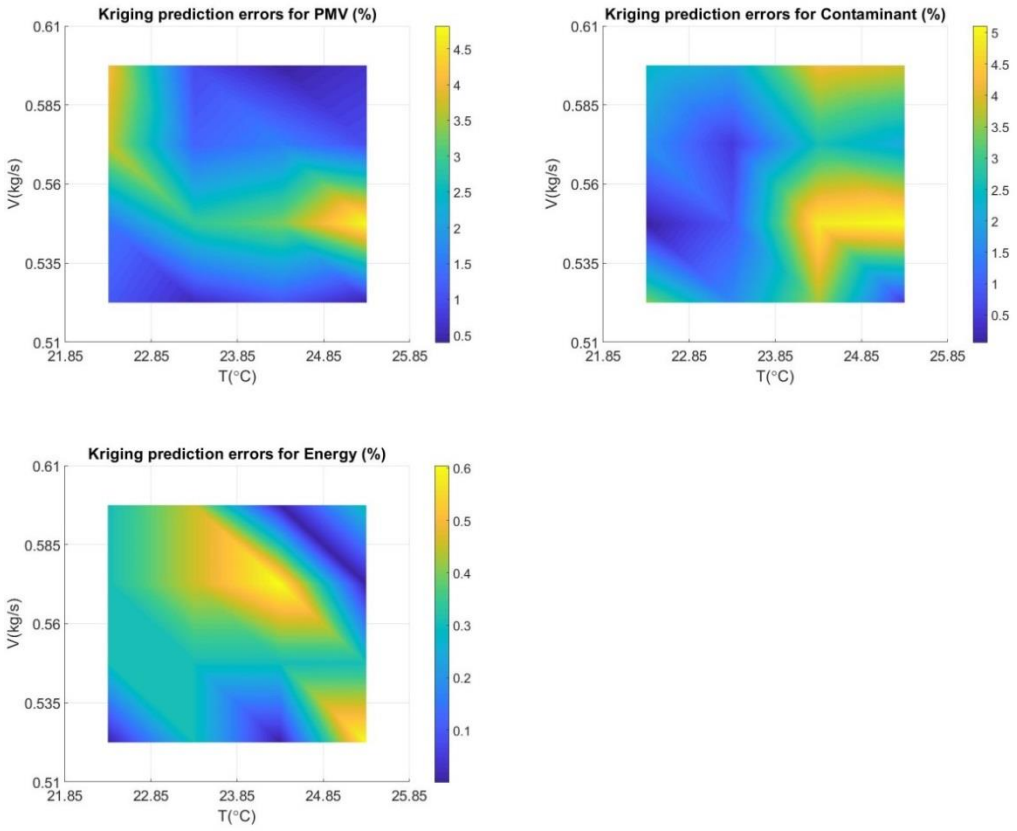

Figure 5.11 Percentage errors of Kriging prediction for the three objectives - PMV, Contaminant and Energy, respectively.

5.4.3 **Design Optimization using NSPSO with High-fidelity Kriging**

As mentioned above, a multi-objective optimization is carried out using NSPSO where the Kriging method is used for evaluating the objective values for replacing the CFD simulations. In this study, NSPSO algorithm has been established and implemented using MALAB R2017b. The algorithm considers two design input variables (i.e. supply air temperature and flow rate) and search the optimal design solutions based on the three objectives (i.e. PMV, energy and contaminant concentration). The size of swarm population was 100 and the maximum iteration

number was 200 (this configuration balances the accuracy and computational cost). In each generation, the validated Kriging surrogate was used to calculate the fitness values for the whole particle population. The final optimal trade-off solutions (i.e. pareto front) obtained from the NSPSO within the objective space are shown in Figure 5.12. Meanwhile, the corresponding input distribution of the pareto front within the design space is also illustrated in Figure 5.13.

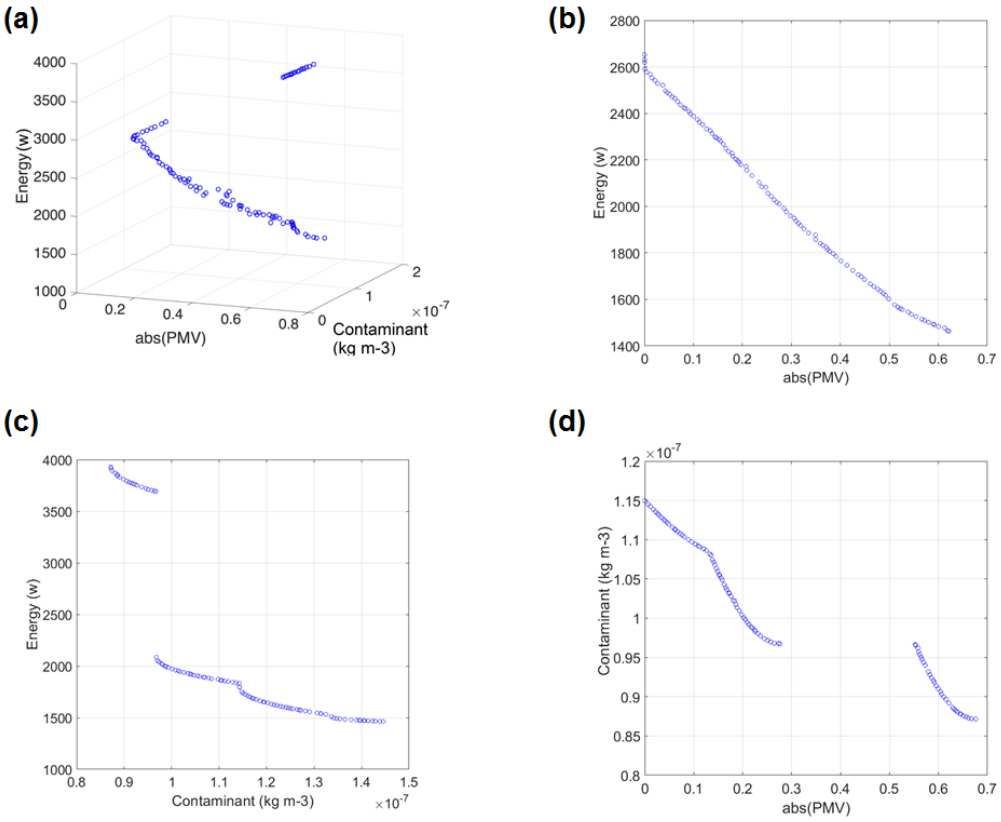

Figure 5.12 Optimal solutions ('Pareto Front') given by NSPSO in objective space.

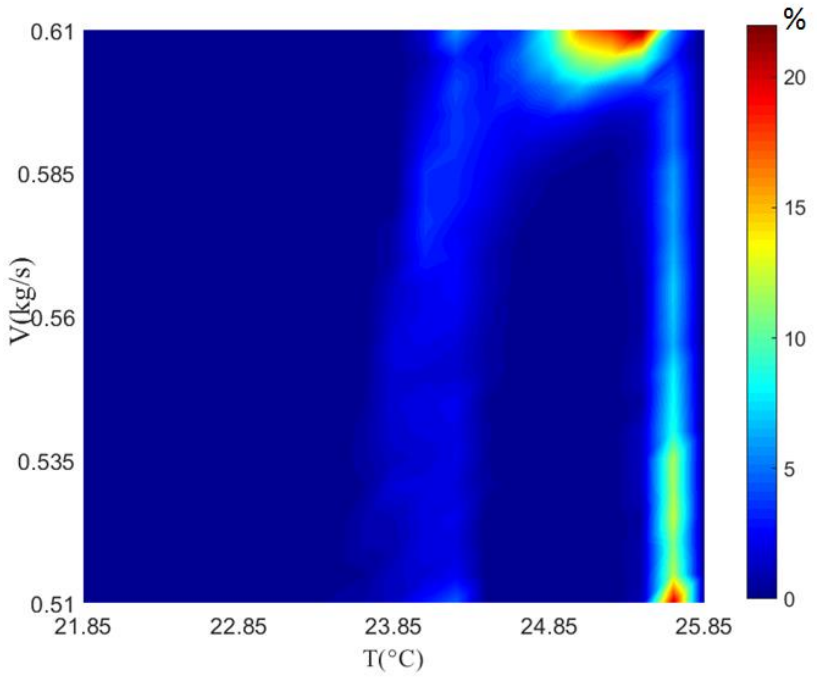

Figure 5.13 Probability density of optimal solutions in design space.

Figure 5.12 clearly shows the trade-off relationships among the design objectives, where Figure 5.12(a) gives a 3D Pareto front visualizing the trade-off among all the three objectives. For the ease of reading, the 2D Pareto fronts considering only two conflicting objectives are shown in Figure 5.12(b, c, d). In Figure 5.12, the pareto front represents a trade-off solution which are equally good in terms of multi-objective considerations. It is worth noting that the solutions listed in Table 5.1 are only three single points as part of the Pareto front. As demonstrated, the optimization framework using NSPSO and Kriging can entirely avoid using the weighting factor in objective function and obtains a set of trade-off solutions within the objective space in a fast, reliable and automated manner. The optimized solutions allow designers to evaluate the best options of the complex problem and choose the best design solutions based on their professional judgement and user preferences.

5.4.4 Enhancement in optimization process using Multi-fidelity Kriging

In previous literatures (Yuan et al. 1999, Lin et al. 2009, Zhou and Haghighat 2009a,b, Li et al. 2012, Li et al. 2013, Liu et al. 2015), the simplified manikin models such as rectangle or cylinder blocks are used in CFD simulations to study the indoor environment. As discussed in our previous paper (Yan et al. 2016), using simplified manikin model is unable to get an accurate CFD predictions. The prediction errors are found particularly jeopardize the manikin. In the present study, computational model using very fine 3D-scanned thermal manikin models (see Figure 5.4) was adopted to accurately capture the thermal response from passengers in the CFD simulations. Unfortunately, the finely spaced surface and volume mesh at the vicinity of the manikins have also dramatically increased the cost of computational resources. For example, for the given computational domain of this study, running a CFD steady state simulation using 2-core CPU at 2.9 GHz with the fine mesh requires roughly 37 hours to obtain a converged solution. For generating sufficient training samples for the Kriging method, a total of 25 simulations are required which translates into an approximately of 925 hours. It obviously poses a significant burden to the computational resources even with the modern parallel computing technology. In attempting to lower the computational time and cost, a Multi-Fidelity Kriging algorithm has been developed by using part of high-fidelity CFD results (fine mesh) and other part of low-fidelity CFD results (coarse mesh) as the training database.

To demonstrate the performance of the multi-fidelity Kriging, a one-dimensional wave function with low and high-fidelity attributes is adopted for the assessment which is given by the following equations:

$$\begin{aligned} f_{HF}(\mathrm{x}) &= 0.5\sin(10\pi \mathrm{x})/\mathrm{x} + (\mathrm{x}-1)^4 + 3\cos(\pi \mathrm{x}) + 2\log(\mathrm{x}) \\ f_{LF}(\mathrm{x}) &= 0.5\sin(10\pi \mathrm{x})/\mathrm{x} + (\mathrm{x}-1)^4 \end{aligned} \tag{5.12}$$

Training samples were then extracted from the low and high fidelity function as an input for Kriging process. Figure 5.14(a) shows the true

plot of f_{HF} (black solid line) and the Kriging prediction (red dash curve) with 5 samples, from which one can observe that there is a huge difference between the true values and the predicted values, i.e. in this case, the function has been under-sampled due to too small amounts of samples. One method to improve prediction accuracy is to add more samples (see as Figure 5.14(b) with 20 samples). However, these additional high-fidelity training samples would boost the computational cost significantly. Figure 5.14(c) shows the Multi-fidelity Kriging prediction of the same process f_{HF} but with only 5 high-fidelity training samples plus 20 training samples from the low-fidelity process f_{LF}. The comparison between Figure 5.14(b) and Figure 5.14(c) indicates the replacement of high-fidelity samples with low-fidelity samples does not affect the prediction accuracy but could help save lots of computational costs. The above assessment clearly demonstrated the significant advantage of the Multi-fidelity Kriging in correlating the fine and coarse training data input and its capacity in producing high accuracy predictions with limit computational cost.

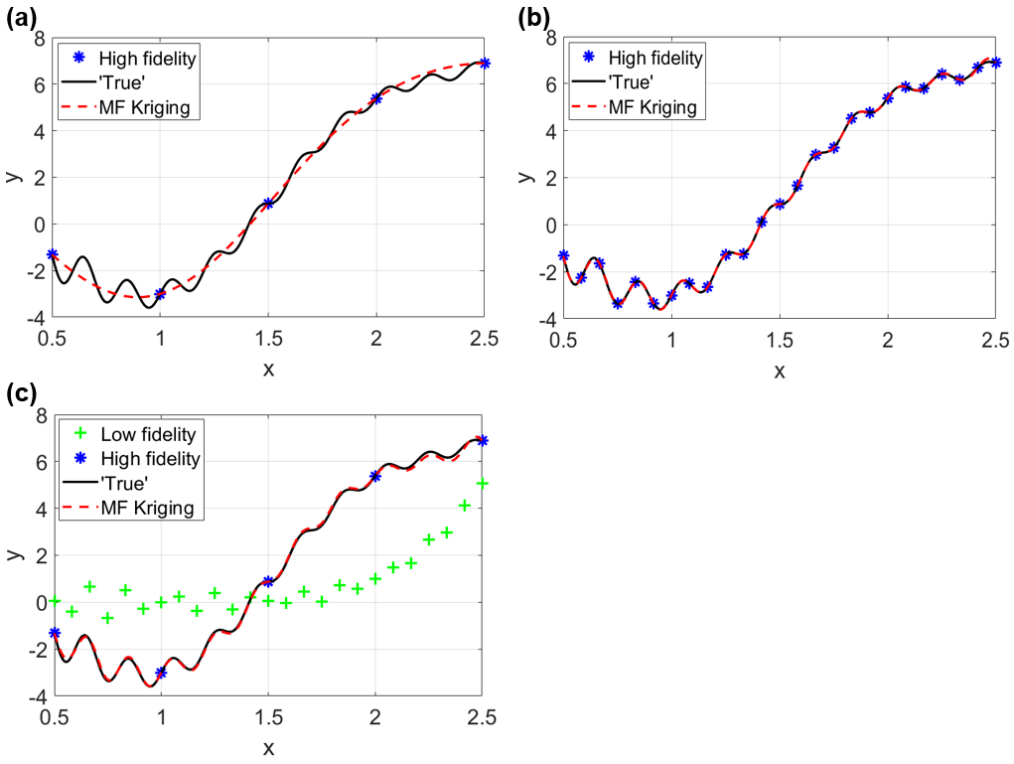

Figure 5.14 An example of 1-D Multi-fidelity Kriging prediction results. (a) The prediction using only a few of high-fidelity data is under sampled and there are huge differences between the prediction values and the true values. (b) The prediction using a large amount of high-fidelity data is well sampled. (c) The prediction using a few of high-fidelity data plus a large amount of low-fidelity data achieves good prediction accuracy, where the low-fidelity data are inaccurate but help to correct the prediction.

5.4.5 Multi-fidelity training samples for HST cabin ventilation problem

In this study, the cabin geometry model was discretized using unstructured computational mesh by ANSYS ICEM. To accurately capture the VOCs dispersion, high quality fine meshes were generated around diffusers and 3D-scanned manikins. Furthermore, for resolving the boundary layer, inflation layers with a grid expansion ration was applied around manikin surface. A grid sensitivity study was conducted over four different grid resolutions ranging from coarse mesh (i.e. ~2.8 million grids), good-mesh set (i.e. ~3.6 million grids), fine-mesh set (i.e. ~6.0 million grids) and extremely-fine-mesh set (i.e. ~7.7 million grids). The grid sensitivity study results of air flow velocity along the central line of

the cabin were illustrated in Figure 5.15. As depicted, the fine-mesh predictions have less than 5% difference comparing the one from extremely-fine-mesh set. The result clearly demonstrated that the fine mesh adopted for all high-fidelity CFD training mesh are grid independent. Meanwhile, it is clear that the predictions of the coarse mesh pose considerably significant errors (i.e. some of the points are over 300%) compared to the fine and extremely-fine mesh. To demonstrate the robustness of the Multi-fidelity Kriging, predictions from the coarse mesh are adopted as low-fidelity training samples.

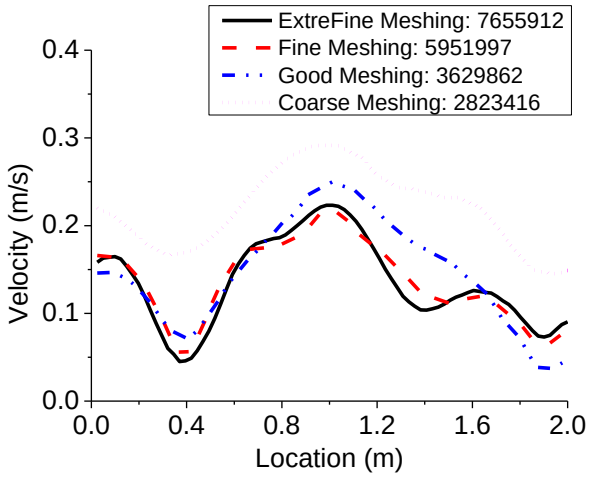

Figure 5.15 Mesh independent testing results.

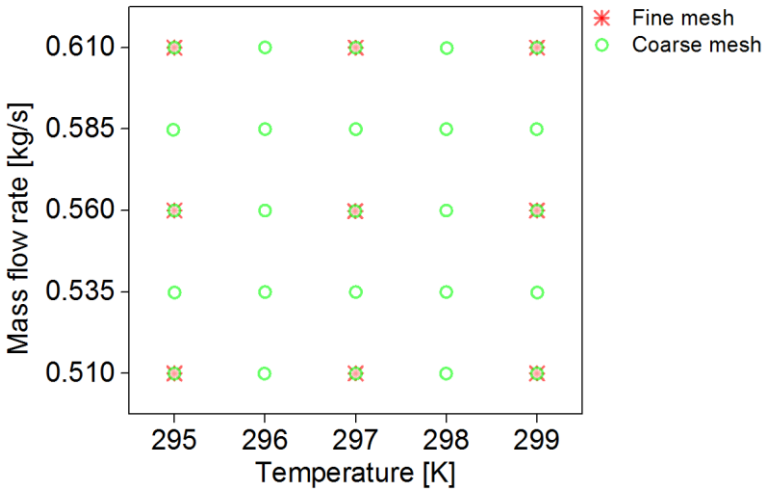

Figure 5.16 Groups of multi-fidelity CFD cases.

In this study, we adopted the fine-mesh set (6.0 million grids) and coarse-mesh set (2.8 million grids) to generate the multi-fidelity training cases. The distribution of the multi-fidelity cases is shown in Figure 5.16, where 9 fine mesh cases plus 25 coarse mesh cases were generated. It is noted that at the 9 locations of the fine mesh cases, the corresponding coarse mesh cases were also performed (i.e. the overlap points in Figure 5.16). With the CFD simulation results from all the multi-fidelity cases (34 in total), we can conduct a multi-fidelity surrogate using Multi-fidelity Kriging algorithm. In order to verify the accuracy of the multi-fidelity surrogate results, Figure 5.17 shows the comparisons of the multi-fidelity predictions for the three design objectives in comparison to the predictions of the high-fidelity Kriging. As depicted, the predictions of the Multi-fidelity Kriging compare extremely well with the predictions of the high fidelity Kriging. The maximum difference between two Kriging is less than 5% which clearly demonstrates the accuracy and robustness of the Multi-fidelity Kriging.

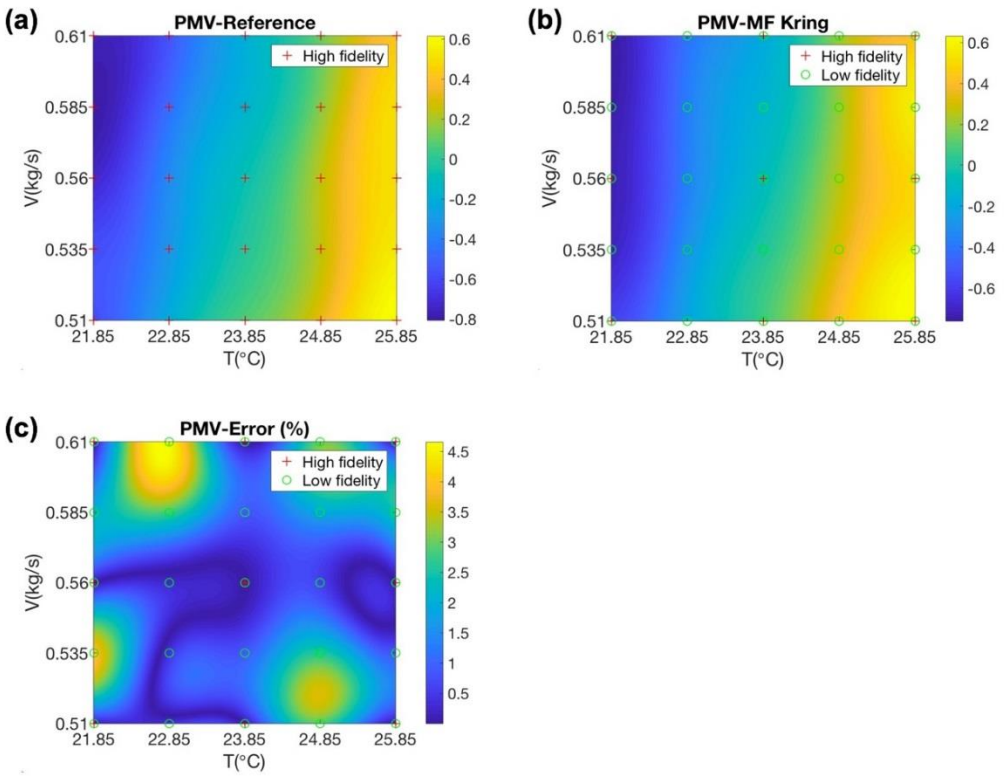

Figure 5.17 Comparison of PMV prediction between Multi-fidelity Kriging and fully high-fidelity Kriging.

The Multi-fidelity Kriging model is then coupled with the NSPSO for design optimization searching the optimal solution within the design space. Figure 5.18 shows the comparison of the Pareto Front obtained by using the high-fidelity (see all the blue dots) and Multi-fidelity Kriging (see all the red dots). The comparison clearly indicates that the Multi-fidelity Kriging is capable to yield reliable and accurate prediction for the NSPSO optimization procedure. The resultant Pareto front shares the overall characteristics and captures the quantitative values as the high-fidelity Kriging method. The maximum prediction error is maintained at 8.16% which is considerably low for ventilation design problem.

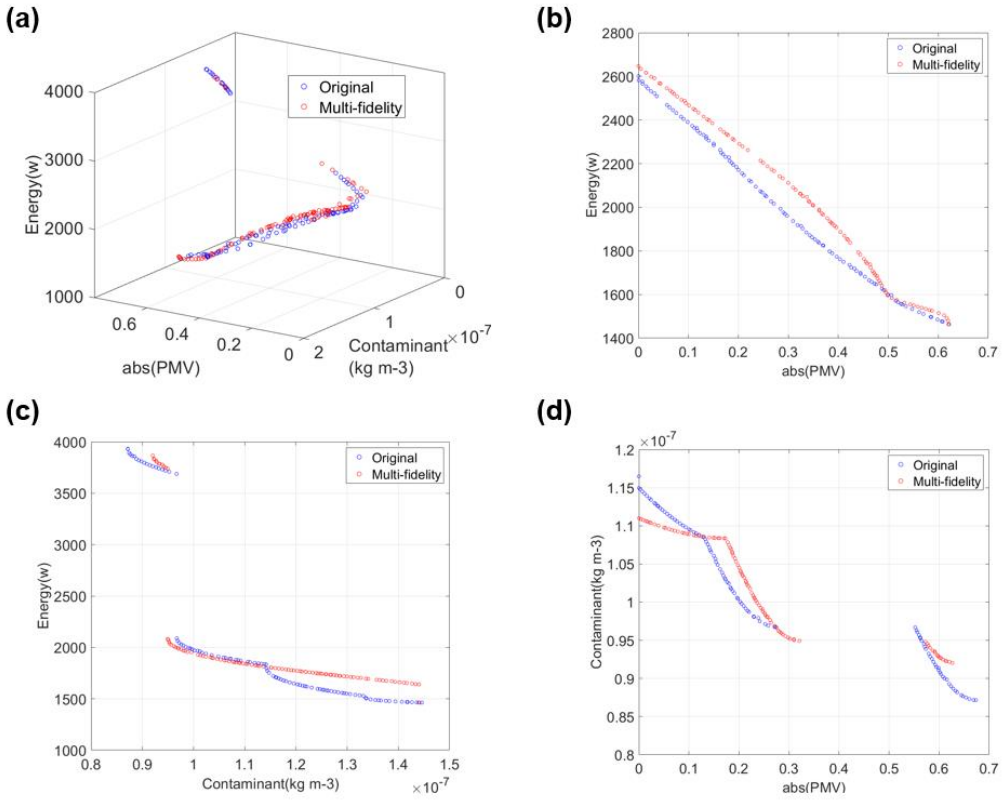

Figure 5.18 Comparison of optimal solutions ('Pareto Front') in objective space between using original high-fidelity Kriging and Multi-fidelity Kriging.

On the other hand, the Multi-fidelity Kriging offers generous saving of the total computational time (i.e. the total of CFD simulation and the optimization computational time) comparing to the high-fidelity Kriging approach. A summary of the total computational time (including optimization time) of the two approaches are tabulated in Table 5.2. With Multi-fidelity Kriging, 16 training samples extracted from the coarse mesh were selected in replacing the fine mesh training sample. At the same time, 9 training samples from the coarse mesh overlapped with the fine mesh samples were also included to establish the correlation between fine and coarse mesh CFD prediction. The table shows that using Multi-fidelity Kriging could reduce up to 35.61% of the total computational time which is over 1/3 of the total time required for the optimization procedures.

Table 5.2 Comparisons of CPU time consumptions (100 particles in the optimization process).

Approach	Procedure	CPU Time (s)	Iterations	Total CPU Time (s)
Traditional	CFD simulation	1.332E+05	25	3.330E+06
	Single objective PSO	2.258E+00	>100	>2.258E+02
NSPSO + Kriging	CFD (Fine mesh)	1.332E+05	25	3.330E+06
	NSPSO	1.653E+02	1	1.653E+02
NSPSO + Multi-fidelity Kriging	CFD (Fine mesh)	1.332E+05	9	1.199E+06
	CFD (Coarse mesh)	3.780E+04	25	9.450E+05
	NSPSO	2.371E+02	1	2.371E+02
Total Saved CPU Time	1.186E+06(s) (**35.61%**)			

5.5 Conclusions

A multi-objective design approach has been proposed in this paper by incorporating the nondominated sorting based particle swarm optimization (NSPSO) algorithm with the Kriging method to searching the optimal design condition of the ventilation system in a fully occupied high-speed train (HST) cabin. The advantage of NSPSO is the capability to provide multiple trade-off solutions in one simulation run and to give a clear visualization of solutions in both design space and objective space. Another benefit of adopting NSPSO is that users can easily select alternative solutions according to their experience and preference, rather than struggling to choose appropriate weightings at the beginning of traditional design process.

In order to validate the proposed approach, a high resolution computational model of Chinese CRH2 train cabin was created with ANSYS Fluent package, where the fine 3D-scanned thermal manikin models were used to improve the computational accuracy. To address the time-consuming problem in CFD simulations, especially when we introduced the high resolution computational thermal manikin model, a Multi-fidelity Kriging surrogate was also developed in this study. The

simulation results demonstrate that the proposed new design procedure is capable to save up to 35.61% of the total computational time compared to the traditional single objective approach, while maintaining an acceptable predictive error. In this paper, we used HST cabin as a test bed and validate the proposed methodology. In the future, other application such as office compartments and airplane cabin could also be tested.

Chapter 6

A new method to cut powerline wildfire risk caused by vegetation conduction ignition using CNN

The main findings of this chapter have been published in:

- **Li, N.,** Shang, Y., Tu, J. & Cheung, S. C. (2019). A new method to cut powerline wildfire risk caused by vegetation conduction ignition in Victoria, Australia, using Convolutional Neural Networks. *International Journal of Wildland Fire*. (under review). ***IF = 2.445, Q1***

The occurrence of wildland fires can pose significant consequences on human lives, assets and ecologic environment. A number of wildfires that occurred in Victoria, Australia were caused by powerline faults. There are strong evidences showing that the wildfires caused by electrical fault burn larger areas and are associated with more severe consequences than fires from other causes. One of the most common ignition processes during powerline faults is ignition by electricity conduction through the vegetation when tree branches contact live power wires. In order to investigate this ignition process, lots of experiments and tests in relation to powerline vegetation fault were conducted in the Powerline Bushfire Safety Program supported by the Australia's Victorian Government and all the experimental data are available to public. When electrical fault happens, if we could detect the fault in an early stage and cut down the power, the fire occurrence can be avoided. Aiming to detect the powerline vegetation fault as soon as possible after fault occurs, this paper developed a new model using Convolutional Neural Networks (CNN), which is able to detect the fault species immediately when fault occurs and also send alarms if the ignition process is already in a late stage,

according to the spectral data of electrical signals caused by vegetation faults. The proposed model was tested and validated by the experimental data from the Powerline Bushfire Safety Program.

6.1 Introduction

Wildland fires can pose significant consequences on life, assets and ecologic environment. In the past decade, wildfires have caused huge losses of human lives and properties around the world including in Greece (Amiridis et al. 2012), the United States (Keeley et al. 2009), Russia (Vivchar 2011) and Australia (Chafer et al. 2004, Stephenson et al. 2013, Collins et al. 2015). Especially, in Australia, over the past century, 20% of the total building losses in nature hazards arise from bushfires (Mcaneney et al. 2009). Understanding the ignition process of the wildfire is essential to reducing the risk of future wildland fires.

A number of wildfires occurred in Victoria, Australia, were attributed to electricity distribution infrastructure, including 12th February 1977 (Mcarthur et al. 1982) and the Ash Wednesday 1983 (Authority 1983, Valent 1984). Although electrical fault is not the top cause of wildfire, these fires resulted from powerline faults have been proved to burn larger areas and thus bring more severe consequences than fires caused by other ignitions (Miller et al. 2017). The most recent example was the Black Saturday bushfires in Victoria in February 2009, which burnt over 270,000 ha, killed 173 people, destroyed 2,029 homes and caused $4.4 billion of economic (Teague et al. 2010, Whittaker et al. 2013, Marxsen 2016). There was clear evidence showing that six of the major fires on Black Saturday were caused by powerline faults (Miller et al. 2017). Following the catastrophic wildland fire of Black Saturday in 2009, the Victorian Bushfires Royal Commission was established to investigate how to better understand bushfires and prevent its occurrence in the future.

Electrical faults can ignite wildfires through three major ignition mechanisms: high voltage arcs near vegetation; molten metal particles emitted when vegetation contacts power wires; and high voltage current through vegetation (Coldham 2010, Russell et al. 2012). When these hot

sources drop to the ground, they can ignite nearby fuels such as dry grass and tree leaf, causing small fires (Urban et al. 2015). If this occurs in days with extreme weather conditions, the small fires can spread very quickly and would result in devastating wildfires (Fernandez-Pello et al. 2015).

The most complex ignition process during powerline fault is ignition from the electricity conduction through the vegetation. The latest research on this type of vegetation fault in the Powerline Bushfire Safety program (PBSP) established by the Victorian Government was conducted in 2015. In the project, over 1,000 tests were performed on 24 species to understand the ignition process and provide database of the fault signatures of different species (Marxsen 2016). Although the project delivered the ignition process from conduction of high voltage electricity through vegetation and provided a statistic valid ranking of bushfire risks for 12 selected species (Table 6.1), a more scientific approach of species fire risk and live fault detection are yet to be studied.

Table 6.1 lists the statistic fire probability of different plant species for experimental test conducted in PBSP where CI represents confidence interval. It is clearly indicated that different species have different fire probability after touching powerline, where the worst fire risk is Willow (Salix Species) and the best is Peppercorn (Schinus Molle), while other species fall between these extremes (Marxsen 2016). In addition, the previous research works in PBSP also delivered detailed analysis of ignition process caused by vegetation fault and provided more than 300GB of data for all the experimental tests which include photos of samples, video recordings, IR camera videos, fault signal recordings, etc. An efficient approach to analyse these data is not yet proposed. On the other hand, traditional protection system relies on a current limit device which uses a pre-set fixed limit value. However, from the ignition process research, the current limit should vary from plant species to avoid over-sensitivity. Therefore, a dynamic detection and protection system is required to provide better powerline protection, which is able to cut off powerline at the right time according to live fault species detection.

Table 6.1 Species fire risk ranking - fire probability for tests with a one amp current limit.

Rank	Species	Number of tests	Average fire probability	Min fire probability (95% CI)	Max fire probability (95% CI)
1	Salix Sp.	88	1.00	0.72	1.00
2	F. Angustifolia	105	0.58	0.32	0.81
3	A. Mearnsii	92	0.57	0.33	0.79
4	Pinus Radiata	56	0.55	0.28	0.79
5	Eu. Baxteri	97	0.53	0.32	0.73
6	Eu. Viminalis	95	0.50	0.25	0.75
7	A. Melanoxylon	52	0.23	0.07	0.48
8	C. Glaucophyllus	43	0.21	0.06	0.46
9	A. Pycnantha	29	0.10	0.01	0.37
10	P. Undulatum	103	0.07	0.00	0.27
11	A. Verticillata	69	0.05	0.00	0.21
12	S. Molle	28	0.00	0.00	0.21

In this paper, we proposed a new live model to detect the vegetation fault signatures using Convolutional Neural Networks (CNN). A scientific approach to classify the species according to the fire risks was also presented in the paper. In order to verify the proposed model, three typical vegetation species (two extreme cases in Table 6.1, i.e. Salix Sp. and S. Molle, and one with the most test data - F. Angustifolia) were selected from the experiment conducted in the PBSP. The detection accuracy and the fire risk ranking for the selected species are discussed in the paper.

6.2 Methodology

6.2.1 History of the vegetation conduction ignition test

Following the disastrous wildfires (i.e. Black Saturday) in February 2009, the Victorian Government of Australia announced a $750 million Powerline Bushfire Safety program (PBSP) in 2011; aiming to reduce the

risk of bushfires started from electrical faults without causing significant impact on electricity supply reliability. Within the program, the world first research project focused on vegetation conduction ignition was completed in 2015 at Springvale substation, Victoria (Figure 6.1). The project studied the ignition processes and recorded corresponding fault signatures (i.e. electrical signals caused by vegetation faults) for three different fault types: branch across wires (phase-phase), branch touching wire (phase – earth), and wire into vegetation. A total of 1,038 experiment tests were conducted on 24 plant species in the project. A detailed report about the project can be found in (Marxsen 2016) and the full authorised data set including photos, videos, test logs and report, are available on the DataVic website. Here we give a brief description of the vegetation conduction ignition process related to this paper.

Figure 6.1 Vegetation conduction ignition test facility at Springvale.

According to the previous research reported in (Marxsen 2016), the vegetation conduction ignition process follows a consistent sequence of stages illustrated in Figure 6.2, which shows the direct fault current recording with time in a phase to earth test of a Desert Ash sample. The process can be divided into four stages: stage one is development of full conductor-vegetation contact where showing an increasing of fault current; stage two is expulsion of moisture where showing a decreasing of fault current; stage three is progressive charring of bark extending

from the thinner end of the vegetation sample where significant fluctuation of fault current can be viewed; and final stage is flashover when flame bridges the high voltage conductors when cutting off the power supply as the fault current has reached the current limit 2A. Figure 6.3 gives an example of the four-stage development of fire after the vegetation fault occurs. In stage one, each conductor-branch contact point was progressively wrapped in a ball of plasma or flame which equivalently enlarges the conductor-branch contact area leading to the rising of fault current. Following the first peak of current, in stage two, the current starts decreasing. This is because in this stage, the moisture contained in the branch turns into steam and leaves the branch. Sometimes jets of steam or even streams of water could be seen issuing from the branch. In stage 3, after moisture expulsion starts to fade, flame slowly spread along the branch from one or both ends and intermittent electric arcs appearing the flame would short-circuit the burning section of branch and cause large fluctuations in the current. Finally, the whole branch was on fire (flashover) in stage four and the power supply was cut off by the powerline protection system.

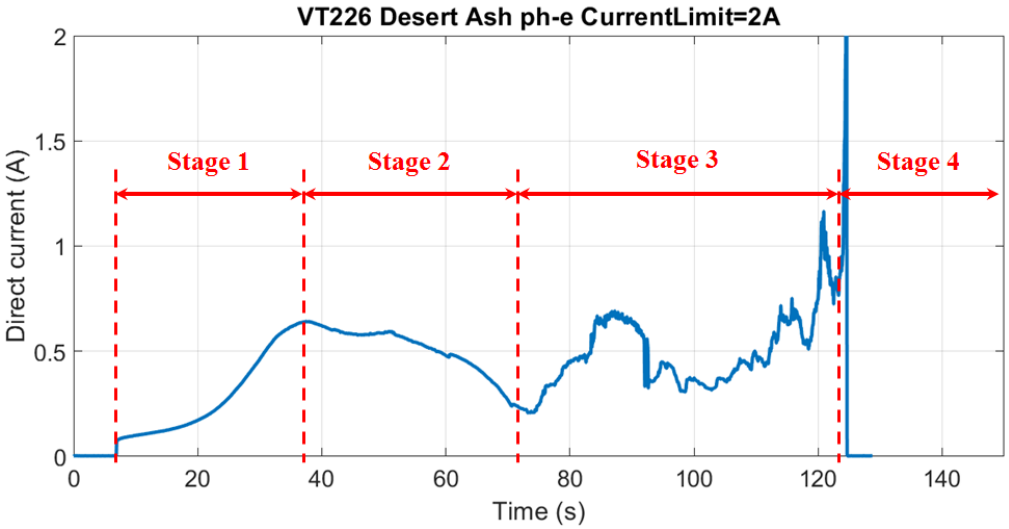

Figure 6.2 Typical variation of fault current in the four stages of vegetation fault development.

Figure 6.3 An example of fire development during the vegetation conduction ignition process (Marxsen 2016).

The experimental tests were conducted on a total of 24 plant species, and they found different species have similar fault current patterns, however the values varies a lot among species. One of the critical values is the first peak current as once entering into stage two, the ignition process does not normally terminate in stage two as the current starts decreasing (i.e. the powerline protection system does not cut off the power supply before the first peak if the pre-set current limit is greater than the peak value), it usually continues into stage three and four, which would have a great probability of causing bushfire. Since different vegetation species have different first peak current values, it is hard to choose a proper pre-set current limit (neither too sensitive nor too slow) for the powerline protection system. In this paper, we developed a new method to cut powerline wildfire risk using CNN, which is able to detect the fault species immediately when fault occurs and also send alarms if the ignition process is already in a late stage (passing stage one).

6.2.2 **Convolutional Neural Networks**

First introduced by (Fukushima 1980), the Convolutional Neural Networks (CNN) is a class of deep artificial neural networks, which has been successfully applied to analysing 2-D visual imagery, such as hand-written character recognition (Lecun 1989, Lecun and Bengio 1995, Niu and Suen 2012). A typical structure of CNN is shown in Figure 6.4, which consists of an input and an output layer, as well as multiple hidden layers in between. The hidden layers of a CNN typically consist of convolutional layers, subsampling (pooling) layers and fully connected layers (see Figure 6.4). The convolutional layers apply convolution filters to the original input to get feature maps which will be passed to the next layer. Following convolutional layer, there may be a pooling layer which combines the outputs of neuron clusters at one layer into a single neuron. There are several methods of pooling, including max pooling, average pooling and linear combination pooling. By applying convolutional and pooling, the CNN significantly reduces the number of neurons, allowing the network to be deeper with fewer parameters (Aghdam and Heravi 2017), and making the network practical for training with limited computational resources. It is worth noting that there can be multiple convolutional layers and pooling layers in a CNN and more layers makes the network deeper. Finally, the fully connected layer receives the information from last hidden layer and turns it into high-level reasoning output.

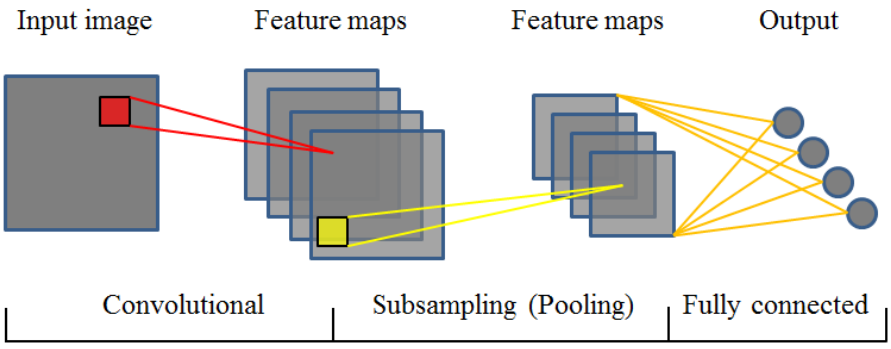

Figure 6.4 A typical CNN structure.

Since the first successful application of LeNet on hand-written digits recognition (Lecun et al. 1990), the CNN has shown great potential on solving different problems including image classification (Ciresan et al. 2011), face recognition (Parkhi et al. 2015), natural language processing (Kumar et al. 2016) and object tracking (Hong et al. 2015). Since the CNN has shown great potential in processing images and making predictions based on images, in this paper, we applied CNN to detect the powerline fault species, according to the spectral data of electrical signals caused by vegetation faults.

6.3 Results and discussions

6.3.1 Vegetation fault detection and evaluation model

In order to perform vegetation fault detection and fire risk evaluation, we proposed a two-step detection model. In step one, we build mathematical models of the fault current (see Figure 6.2 for an example) and calculate the critical point (near the peak in stage one of ignition process). Then we can give rankings of fire risk for different species according to the time and current value at the critical point. In step two of detection, we create and train a CNN model to recognize which species cause the fault according to the fault current spectrogram (Figure 6.5). Afterwards, combining the fire risk rankings given in step one, we provide a comprehensive solution for powerline vegetation fault detection. Details of the two-step model are described as followings.

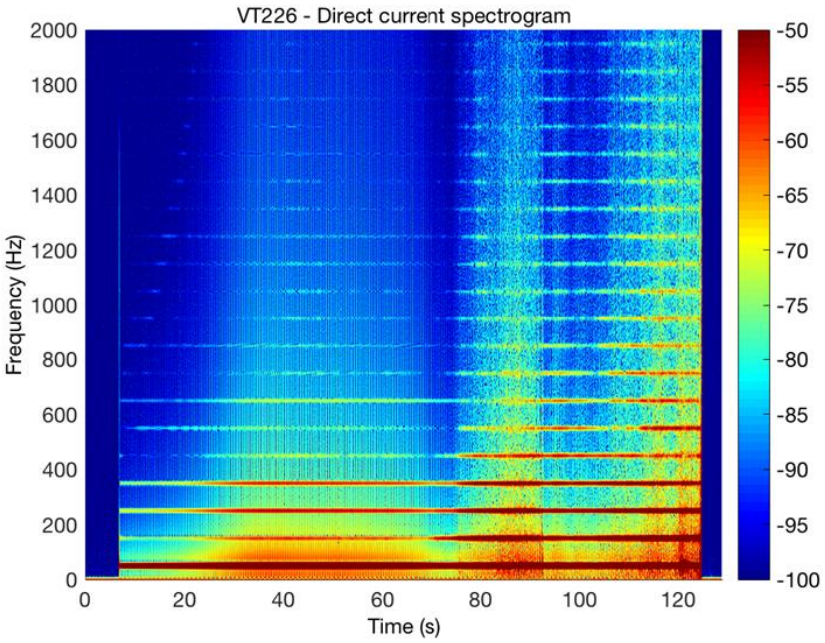

Figure 6.5 An example of direct fault current spectrogram.

a) *Step one of detection*

As discussed in Section 6.2.1, it would be hard to get fault signal detected if the ignition process has passed stage one (i.e. fault current has peaked already). The protection system must detect the fault signal before stage two, otherwise it would pose a great risk of causing bushfire. Therefore, the peak point here is a critical detection point and we should pay more attention to the fault signal in stage one.

Figure 6.6 shows the fault current (blue curve) in stage one for test VT226 (i.e. the segment before the first peak in Figure 6.2). The red curve in Figure 6.6 is the derivative of the current where there is a peak (red circled in Figure 6.6) located just before the current peak. In this paper, we call this circled peak 'Boiling Point' (BP) as the moisture contained in the branch turns into steam and leaves the branch at this time, which can be observed from the video recording. Before reaching BP, the fault current increases sharply while after BP, the current rises up slowly till reaching the first peak. In our model, we take BP as the critical point instead of the current peak. There are two reasons why we do so. Firstly, in many experimental tests, the power was cut off when the fault current

was still going up due to the protection system has been triggered before reaching current peak. In those cases, we are not able to get the peak point. On the other hand, BP is always earlier than first current peak, which means it is even safer if we could detect the fault signal at BP.

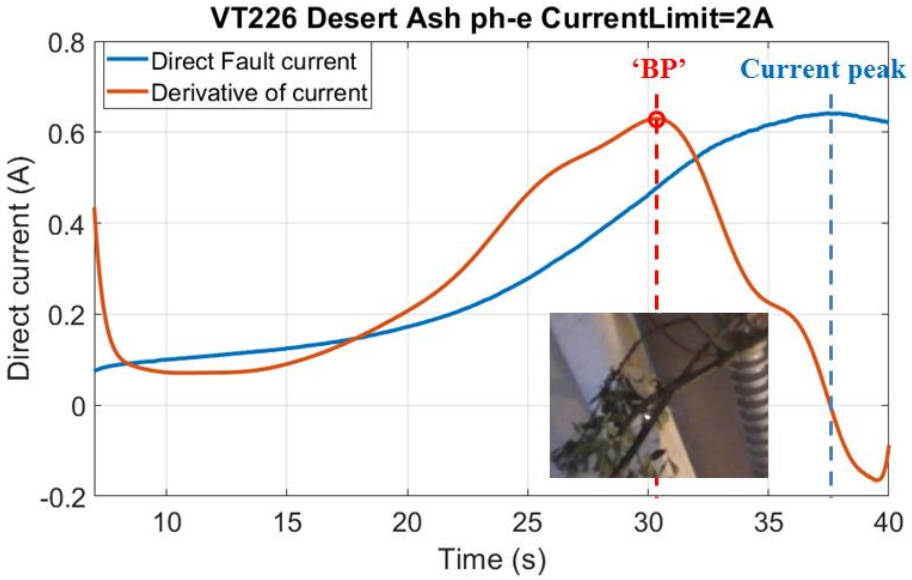

Figure 6.6 Fault current in ignition stage one.

In order to find BP, we need to fit a mathematic function of the fault current curve. We used a piecewise function to do the curve fitting:

$$I = \frac{I_{t=0}}{1 + a(e^{bt} - 1) + Piecewise\{[0, t \leq BP], [c(e^{d(t-BP)} - 1), t > BP]\}}$$

where $I_{t=0}$ is the initial current, a, b, c, d are the fitting hyperparameters. We did curve fitting for all the test results of different species and recorded the time when moisture boiling occurs and the value of fault current at that time. The information at BP could provide a reference when we give each species a ranking number in terms of causing bushfire according to the following two rules: the earlier BP occurs, the bigger chance of causing fire as shorter time left for reaction; the smaller current at BP, the bigger chance of causing fire because it is harder to be detected by protection system.

b) *Step two of detection*

In this step, rather than using fault current curve which was used in step one, we use spectrogram of fault signal (see Figure 6.5 for an example) as input to create a vegetation fault live detection model. Each spectrogram can be treated as a 2D image and CNN is one of the most promising techniques to deal with image detection. The first step in CNN is to get training samples and label these samples. In this study, we only have around 1000 test results available in total, which is not enough for training CNN; on the other hand, the whole image of spectrogram is too big for input of CNN. In order to get sufficient samples and to reduce the sample size, we split the full spectrogram into small sub-images using a sliding sample window. Figure 6.7 gives an example of this data augmentation process. The width of sample window is 12 pixels (i.e. ~1s) and the moving step is one pixel per sub-sample. By doing this, we can get around 64 sub-samples for test VT609. The sub-samples before 'BP' are labeled as 'species name + fault type' (e.g. 'Desert Ash ph-e' in Figure 6.7) and those after 'BP' are all labeled as 'danger' regardless of the species name. We did the similar procedure for all the available test results and next step is using the labeled samples to train a CNN model.

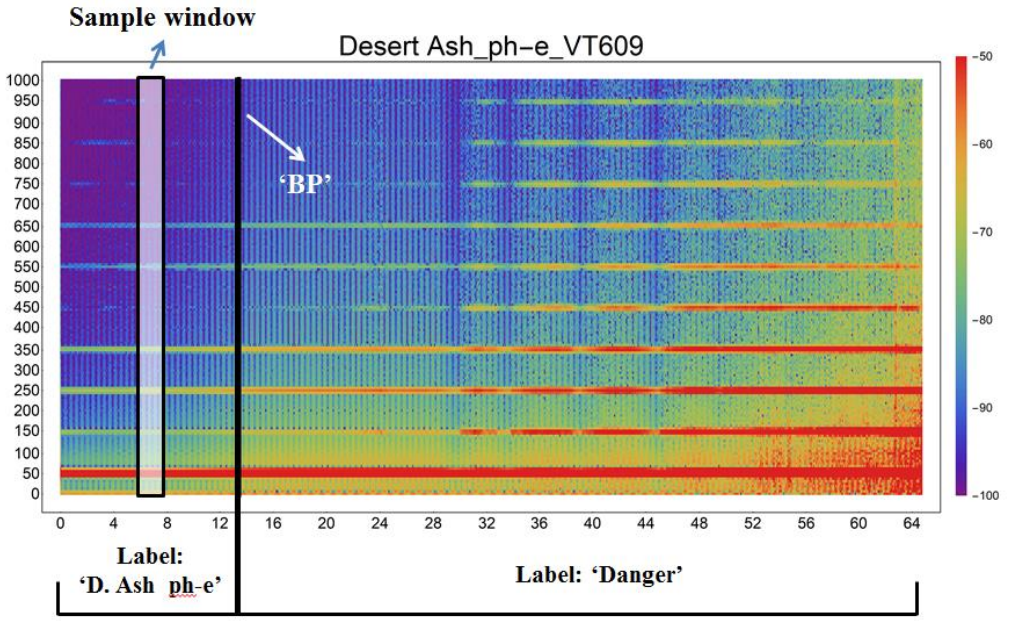

Figure 6.7 Sample splitting and labeling.

Refer to the typical structure of CNN shown in Figure 6.4, in this paper, we created a CNN model indicated in Figure 6.8. As shown in Figure 6.8, the model contains 13 layers in total including one input layer, several convolution ramp pooling layers, fully connection layers and final output layer. Figure 6.9 gives detailed information of each layer including layer name, data type and data size. Through these layers, raw images get filtered, rectified and pooled to create a set of shrunken, feature-filtered images. These can be filtered and shrunken again and again. Each time, the features become larger and more complex and the images become more compact. The lower layers represent simple aspects of the image, such as edges and bright spots, while the higher layers can represent increasingly sophisticated aspects of the image, such as shapes and patterns. With respect of the configuration of the proposed CNN structure, we finalized using a total of 13 layers by experiments. In the experiments, we found that either less layers or more layers could lower the performance of the network, due to underfitting (too few layers) or overfitting (too many layers). As we all know, the tuning process of hyper-parameters of the CNN are still based on experiment and experience, thus 13 layers are a result of experiment in this study. Figure 6.10 gives an example of what the features look like after each layer using a subsample from Figure 6.7 (t=5s). Figure 6.10(a) is the raw sample image and Figure 6.10(b) shows the feature maps after first convolutional calculation. Figure 6.10(c) gives the shrunken feature maps after ramp and pooling. Figure 6.10(d) indicates the deeper feature maps after second loop of convolution, ramp and pooling. Figure 6.10(e, f and g) demonstrates the outputs from fully connection layers, where Figure 6.10(g) is the final layer that outputs the corresponding category with highest possibility value (i.e. 'Desert Ash ph-e' in this example).

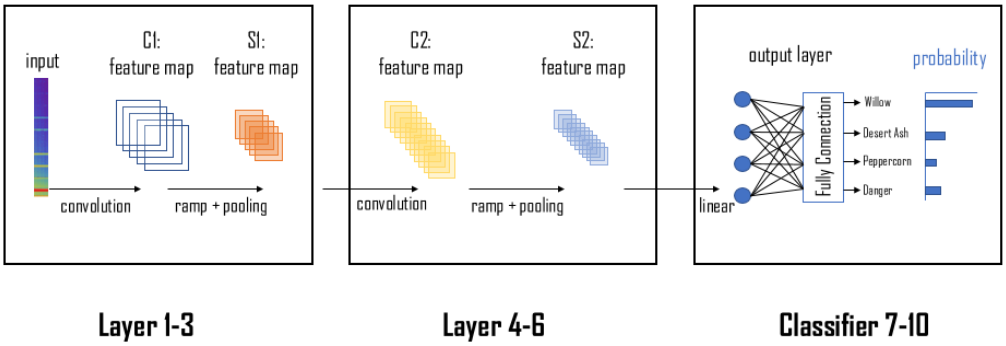

Figure 6.8 Proposed CNN structure.

	Input	3-tensor (size: 1 × 165 × 12)
1	ConvolutionLayer	3-tensor (size: 4 × 156 × 8)
2	Ramp	3-tensor (size: 4 × 156 × 8)
3	PoolingLayer	3-tensor (size: 4 × 78 × 4)
4	ConvolutionLayer	3-tensor (size: 13 × 74 × 2)
5	Ramp	3-tensor (size: 13 × 74 × 2)
6	PoolingLayer	3-tensor (size: 13 × 37 × 1)
7	FlattenLayer	vector (size: 481)
8	LinearLayer	vector (size: 200)
9	Ramp	vector (size: 200)
10	LinearLayer	vector (size: 13)
11	SoftmaxLayer	vector (size: 13)
	Output	class

Figure 6.9 Details of the proposed CNN layers.

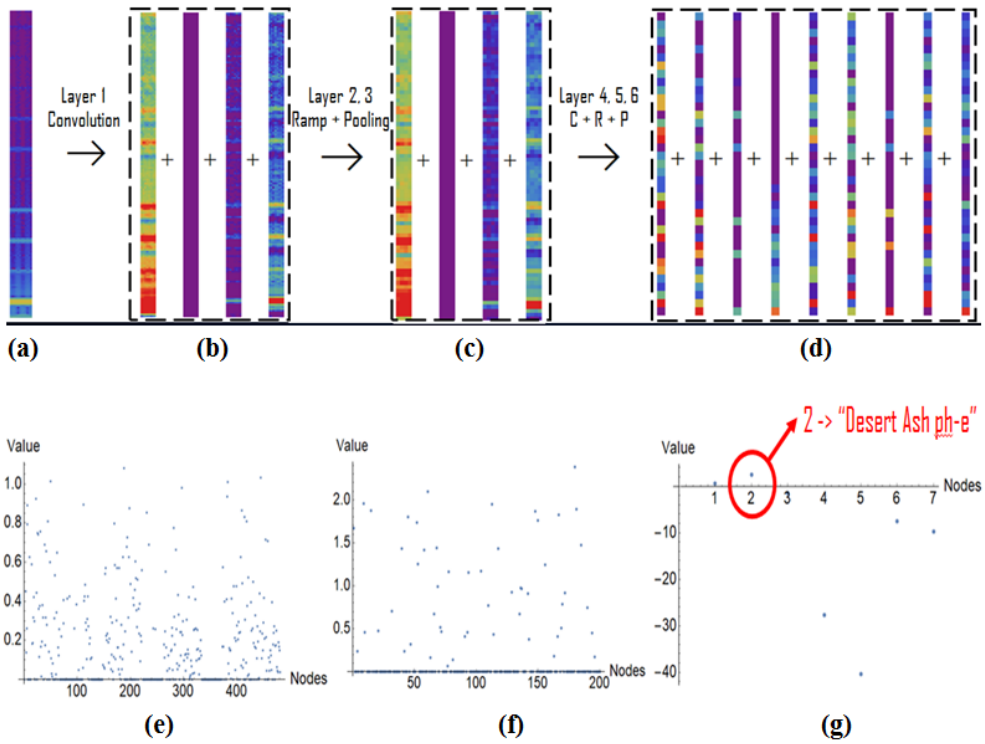

Figure 6.10 Visualization of CNN layers.

6.3.2 **Validation Results**

In order to validate the proposed vegetation fault detection approach, we selected three typical species to test our models: Salix Species (Willow), Franxinus Angustifolia (Desert Ash) and Schinus Molle (Peppercorn).

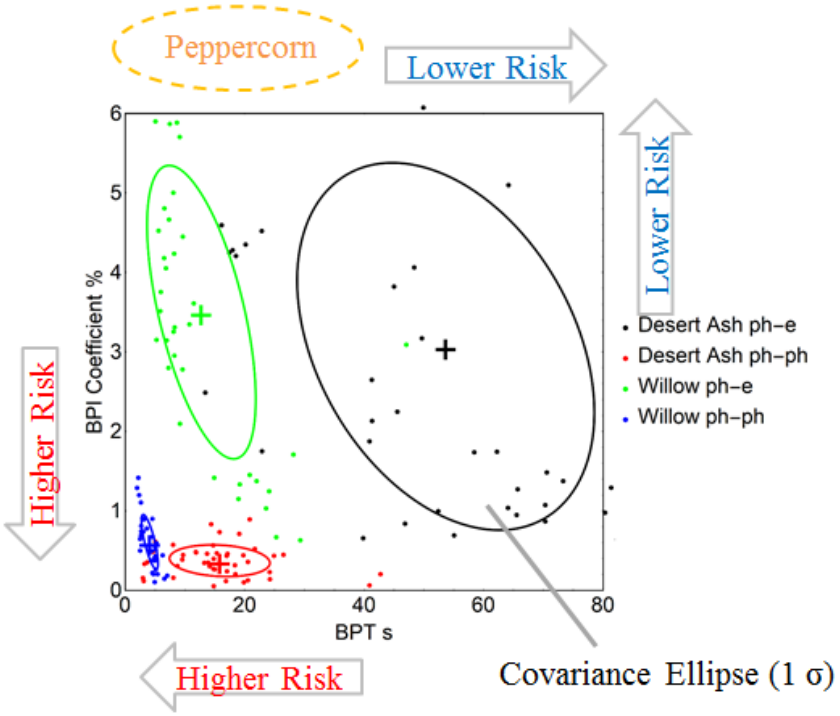

Figure 6.11 Fire risk rankings of three typical species: Willow, Desert Ash and Peppercorn.

Figure 6.11 shows the step one results of the chosen species, i.e. fire risk rankings. As we described in Section 6.3.1(a), we use the fault current and time at 'BP' to guide the ranking. In Figure 6.11, the BP current (BPI) and BP time (BPT) for different species and different fault types are scattered using different colors. In addition, corresponding covariance ellipses are plotted as well. It is noted that there are two types of fault (i.e. phase – phase and phase – earth), where the typical background current values are different (300A and 20A, respectively). In order to compare fire risks among different species and different fault types, in Figure 6.11, the y axis is BPI coefficient (i.e. BPI divided by background current) instead of the absolute value of fault current. From Figure 6.11, it can be observed that through this approach, we can separate the species and fault types and allocate fire risk ranking levels according to the discussion in Section 6.3.1(a). It is worth noting that for Peppercorn, all experiment tests were terminated before BP as power protection system triggered. It means the fault current of Peppercorn rises sharply and it has a much bigger BPI than other two species (see Figure 6.12 for comparisons). Thus the Peppercorn should locate at the top of the plot in Figure 6.11 (orange

dashed ellipse). According to the discussion in Section 6.3.1(a): the earlier BP occurs and the smaller current at BP, the bigger chance of causing fire, we could give fire risk rankings of the species shown in Figure 6.11: Willow > Desert Ash > Peppercorn and ph – ph > ph – e. A summary of average fire probability of these three species in the experimental tests can be found in Table 6.1 from Marxsen's report (Marxsen 2016), which is consistent with the rankings given by this paper (i.e. Willow: 100%, Desert Ash: 58%, Peppercorn: 0%).

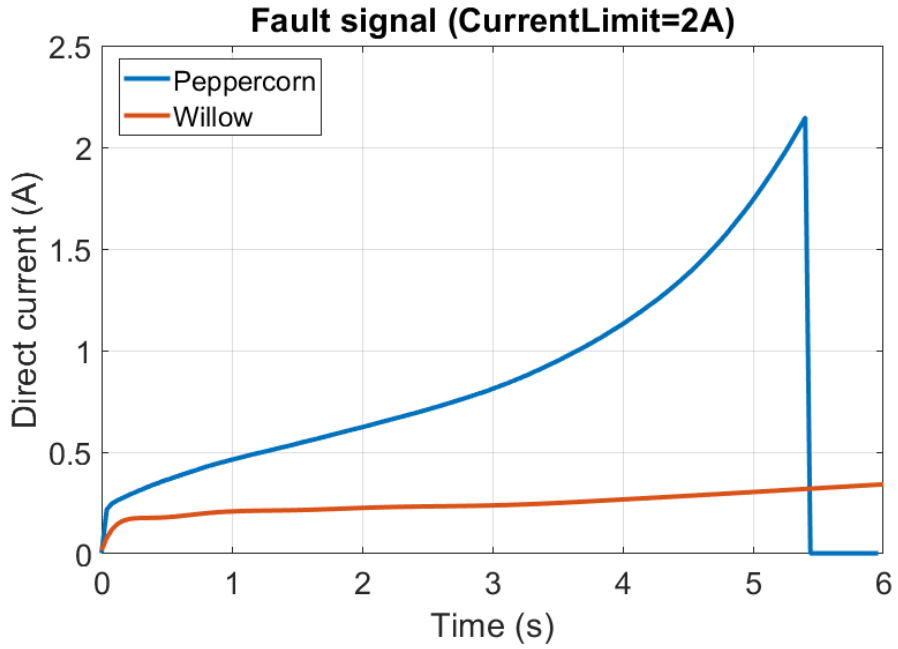

Figure 6.12 Comparison of fault current in ignition stage one between Willow (VT153) and Peppercorn (VT973).

After getting the fire risk rankings, the next step is live vegetation detection using CNN. As discussed before, in this step, we use spectrogram of fault signal as input to predict which species is touching powerline when fault occurs. A CNN model was trained to do the prediction. Details of the model structure and data sampling have been discussed in previous sections (Figure 6.7~Figure 6.9). During the training process, 80% of all available samples were used as training set and the rest 20% were used for test purpose. Table 6.2 lists the prediction accuracy of the trained CNN for the selected three species.

Table 6.2 CNN prediction accuracy for Willow, Desert Ash, Peppercorn.

Label	Data group	Accuracy	Number of samples
'Dangerous'	Training	96.6%	62,507
	Test	93.2%	15,816
'Desert Ash ph-e'	Training	87.5%	25,213
	Test	85.4%	6,085
'Desert Ash ph-ph'	Training	93.4%	7,767
	Test	73.0%	1,482
'Peppercorn ph-e'	Training	97.4%	737
	Test	84.3%	481
'Peppercorn ph-ph'	Training	78.6%	14
	Test	N/A	0
'Willow ph-e'	Training	82.0%	4,427
	Test	79.7%	1,394
'Willow ph-ph'	Training	75.8%	1,505
	Test	75.6%	364

In Table 6.2, it is noted that for three species, 7 labels in total are used, which include one 'Danger' representing fault stages passing 'BP' for all species, the rest 6 labels represent fault stages before 'BP' (safe for now) for different species and different fault types. In this paper, the 'Dangerous' is the priority to be detected because this situation is urgent, while other statuses are relatively safe. Therefore, the goal of the algorithm is to detect 'Danger' with high accuracy. From the prediction results listed in Table 6.2, we can find that good detection accuracy (greater than 93%) has been achieved for 'Danger' in both training set (96.6%) and test set (93.2%). It is worth noting that for the species and fault type detection, the accuracy is calculated as the following rule: incorrect results are defined as the species or fault types are incorrectly predicted, not including detected as 'dangerous' situation. For example, if the 'Willow ph-e' was predicted as 'Desert Ash ph-e' by our model, it is counted as incorrect results, and in this situation, it is very dangerous, because Willow has higher fire risk than Desert Ash. However, if the

'Willow ph-e' was predicted 'Dangerous', it will not be counted as incorrect results, because it is possible that the 'Dangerous' is caused by Willow and our model just gives an alert at an early stage which is not causing dangerous problem as before. Also, from the Table 6.2, one can notice that some training accuracy rates are relatively low (peppercorn and willow), and this is mainly because we are lack of enough samples (<2,000) for training the deep network. The accuracy rates would be improved once we collect more samples from the lab experiments. Figure 6.13 gives such an example of CNN live detection results for 'Willow ph-e', where the bottom color bar indicates the detection results (blue: Willow ph-e fault stage one, black: Danger). From Figure 6.13, we can see that more than half samples from fault stage one were incorrectly classified as 'Danger'. As Willow is the most dangerous species (i.e. highest fire risk), it is acceptable to send a fire alarm when fault is in stage one (i.e. before 'BP'). It could bring mistake alert. However, from the aspect of safety, it is better than late or no alert. Therefore, this scenario does not pose a negative impact to the performance of our model. Figure 6.14 gives more live detection results for the other two species (Desert Ash and Peppercorn), which demonstrates the good performance of our CNN model.

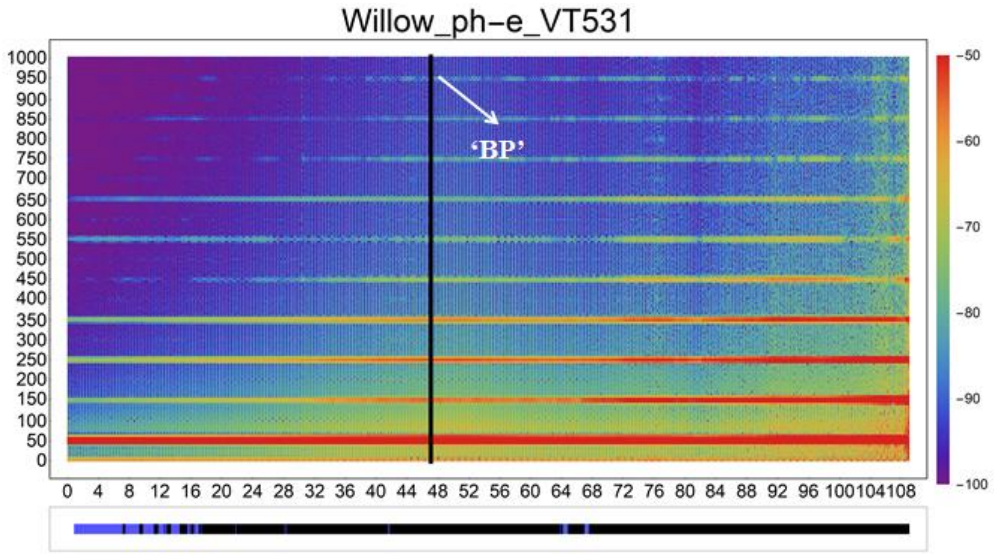

Figure 6.13 An example of live detection results for Willow ph-e.

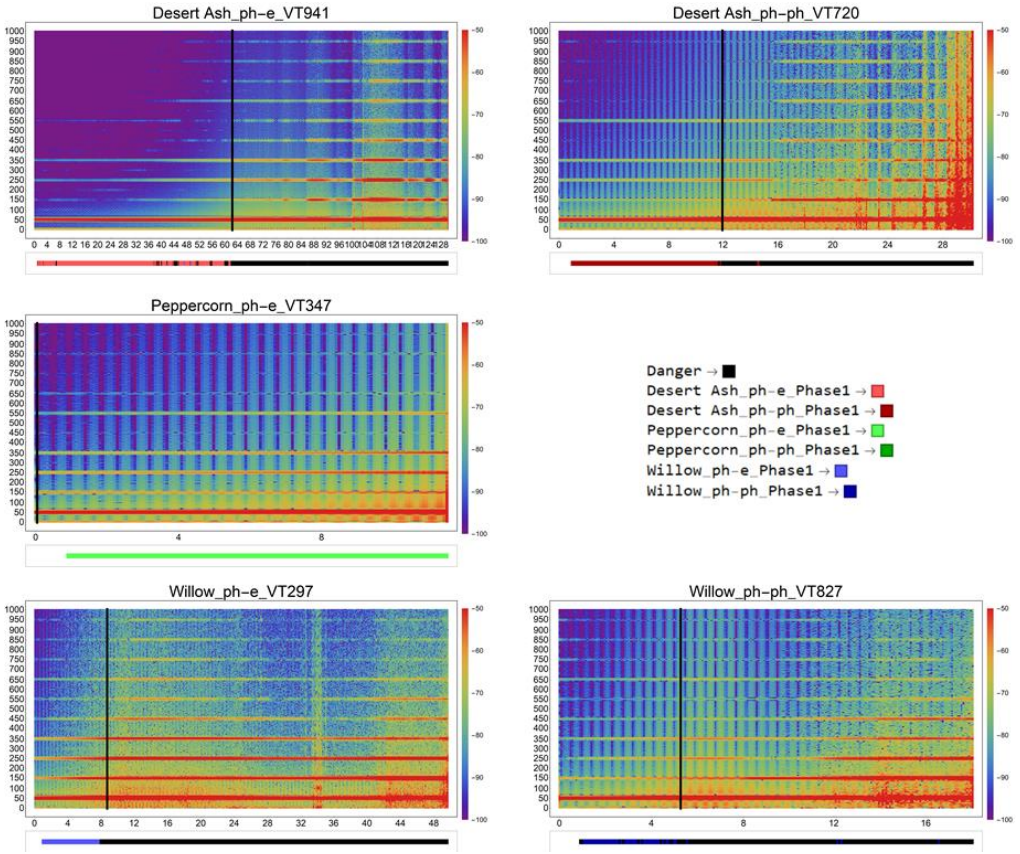

Figure 6.14 Live detection results.

6.3.3 **Noise impact**

The data used in above sections are from laboratory experimental test (i.e. the data are clean). However, the on-site measured data must contain noise and our model should take the noise signal into account. There are two steps in our model where step one is a static detection and step two is a live detection. In step one, we can use clean data obtained from lab test to give fire risk rankings of different species, while in step two, since it is live detection, noise in field test has to be considered. Figure 6.15 gives an example of extreme case where the fault current has large fluctuations after 'BP' (starting from t=30s). From the live detection results shown in the bottom image in Figure 6.15, we can find that those fluctuations affect the detection accuracy and result in detection errors (circled in bottom color bar). However, these narrow detection errors surrounded by correct results (black bars) are easily to be ignored, because we are looking at a continuous fault ignition process and those signals that

suddenly appear and exist no more than 1 second could be treated as wrong signals and therefore should be ignored.

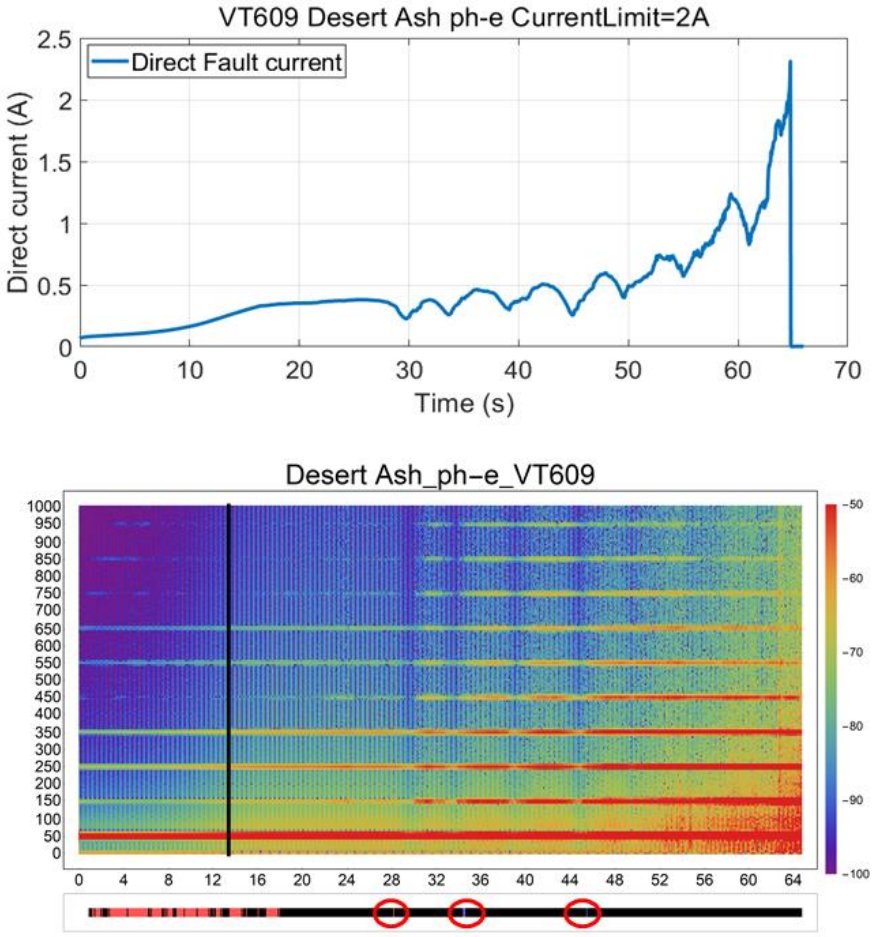

Figure 6.15 An example of extreme case with large current fluctuations after 'BP'.

Furthermore, since the network noise data are available on the DataVic website, we could add the noise signal and the fault signal together to represent the on-site signal which contains noise. Figure 6.16 shows the noise direct current curve and corresponding spectrogram graph. It is noticed from spectrogram that the noise is mainly composed of 50Hz signals which is reasonable as the utility frequency of powerline is 50Hz. In order to study the influence of noise to the CNN live detection model, we add the noise spectrogram with different levels of gain (1X, 2X, 5X) to the aforementioned extreme case (VT609 shown in Figure 6.15), and then test the detection model (trained with clean data) using the data with

noise. The comparisons of detection results are shown in Figure 6.17, where image (a) gives the detection result of the original signal, and the rest images indicate the results given by the same detection model with 1X, 2X, 5X noise signal, respectively. In this study, detecting vegetation species correctly after powerline fault occurs and giving fire alert immediately when the fault process is passed 'BP' are the two important goals to cut powerline wildfire risk caused by vegetation conduction ignition. From the results shown in Figure 6.17(b, c, d), we can find that the noise signal makes the fire alert around 4 seconds late (i.e. 'BP' occurs at 13s, while continuous black bar starts from 17s). Although the noise causes 4 seconds delay, comparing to 64 seconds which is the traditional current limit safety device would take to react to fault, our model is much faster (17 seconds). On the other hand, we could retrain the model as long as we collected enough on-site test data (with noise) to achieve a better robustness performance of dealing with noise.

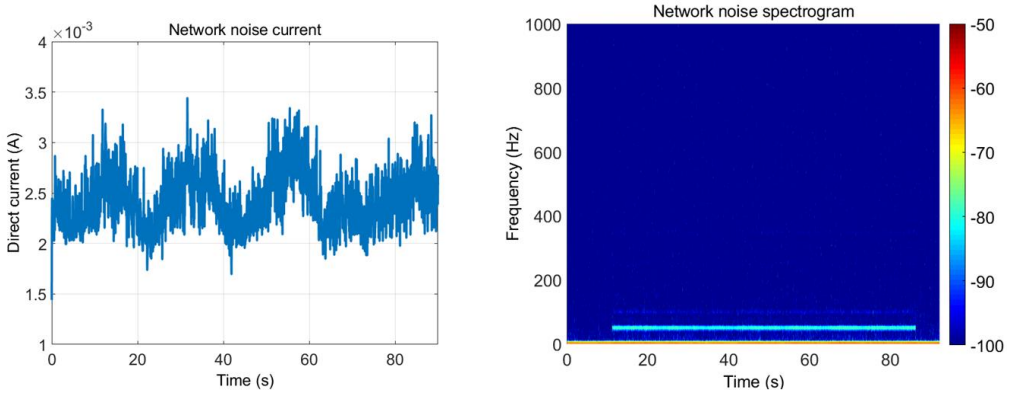

Figure 6.16 Network noise current and spectrogram.

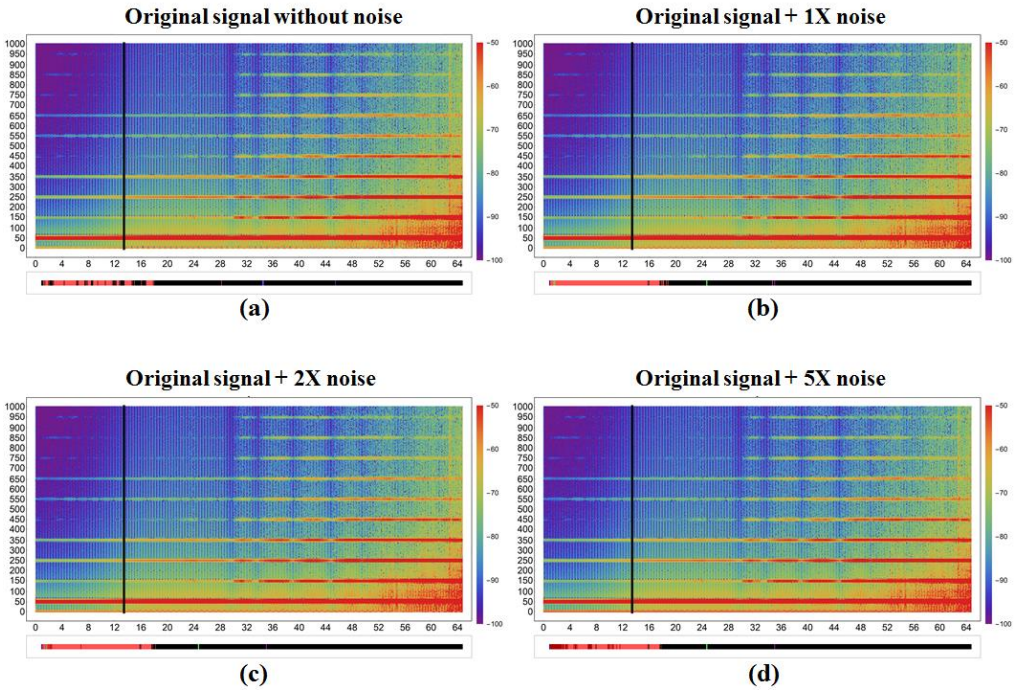

Figure 6.17 Live detection results of Desert Ash (VT609) with impacts of different gains of noise.

6.4 Conclusions

A novel two-step vegetation detection model has been developed and proposed aiming to cut powerline wildfire risk caused by vegetation conduction ignition after powerline fault occurs. In step one, we focus on the fault current signal before the first peak and according to the information at the critical point ('BP'), we give fire risk rankings of different vegetation species. In step two of the model, live detection is implemented using CNN with the fault signal spectrogram as input. The model will monitor the fault signal continuously and will give the fault species and fault type when fault occurs, providing powerline operators with reference of reacting to the fault according to the fire risk ranking of that species given in step one. The live detection model can also send a 'Danger' alert to operators if the fault process has passed safe stage (i.e. over 'BP'). The testing results on three typical species (Willow, Desert Ash, Peppercorn) of fire risk rankings are consistent with the average fire probabilities of these three species in the experimental tests reported in Marxsen's paper (Marxsen 2016). With the limitation of experimental

training data, the live detection CNN model achieves accuracy of 85.3% 93.2% for species detection and danger detection, respectively. Compared to the traditional current limit safety device, our model provides more detailed information during powerline vegetation fault process and makes corresponding actions within shorter time, thus supplies more accurate and faster protection to cut powerline wildfire risk caused by vegetation conduction ignition. Future research include collecting more test data for each species to further increase the model accuracy, bringing more types of species to do analysis, improving the performance of the model against on-site noise.

Chapter 7
Conclusion

This thesis focused on the development of novel multi-objective software platforms to assist engineering design and investigation, especially for simulation-based indoor environment problems, which always involve multiple evaluation criteria. In addition, this thesis aimed to develop new methods to reduce the computational cost associated with the design process. Also, development of new data analysis approach with the assistance of machine learning techniques to handle a large amount of data quickly and accurately was a part of this thesis. The main contributions from this thesis are:

- A new design scheme was developed to achieve multi-objective optimization to help indoor ventilation system design, without having to use any weighting factors, and is capable to explore the whole design space, to provide multiple solutions.
- Developed a response-adaptive sampling approach, which allocates the sampling points in an adaptive way to save computational resources in the CFD-assisted HVAC system design process.
- A multi-fidelity surrogate model of CFD simulation was developed, which uses both high-fidelity results (fine meshing) and low-fidelity results (coarse meshing), thus is able to save computational cost without sacrificing the simulation accuracy.
- Developed a new data analysis approach using CNN to help find features from large amount of data collected from the lab and help reduce powerline wildfire risk caused by vegetation conduction ignition after powerline fault occurs.

7.1 Details of the contributions

7.1.1 Multi-objective optimization of HVAC system in an office room using NSPSO

Chapter 3 presents a CFD-assisted design procedure of the ventilation system in a typical office compartment. A multi-objective optimization platform has been proposed and developed by incorporating the Nondominated Sorting based Particle Swarm Optimization (NSPSO) algorithm with the Kriging method. To remedy the drawback of some previous studies, the NSPSO algorithm removes the necessity of using weighting factors in constructing the objective function and obtains the corresponding trade-off solutions (i.e. Pareto Front) for the given objective space. With the visualization of solutions in objective space, designers could easily pick up the most appropriate design solution according to their own judgments and preferences, rather than being struggled to decide the value of weighting factor in advance.

Special attention is also taken to minimize the computational cost where considerably large training sample based on Computational Fluid Dynamics (CFD) simulations are usually required for the surrogate modelling. The Kriging method where the best linear unbiased value of the unobserved fields is estimated based on the known sampled data is adopted in the present study. One particular advantage of the Kriging method is its capability in achieving high prediction accuracy with a relatively small training sample size. Predictions from the Kriging method are compared and assessed with the CFD predictions. The comparison has shown that the Kriging method provides excellent accuracy in prediction with the maximum error of 5.12%. In addition, with the proposed adaptive sampling procedure, further reduction of the computational cost could be realized. Based on the given case study, the optimization platform achieves a saving of 46.6% of CPU time without sacrificing the accuracy of the optimal solution.

7.1.2 Multi-fidelity surrogate algorithm for fire origin determination

Chapter 4 demonstrates a reverse engineering solution to fire origin determination using CFD. A novel Multi-fidelity Kriging algorithm for fire forensic investigations was proposed in this chapter, which is capable to quantitatively determine the fire origin based on the soot deposition patterns that could be measured in the post-fire fire scene or predicted by means of numerical simulations. Aiming to develop an alternative methodology assisting the fire investigation, the proposed methodology does not involve expert interpretations and professional judgments in process, which could eliminate the inevitable subjective bias and human error in the determination process. A series of fire scenarios (i.e. totally 41 cases) based on a single compartment fire situation was utilized for assessing the performance of the proposed Multi-fidelity Kriging algorithm. Soot deposition patterns on the compartment wall were predicted by the widely adopted fire CFD package - Fire Dynamics Simulator (FDS). The predicted soot deposition patterns on three vertical wall surfaces were extracted and parameterized into 4 parameters describing the boundary line of the pattern. Performance assessment shows the algorithm could capture the correlation between the high-fidelity and low-fidelity simulation results and produce accurate predictions almost comparable to those made by all high-fidelity training samples. Taking 95% as the confidence interval, the prediction errors of using high-fidelity and multi-fidelity training data are of 0.89m and 0.90m, respectively. More importantly, without scarifying the accuracy, the multi-fidelity Kriging algorithm managed to reduce the associated computational time for constructing the training sample by 58%.

7.1.3 Multi-objective optimization design of HST cabin ventilation system

As a subsequent study of the research works described in chapter 3, a more complex ventilation system design was demonstrated in chapter 5, where a new design approach has been proposed by incorporating the

nondominated sorting based particle swarm optimization (NSPSO) algorithm with the Multi-fidelity Kriging method to searching the optimal design condition of the ventilation system in a fully occupied high-speed train (HST) cabin. The advantage of NSPSO is the capability to provide multiple trade-off solutions in one simulation run and to give a clear visualization of solutions in both design space and objective space. Another benefit of adopting NSPSO is that users can easily select alternative solutions according to their experience and preference, rather than struggling to choose appropriate weightings at the beginning of traditional design process.

In order to validate the proposed approach, a high-resolution computational model of Chinese CRH2 train cabin was created with ANSYS Fluent package, where the fine 3D-scanned thermal manikin models were used to improve the computational accuracy. To address the time-consuming problem in CFD simulations, especially when we introduced the high resolution computational thermal manikin model, a Multi-fidelity Kriging surrogate was also developed in this study. The simulation results demonstrate that the proposed new design procedure is capable to save up to 35.61% of the total computational time compared to the traditional single objective approach, while maintaining an acceptable predictive error.

7.1.4 A new method to cut powerline wildfire risk caused by vegetation conduction ignition using CNN

Chapter 6 presents a new data analysis method using CNN to find data features from big data. A novel two-step vegetation detection model has been developed and proposed aiming to cut powerline wildfire risk caused by vegetation conduction ignition after powerline fault occurs. In step one, we focus on the fault current signal before the first peak and according to the information at the critical point (‘BP’), we give fire risk rankings of different vegetation species. In step two of the model, live detection is implemented using CNN with the fault signal spectrogram as input. The model will monitor the fault signal continuously and will give the fault species and fault type when a fault occurs, providing powerline

operators with reference of reacting to the fault according to the fire risk ranking of that species given in step one. The live detection model can also send a 'Danger' alert to operators if the fault process has passed safe stage (i.e. over 'BP'). The testing results on three typical species (Willow, Desert Ash, Peppercorn) of fire risk rankings are consistent with the average fire probabilities of these three species in the experimental tests reported in Marxsen's paper (Marxsen 2016). With the limitation of experimental training data, the live detection CNN model achieves accuracy of 85.3% 93.2% for species detection and danger detection, respectively. Compared to the traditional current limit safety device, our model provides more detailed information during powerline vegetation fault process and makes corresponding actions within a shorter time, thus supplies more accurate and faster protection to cut powerline wildfire risk caused by vegetation conduction ignition.

7.2 Future works

7.2.1 CFD-assisted fire origin determination

In Chapter 4, as a preliminary validation of the proposed approach, the authors built a single compartment fire model, which is a relatively simple geometry model. Prediction of fire origin in complex geometries (e.g. multi-compartments) is subject to on-going research work. In addition, in chapter 4, we used a lightly sooting fuel - propane as the burning material in the CFD model, where the level of soot concentration is 10^{-5} kg/m^2. However, in reality, most of the building fires involve burning heavily sooting materials such as foamed furniture and natural or synthetic carpets. Using propone as the primary fuel in this study representing a conservative assumption in terms of soot deposition on the wall. In practical building fires, it is expected to have more rigorous soot depositions on the walls. In addition, more scenarios with post-flashover fire will be studied in the future as all the study cases described in chapter 4 were pre-flashover.

7.2.2 **CFD-assisted ventilation system design**

In this thesis, we used single room compartment (Chapter 3) and high-speed train cabin (Chapter 5) as test beds and validate the proposed methodology to help multi-objective indoor ventilation system design. In the future, other applications such as multiple office compartments and aeroplane cabin also need to be tested. In addition, in order to further validate the proposed multi-objective optimization and multi-fidelity surrogate methods, different locations of diffusers, different angle of supply air and cabin environment in different seasons need to be considered as well. Also, the application of the multi-objective optimization framework consisting of the NSPSO and Kriging methods to assist outdoor ventilation system, such as wind energy system design, is also worth of investigating. Subject to financial support, doing experimental test to further validate the computational models is also within our future scopes.

7.2.3 **CNN-assisted vegetation detection**

Future research regarding the CNN-assisted vegetation detection includes collecting more test data for each species to further increase the model accuracy, bringing more types of species to do analysis, improving the performance of the model against on-site noise.

Bibliography

Acikgoz O, Çebi A, Dalkilic AS, Koca A, Çetin G, Gemici Z, Wongwises S (2017). A novel ANN-based approach to estimate heat transfer coefficients in radiant wall heating systems. Energy and Buildings, 144: 401-415.

Aghdam HH, Heravi EJ (2017). Guide to Convolutional Neural Networks: A Practical Application to Traffic-Sign Detection and Classification. Springer.

Alpaydin E (2014). Introduction to machine learning. MIT press.

Amiridis V, Zerefos C, Kazadzis S, Gerasopoulos E, Eleftheratos K, Vrekoussis M, Stohl A, Mamouri RE, Kokkalis P, Papayannis A, Eleftheriadis K, Diapouli E, Keramitsoglou I, Kontoes C, Kotroni V, Lagouvardos K, Marinou E, Giannakaki E, Kostopoulou E, Giannakopoulos C, Richter A, Burrows JP, Mihalopoulos N (2012). Impact of the 2009 Attica wild fires on the air quality in urban Athens. Atmospheric Environment, 46: 536-544.

Ashrae (2004). ANSI ASHRAE Standard 55 -2004 Thermal Environmental Conditions for Human Occupancy.

Association NFP (2013). NFPA 921: Guide for Fire & Explosion Investigations. Technical Committee on Fire Investigations.

Atiya A, Ji CJITONN (1997). How initial conditions affect generalization performance in large networks. 8: 448-451.

Austin J, Brimblecombe P, Sturges W (2002). Air pollution science for the 21st century. Oxford: Elsevier Science Ltd.

Authority CF (1983). The major fires originating 16th February, 1983. Vaughan Printing, Melbourne, 39pp.

Batterman S, Peng C-U (2010). TVOC and CO2 Concentrations as Indicators in Indoor Air Quality Studies. American Industrial Hygiene Association Journal, 56: 55-65.

Beyler CL (2009). Analysis of the fire investigation methods and procedures used in the criminal arson cases against Ernest Ray Willis and Cameron Todd Willingham. Hughes Associates.

Bishop CM (2006). Pattern recognition and machine learning. springer.

Bre F, Gimenez JM, Fachinotti VD (2018). Prediction of wind pressure coefficients on building surfaces using artificial neural networks. Energy and Buildings, 158: 1429-1441.

Bryner N, Madrzykowski D, Grosshandler W (2007). Reconstructing the station nightclub fire–computer modeling of the fire growth and spread. In: International interflam conference, 11th proceedings, pp.3-5.

Buratti C, Palladino D, Ricciardi P (2016). Application of a new 13-value thermal comfort scale to moderate environments. Applied Energy, 180: 859-866.

Buratti C, Ricciardi P (2009). Adaptive analysis of thermal comfort in university classrooms: correlation between experimental data and mathematical models. Building and Environment, 44: 674-687.

Buratti C, Ricciardi P, Vergoni M (2013). HVAC systems testing and check: A simplified model to predict thermal comfort conditions in moderate environments. Applied energy, 104: 117-127.

Cao X, Li J, Liu J, Yang WJB, Environment (2016). 2D-PIV measurement of isothermal air jets from a multi-slot diffuser in aircraft cabin environment. 99: 44-58.

Cao X, Liu J, Jiang N, Chen Q (2014). Particle image velocimetry measurement of indoor airflow field: A review of the technologies and applications. Energy and Buildings, 69: 367-380.

Cardinale N, Stefanizzi P, Rospi G, Augenti V (2010). Thermal performance of a mobile home with light envelope. Building Simulation, 3: 331-338.

Carrese R, Sobester A, Winarto H, Li X (2011). Swarm heuristic for identifying preferred solutions in surrogate-based multi-objective engineering design. AIAA Journal, 49: 1437-1449.

Chafer CJ, Noonan M, Macnaught E (2004). The post-fire measurement of fire severity and intensity in the Christmas 2001 Sydney wildfires. International Journal of Wildland Fire, 13: 227-240.

Chao CYH, Wan MP, Morawska L, Johnson GR, Ristovski ZD, Hargreaves M, Mengersen K, Corbett S, Li Y, Xie X, Katoshevski D (2009). Characterization of expiration air jets and droplet size distributions immediately at the mouth opening. Journal of Aerosol Science, 40: 122-133.

Charnay L, Ångström H-E, Andersson L, Palm B, Östling L (2001). CFD Optimization of an EGR Cooler for Heavy-Duty Diesel Engines.

Chen C, Liu W, Li F, Lin C-H, Liu J, Pei J, Chen Q (2013). A hybrid model for investigating transient particle transport in enclosed environments. Building and Environment, 62: 45-54.

Chen Q, Lee K, Mazumdar S, Poussou S, Wang L, Wang M, Zhang Z (2010). Ventilation performance prediction for buildings: Model assessment. Building and Environment, 45: 295-303.

Chen R, Hu B, Liu Y, Xu J, Yang G, Xu D, Chen C (2016). Beyond PM2.5: The role of ultrafine particles on adverse health effects of air pollution. Biochim Biophys Acta, 1860: 2844-55.

Cheung SC, Yeoh G (2009). A fully-coupled simulation of vortical structures in a large-scale buoyant pool fire. International Journal of Thermal Sciences, 48: 2187-2202.

Chi JH (2013). Reconstruction of an inn fire scene using the Fire Dynamics Simulator (FDS) program. Journal of forensic sciences, 58.

Ciresan DC, Meier U, Masci J, Maria Gambardella L, Schmidhuber J (2011). Flexible, high performance convolutional neural networks for image classification. In: IJCAI Proceedings-International Joint Conference on Artificial Intelligence, pp.1237.

Ciro WD, Eddings EG, Sarofim AF (2006). Experimental and numerical investigation of transient soot buildup on a cylindrical container immersed in a jet fuel pool fire. Combustion science and technology, 178: 2199-2218.

Clerc M, Kennedy J (2002). The particle swarm-explosion, stability, and convergence in a multidimensional complex space. IEEE Transactions on Evolutionary Computation, 6: 58-73.

Cohan BD. 2010. Verification and validation of a candidate soot deposition model In fire dynamics simulator version 5.5. 1.

Coldham D (2010). Bushfire ignition from electric faults, a review of technical literature. HRL Australia, Report No: HCL/2010/440.

Collins KM, Price OF, Penman TD (2015). Spatial patterns of wildfire ignitions in south-eastern Australia. International Journal of Wildland Fire, 24: 1098-1108.

Cressie N (1990). The origins of kriging. Mathematical geology, 22: 239-252.

Crosby, L J (1973). Computer Simulation in Genetics. London: John Wiley & Sons.

De Baar J, Roberts S, Dwight R, Mallol B (2015). Uncertainty quantification for a sailing yacht hull, using multi-fidelity kriging. Computers & Fluids, 123: 185-201.

De Bellis F, Catalano LA (2012). CFD optimization of an immersed particle heat exchanger. Applied Energy, 97: 841-848.

Deb K (2001). Multi-objective optimization using evolutionary algorithms. John Wiley & Sons.

Deb K, Pratap A, Agarwal S, Meyarivan T (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Transactions on Evolutionary Computation, 6: 182-197.

Elkan C (2001). The foundations of cost-sensitive learning. In: International joint conference on artificial intelligence, pp.973-978.

Erdem H (2010). Prediction of the moment capacity of reinforced concrete slabs in fire using artificial neural networks. Advances in Engineering Software, 41: 270-276.

Escombe AR, Oeser CC, Gilman RH, Navincopa M, Ticona E, Pan W, Martinez C, Chacaltana J, Rodriguez R, Moore DA, Friedland JS, Evans CA (2007). Natural ventilation for the prevention of airborne contagion. PLoS Med, 4: e68.

Fanger PO (1972). Thermal comfort: analysis and applications in environmental engineering. New York: McGraw-Hill.

Fernandez-Pello A, Lautenberger C, Rich D, Zak C, Urban J, Hadden R, Scott S, Fereres S (2015). Spot fire ignition of natural fuel beds by hot metal particles, embers, and sparks. Combustion science and technology, 187: 269-295.

Fonseca CM, Fleming PJ (1993). Genetic Algorithms for Multiobjective Optimization: FormulationDiscussion and Generalization. In: Icga, pp.416-423.

Forrester A, Sobester A, Keane A (2008). Engineering design via surrogate modelling: a practical guide. John Wiley & Sons.

Forrester AI, Sóbester A, Keane AJ (2007). Multi-fidelity optimization via surrogate modelling. Proceedings of the royal society of london a: mathematical, physical and engineering sciences, 463: 3251-3269.

Franklin PJ (2007). Indoor air quality and respiratory health of children. Paediatr Respir Rev, 8: 281-6.

Fraser A, Burnell, Donald (1970). Computer Models In Genetics.(Book Review). 147.

Fukushima K (1980). Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. Biological Cybernetics, 36: 193-202.

Galea ER, Wang Z, Veeraswamy A, Jia F, Lawrence PJ, Ewer J (2008). Coupled fire/evacuation analysis of the Station Nightclub fire. Fire Safety Science, 9: 465-476.

Gano SE, Renaud JE, Martin JD, Simpson TW (2006). Update strategies for kriging models used in variable fidelity optimization. Structural and Multidisciplinary Optimization, 32: 287-298.

Gilani SIUH, Khan MH, Pao W (2015). Thermal comfort analysis of PMV model prediction in air conditioned and naturally ventilated buildings. Energy Procedia, 75: 1373-1379.

Gorbett GE, Meacham BJ, Wood CB, Dembsey NA (2017). Structure and Evaluation of the Process for Origin Determination in Compartment Fires. Fire Technology, 53: 301-327.

Gottuk D, Mealy C, Floyd J (2008). Smoke transport and FDS validation. Fire Safety Science, 9: 129-140.

Gupta JK, Lin CH, Chen Q (2011). Transport of expiratory droplets in an aircraft cabin. Indoor Air, 21: 3-11.

Handcock MS, Stein ML (1993). A Bayesian analysis of kriging. Technometrics, 35: 403-410.

Hardwick J, Stout QF (2016). Adaptive Sampling Designs. University of Michigan. http://web.eecs.umich.edu/~qstout/AdaptSample.html.

Hartman J, Beyler A, Riahi S, Beyler C (2012). Smoke oxidation kinetics for application to prediction of clean burn patterns. Fire and Materials, 36: 177-184.

Hassan R, Cohanim B, De Weck O, Venter G (2005). A comparison of particle swarm optimization and the genetic algorithm. In: Proceedings of the 1st AIAA multidisciplinary design optimization specialist conference, pp.1-13.

He Y, Fernando A, Luo M (1998). Determination of interface height from measured parameter profile in enclosure fire experiment. Fire Safety Journal, 31: 19-38.

Hess JL, Smith AMO (1967). Calculation of potential flow about arbitrary bodies. Progress in Aerospace Sciences, 8: 1-138.

Hiyama K, Kato S, Ishida Y (2010). Thermal simulation: Response factor analysis using three-dimensional CFD in the simulation of air conditioning control. Building Simulation, 3: 195-203.

Hong S, You T, Kwak S, Han B (2015). Online tracking by learning discriminative saliency map with convolutional neural network. In: International Conference on Machine Learning, pp.597-606.

Horn J, Nafpliotis N, Goldberg DE (1994). A niched Pareto genetic algorithm for multiobjective optimization. In: Evolutionary Computation, 1994. IEEE World Congress on Computational Intelligence., Proceedings of the First IEEE Conference on, pp.82-87.

Howie R (1990). Respirator performance - the unpleasant reality. Institute of Occupational Medicine.

Hsu D-J, Chuang M-H (2012). In-vivo measurements of micrometer-sized particle deposition in the nasal cavities of taiwanese adults. Aerosol Science and Technology, 46: 631-638.

Iso (2005). ISO 7730:2005 Ergonomics of the thermal environment -- Analytical determination and interpretation of thermal comfort using calculation of the PMV and PPD indices and local thermal comfort criteria. ISO 7730:2005.

Jahn W, Gonzalez O, De Dios Rivera J, Torero JL (2015). Using Computational Fluid Dynamics in the forensic analysis of a prison fire. Forensic science international, 253: e33-e42.

Journel AG, Huijbregts CJ (1978). Mining geostatistics. Academic press.

Keeley JE, Safford H, Fotheringham C, Franklin J, Moritz M (2009). The 2007 southern California wildfires: lessons in complexity. Journal of Forestry, 107: 287-296.

Kelly FJ, Fussell JC (2015). Air pollution and public health: emerging hazards and improved understanding of risk. Environ Geochem Health, 37: 631-49.

Kelly JT, Asgharian B, Kimbell JS, Wong BA (2004). Particle deposition in human nasal airway replicas manufactured by different methods. Part I: Inertial regime particles. Aerosol Science and Technology, 38: 1063-1071.

Kennedy J (2001). Swarm intelligence. San Francisco: Morgan Kaufmann Publishers.

Kennedy J, Eberhart RC (1997). A discrete binary version of the particle swarm algorithm. In: Systems, Man, and Cybernetics, 1997. Computational Cybernetics and Simulation., 1997 IEEE International Conference on, pp.4104-4108.

Ko BC, Cheong K-H, Nam J-Y (2009). Fire detection based on vision sensor and support vector machines. Fire Safety Journal, 44: 322-329.

Kochetov N, Loktionov V, Sidorov A (2015). Using the Star CCM+ software system for modeling the thermal state and natural convection in the melt metal layer during severe accidents in VVER reactors. Thermal Engineering, 62: 663-672.

Konstantinov M, Wagner C (2014). Numerical simulation of the air flow and thermal comfort in a train cabin. In: Proceedings of the second international conference on railway technology: research, development and maintenance,

Konstantinov M, Wagner C (2015). Numerical simulation of the thermal comfort in a train cabin. International Journal of Railway Technology, 4: 69-88.

Koza JR, Bennett FH, Andre D, Keane MA 1996. Automated design of both the topology and sizing of analog electrical circuits using genetic programming. Artificial Intelligence in Design'96. Springer.

Krige DG (1951). A statistical approach to some basic mine valuation problems on the Witwatersrand. Journal of the Southern African Institute of Mining and Metallurgy, 52: 119-139.

Kumar A, Irsoy O, Ondruska P, Iyyer M, Bradbury J, Gulrajani I, Zhong V, Paulus R, Socher R (2016). Ask me anything: Dynamic memory

networks for natural language processing. In: International Conference on Machine Learning, pp.1378-1387.

Kwon H-B, Nam S-W, Kwak J-H (2009). Assessment of the pressure transient inside the passenger cabin of high-speed train using computational fluid dynamics. Journal of the Korean Society for Railway, 12: 65-71.

Laverge J, Janssens A (2013). Optimization of design flow rates and component sizing for residential ventilation. Building and Environment, 65: 81-89.

Lecun Y (1989). Generalization and network design strategies. Connectionism in perspective: 143-155.

Lecun Y, Bengio Y (1995). Convolutional networks for images, speech, and time series. The handbook of brain theory and neural networks, 3361: 1995.

Lecun Y, Boser BE, Denker JS, Henderson D, Howard RE, Hubbard WE, Jackel LD (1990). Handwritten digit recognition with a back-propagation network. In: Advances in neural information processing systems, pp.396-404.

Lee EW, Yuen RK, Lo S, Lam K, Yeoh G (2004). A novel artificial neural network fire model for prediction of thermal interface location in single compartment fire. Fire Safety Journal, 39: 67-87.

Lee TW, Singh H, Lee J, Jeong HM, Sturm D (2011). Computational simulations and optimization of flow and temperature distributions in a large-scale power plant building. Building Simulation, 4: 341-349.

Lee W, Yuen K, Lo S, Lam K (2000). Prediction of sprinkler actuation time using the artificial neural networks. J Build Surv, 2: 10-13.

Lewis JP (1995). Fast normalized cross-correlation. Vision interface, 10: 120-123.

Li F, Liu J, Pei J, Lin C-H, Chen Q (2014a). Experimental study of gaseous and particulate contaminants distribution in an aircraft cabin. Atmospheric Environment, 85: 223-233.

Li F, Liu JJ, Pei JJ, Lin CH, Chen QY (2014b). Experimental study of gaseous and particulate contaminants distribution in an aircraft cabin. Atmospheric Environment, 85: 223-233.

Li J, Cao X, Liu J, Wang C, Zhang Y (2015). Global airflow field distribution in a cabin mock-up measured via large-scale 2D-PIV. Building and Environment, 93: 234-244.

Li J, Liu J, Cao X, Jiang N (2016). Experimental study of transient air distribution of a jet collision region in an aircraft cabin mock-up. Energy and Buildings, 127: 786-793.

Li J, Liu J, Dai S, Guo Y, Jiang N, Yang W (2018a). PIV experimental research on gasper jets interacting with the main ventilation in an aircraft cabin. Building and Environment, 138: 149-159.

Li K, Su H, Chu J, Xu C (2012). A fast-POD model for simulation and control of indoor thermal environment of buildings. Building and Environment, 60: 150-157.

Li K, Xue W, Xu C, Su H (2013). Optimization of ventilation system operation in office environment using POD model reduction and genetic algorithm. Energy and Buildings, 67: 34-43.

Li M, Yan Y, Zhao B, Tu J, Liu J, Li F, Wang C (2018b). Assessment of turbulence models and air supply opening models for CFD modelling of airflow and gaseous contaminant distributions in aircraft cabins. Indoor Built Environment, 27: 606-621.

Li N, Cheung SCP, Li X, Tu J (2017). Multi-objective optimization of HVAC system using NSPSO and Kriging algorithms—A case study. Building Simulation, 10: 769-781.

Li X (2003a). A non-dominated sorting particle swarm optimizer for multiobjective optimization. In: Proceedings of Genetic and Evolutionary Computation Conference, Berlin, Heidelberg, pp.37-48.

Li X (2003b). A non-dominated sorting particle swarm optimizer for multiobjective optimization.

Li Z, Khananian A, Fraser RH, Cihlar J (2001). Automatic detection of fire smoke using artificial neural networks and threshold approaches applied to AVHRR imagery. IEEE Transactions on Geoscience and Remote Sensing, 39: 1859-1870.

Limane A, Fellouah H, Galanis N (2018). Three-dimensional OpenFOAM simulation to evaluate the thermal comfort of occupants, indoor air quality and heat losses inside an indoor swimming pool. Energy and Buildings, 167: 49-68.

Lin Z, Chow TT, Tsang CF, Chan LS, Fong KF (2005). Effect of air supply temperature on the performance of displacement ventilation (Part I)--thermal comfort.(using heating, ventilating and air-conditioning systems). Indoor and Built Environment, 14: 103.

Lin Z, Chow TT, Tsang CF, Fong KF, Chan LS, Shum WS, Tsai L (2009). Effect of internal partitions on the performance of under floor air supply ventilation in a typical office environment. Building and Environment, 44: 534-545.

Liu S, Schiavon S, Kabanshi A, Nazaroff WW (2017). Predicted percentage dissatisfied with ankle draft. Indoor Air, 27: 852-862.

Liu W, Duan R, Chen C, Lin C-H, Chen Q (2015). Inverse design of the thermal environment in an airliner cabin by use of the CFD-based adjoint method. Energy and Buildings, 104: 147-155.

Liu W, Mazumdar S, Zhang Z, Poussou SB, Liu J, Lin C-H, Chen Q (2012). State-of-the-art methods for studying air distributions in commercial airliner cabins. Building and Environment, 47: 5-12.

Liu W, Wen J, Lin C-H, Liu J, Long Z, Chen Q (2013). Evaluation of various categories of turbulence models for predicting air distribution in an airliner cabin. Building and Environment, 65: 118-131.

Lo S, Liu M, Zhang P, Yuen RK (2009). An artificial neural-network based predictive model for pre-evacuation human response in domestic building fire. Fire Technology, 45: 431-449.

Lu W-Z, Xue Y (2014). Prediction of particulate matter at street level using artificial neural networks coupling with chaotic particle swarm optimization algorithm. Building and Environment, 78: 111-117.

Luh GC, Lin CY (2011). Optimal design of truss-structures using particle swarm optimization. Computers and Structures, 89: 2221-2232.

Lundström JN, Olsson MJ (2010). Functional neuronal processing of human body odors. Vitamins and hormones, 83: 1-23.

Maeda EE, Formaggio AR, Shimabukuro YE, Arcoverde GFB, Hansen MC (2009). Predicting forest fire in the Brazilian Amazon using MODIS imagery and artificial neural networks. International Journal of Applied Earth Observation and Geoinformation, 11: 265-272.

Martínez SZ, Coello CaC (2013). Use of Gradient-Free Mathematical Programming Techniques to Improve the Performance of Multi-Objective Evolutionary Algorithms. Science.

Marxsen T (2016). New technology to cut Victoria's powerline fire risk. In: Arboriculture Australia Conference, Australia,

Mcaneney J, Chen K, Pitman A (2009). 100-years of Australian bushfire property losses: Is the risk significant and is it increasing? Journal of environmental management, 90: 2819-2822.

Mcarthur AG, Cheney NP, Barber J (1982). The fires of 12 February 1977 in the Western District of Victoria. CSIRO.

Mcgrattan KB, Baum HR, Rehm RG (1998). Large eddy simulations of smoke movement. Fire Safety Journal, 30: 161-178.

Milke JA, Mcavoy TJ (1995). Analysis of signature patterns for discriminating fire detection with multiple sensors. Fire Technology, 31: 120-136.

Miller C, Plucinski M, Sullivan A, Stephenson A, Huston C, Charman K, Prakash M, Dunstall S (2017). Electrically caused wildfires in Victoria, Australia are over-represented when fire danger is elevated. Landscape and Urban Planning, 167: 267-274.

Mirkhani SA, Gharagheizi F, Sattari M (2012). A QSPR model for prediction of diffusion coefficient of non-electrolyte organic compounds in air at ambient condition. Chemosphere, 86: 959-66.

Mousa WaY, Lang W, Auer T, Yousef WA (2017). A pattern recognition approach for modeling the air change rates in naturally ventilated buildings from limited steady-state CFD simulations. Energy and Buildings, 155: 54-65.

Nematchoua MK, Tchinda R, Ricciardi P, Djongyang N (2014). A field study on thermal comfort in naturally-ventilated buildings located in the equatorial climatic region of Cameroon. Renewable and Sustainable Energy Reviews, 39: 381-393.

Niu X-X, Suen CY (2012). A novel hybrid CNN–SVM classifier for recognizing handwritten digits. Pattern Recognition, 45: 1318-1325.

Okayama Y (1991). A primitive study of a fire detection method controlled by artificial neural net. Fire Safety Journal, 17: 535-553.

Parkhi OM, Vedaldi A, Zisserman A (2015). Deep Face Recognition. In: BMVC, pp.6.

Peri D, Rossetti M, Campana EF (2001). Design optimization of ship hulls via CFD techniques. Journal of Ship Research, 45: 140-149.

Pu L, Li Y, Xiao F, Ma Z, Qi D, Shen S (2014). Effects of different inlet vent positions on the uniformity of humidity inside a building chamber. Energy and Buildings, 76: 565-571.

Ravikumar P, Prakash D (2009). Analysis of thermal comfort in an office room by varying the dimensions of the windows on adjacent walls using CFD: A case study based on numerical simulation. Building simulation, 2: 187-196.

Redrow J, Mao SL, Celik I, Posada JA, Feng ZG (2011). Modeling the evaporation and dispersion of airborne sputum droplets expelled from a human cough. Building and Environment, 46: 2042-2051.

Riahi S, Beyler CL, Hartman J (2013). Wall smoke deposition from a hot smoke layer. Fire technology, 49: 395-409.

Ricciardi P, Buratti C (2015). Thermal comfort in the Fraschini theatre (Pavia, Italy): Correlation between data from questionnaires, measurements, and mathematical model. Energy and Buildings, 99: 243-252.

Ricciardi P, Ziletti A, Buratti C (2016). Evaluation of thermal comfort in an historical Italian opera theatre by the calculation of the neutral comfort temperature. Building and Environment, 102: 116-127.

Richardson LF (1922). Weather prediction by numerical process. Cambridge: University Press.

Russell BD, Benner CL, Wischkaemper JA (2012). Distribution feeder caused wildfires: Mechanisms and prevention. In: Protective Relay Engineers, 2012 65th Annual Conference for, pp.43-51.

Russell SJ, Norvig P (2016). Artificial intelligence: a modern approach. Malaysia; Pearson Education Limited.

Safi Y, Bouroumi A (2013). Prediction of forest fires using artificial neural networks. Applied Mathematical Sciences, 7: 271-286.

Santos LS, Biscaia Jr EC, Pagano RL, Calado VMA (2012). CFD-optimization algorithm to optimize the energy transport in pultruded polymer composites. Brazilian Journal of Chemical Engineering, 29: 559-566.

Shen T-S, Huang Y-H, Chien S-W (2008). Using fire dynamic simulation (FDS) to reconstruct an arson fire scene. Building and environment, 43: 1036-1045.

Sivia D, Skilling J (2006). Data analysis: a Bayesian tutorial. OUP Oxford.

Srinivas N, Deb K (1994). Muiltiobjective optimization using nondominated sorting in genetic algorithms. Evolutionary computation, 2: 221-248.

Stavrakakis GM, Karadimou DP, Zervas PL, Sarimveis H, Markatos NC (2011). Selection of window sizes for optimizing occupational comfort and hygiene based on computational fluid dynamics and neural networks. Building and Environment, 46: 298-314.

Stavrakakis GM, Markatos NC (2009). Simulation of airflow in one- and two-room enclosures containing a fire source. International Journal of Heat and Mass Transfer, 52: 2690-2703.

Stein ML (2012). Interpolation of spatial data: some theory for kriging. Springer Science & Business Media.

Stephenson C, Handmer J, Betts R (2013). Estimating the economic, social and environmental impacts of wildfires in Australia. Environmental Hazards, 12: 93-111.

Talbot L, Cheng R, Schefer R, Willis D (1980). Thermophoresis of particles in a heated boundary layer. Journal of Fluid Mechanics, 101: 737-758.

Taroni F, Aitken CG, Garbolino P, Biedermann A (2006). Bayesian networks and probabilistic inference in forensic science. Wiley Online Library.

Teague B, Mcleod R, Pascoe S (2010). Final report, 2009 Victorian bushfires royal commission. Parliament of Victoria, Melbourne Victoria, Australia.

Tian W, Han X, Zuo W, Sohn MD (2018). Building energy simulation coupled with CFD for indoor environment: A critical review and recent applications. Energy and Buildings.

Tinsley A, Gorbett G (2013). Fire investigation origin determination survey. Fire and Arson Investigator Journal of the International Association of Arson Investigators, 63: 24-40.

Tu J, Yeoh GH, Liu C (2008). Computational fluid dynamics a practical approach. Amsterdam, Boston.

Tucker AB (2004). Computer science handbook. CRC press.

Turner WJ, Awbi HB (2015). Residential hybrid ventilation: Airflow and heat transfer optimisation of a convector using computational fluid dynamics. In: Building Simulation, pp.65-72.

Uic (2005). International Union of Railways Standard, Heating, Ventilation and Air Conditioning in Coaches. Paris, French: Internaionla Union of Railways.

Urban JL, Zak CD, Fernandez-Pello C (2015). Cellulose spot fire ignition by hot metal particles. Proceedings of the Combustion Institute, 35: 2707-2714.

Valent P (1984). The Ash Wednesday bushfires in Victoria. Medical Journal of Australia, 141: 268-270.

Van Beers W, Kleijnen JP (2004). Kriging interpolation in simulation: a survey. Proceedings of the 2004 Winter Simulation Conference, 1.

Varol Y, Avci E, Koca A, Oztop HF (2007). Prediction of flow fields and temperature distributions due to natural convection in a triangular enclosure using Adaptive-Network-Based Fuzzy Inference System (ANFIS) and Artificial Neural Network (ANN). International Communications in Heat and Mass Transfer, 34: 887-896.

Vivchar A (2011). Wildfires in Russia in 2000–2008: estimates of burnt areas using the satellite MODIS MCD45 data. Remote Sensing Letters, 2: 81-90.

Wang H, Lin M, Chen Y (2014). Performance evaluation of air distribution systems in three different China railway high-speed train cabins using numerical simulation. In: Building simulation, pp.629-638.

Wang H, Zhai Z (2016). Advances in building simulation and computational techniques: A review between 1987 and 2014. Energy and Buildings, 128: 319-335.

Wang J-F, Piechna J, Mueller N (2012). A novel design of composite water turbine using CFD. Journal of Hydrodynamics, Ser. B, 24: 11-16.

Wang X, Wang X, Xing G, Chen J, Lin C-X, Chen Y (2013). Intelligent sensor placement for hot server detection in data centers. Parallel and Distributed Systems, IEEE Transactions on, 24: 1577-1588.

Whittaker J, Haynes K, Handmer J, Mclennan J (2013). Community safety during the 2009 Australian 'Black Saturday'bushfires: an analysis

of household preparedness and response. International journal of wildland fire, 22: 841-849.

Wikle CK, Berliner LM (2007). A Bayesian tutorial for data assimilation. Physica D: Nonlinear Phenomena, 230: 1-16.

Wolkoff P (1991). The possible role of TVOC in the indoor environment. Comments on "The effect of ventilation and air pollution on perceived indoor air quality in five town halls". Energy and Buildings, 17: 253-254.

Wolkoff P (2013). Indoor air pollutants in office environments: assessment of comfort, health, and performance. Int J Hyg Environ Health, 216: 371-94.

Yan Y, Li X, Tu J (2014). Numerical study of passenger thermal effects on the transport characteristics of exhaled droplets in an airliner cabin. In: 13th International Conference on Indoor Air Quality and Climate 2014, pp.1-8.

Yan Y, Li X, Yang L, Tu J (2016). Evaluation of manikin simplification methods for CFD simulations in occupied indoor environments. Energy and Buildings, 127: 611-626.

Yang L, Li M, Li X, Tu J (2018). The effects of diffuser type on thermal flow and contaminant transport in high-speed train (HST) cabins – a numerical study. International Journal of Ventilation, 17: 48-62.

Yang X, Wang J, He S (2012). A SVM approach for vessel fire detection based on image processing. Proceedings of International Conference on Modelling, Identification & Control (ICMIC): 150-153.

Yeoh G, Yuen R, Chen D, Kwok W (2002a). Combustion and heat transfer in compartment fires. Numerical Heat Transfer: Part A: Applications, 42: 153-172.

Yeoh G, Yuen R, Lee E, Chueng S (2002b). Fire and smoke distribution in a two-room compartment structure. International Journal of Numerical Methods for Heat & Fluid Flow, 12: 178-194.

You R, Zhang Y, Zhao X, Lin C-H, Wei D, Liu J, Chen Q (2018). An innovative personalized displacement ventilation system for airliner cabins. Building and Environment, 137: 41-50.

Yuan X, Chen Q, Glicksman LR, Hu Y, Yang X (1999). Measurements and computations of room airflow with displacement ventilation. ASHRAE Transactions, 105: PART 1/.

Yuen A, Yeoh G, Alexander R, Cook M (2014). Fire scene reconstruction of a furnished compartment room in a house fire. Case Studies in Fire Safety, 1: 29-35.

Yuen RK, Lee EW, Lo S, Yeoh G (2006). Prediction of temperature and velocity profiles in a single compartment fire by an improved neural network analysis. Fire safety journal, 41: 478-485.

Zhai Z, Xue Y, Chen Q (2014). Inverse design methods for indoor ventilation systems using CFD-based multi-objective genetic algorithm. Building Simulation, 7: 661-669.

Zhang L, Li Y (2012). Dispersion of coughed droplets in a fully-occupied high-speed rail cabin. Building and Environment, 47: 58-66.

Zhao B (2009). Modeling pressure drop coefficient for cyclone separators: a support vector machine approach. Chemical Engineering Science, 64: 4131-4136.

Zhao J, Zhang Z, Han S, Qu C, Yuan Z-Y, Zhang D (2011). SVM based forest fire detection using static and dynamic features. Comput. Sci. Inf. Syst., 8: 821-841.

Zhou L, Haghighat F (2009a). Optimization of ventilation system design and operation in office environment, Part I: Methodology. Building and Environment, 44: 651-656.

Zhou L, Haghighat F (2009b). Optimization of ventilation systems in office environment, Part II: Results and discussions. Building and Environment, 44: 657-665.

Zhuang R, Li X, Tu J (2014). Should different gaseous contaminants be treated differently in CFD indoor simulations? Air Pollution XXII, 183: 353.