



**University of
Nottingham**
UK | CHINA | MALAYSIA

**Generative modeling-aided data-driven fault detection and
diagnosis for heating, ventilation and air-conditioning systems of
next-generation intelligent buildings**

Fangliang Zhong

Department of Architecture and Built Environment
University of Nottingham

November, 2023

A thesis submitted in partial fulfilment of the regulations for the Degree of
Doctor of Philosophy at the University of Nottingham

Acknowledgements

Embarking on the journey pursuing a PhD is an enduring expedition marked by triumphs and setbacks, which requires a blend of courage, curiosity and self-motivation. First of all, I would like to thank myself for my courage for initiating my PhD journey, my curiosity in applying artificial intelligence (AI) techniques within the realm of the built environment, and my persistence propelling me forward after research setbacks and life's obstacles.

Secondly, I would like to express my deepest gratitude to my supervisors Dr. John Calautit and Prof. Yupeng (Jack) Wu, whose valuable guidance and support have been instrumental in shaping my research. To John, thanks for providing the computational resource essential for my conducting AI research and his dedicated efforts into steering me toward presenting my research rigorously and concisely in journal papers and conference presentations. To Jack, I would like to appreciate your consistently offering invaluable advice on crafting impactful journal papers.

Thirdly, I would like to convey my deepest appreciation to the collective efforts of those who generously provide invaluable online educational resources in AI and Python programming. In particular, I would like to express my sincere thanks to Prof. Andrew Ng and the team at DeepLearning.AI who devotes themselves to empowering aspiring AI learners like me to delve into the realms of AI technology.

Lastly and most importantly, I would like to express my inexpressible gratitude to my parents and grandparents, who unwaveringly support me in pursuing my PhD. 谢谢我的父母和外公外婆对我追求博士学位的坚持不懈的支持。I am grateful to my dog, Liuyi, whose unconditional love has been a wellspring of my spiritual power. In addition, I would also thank my friends Zihan Li and Sihui Liu who are always emotionally available, especially under the toughest phases of the pandemic throughout my PhD journey. I am also thankful to all the other friends who shared their daily joys with me during my PhD journey.

Abstract

Heating, ventilation and air-conditioning (HVAC) systems are responsible for 40-50% of the total building energy usage. The occurrences of HVAC faults lead to an additional 20% or more energy waste for HVAC systems. The automated fault detection and diagnosis (FDD) has the potential to reduce HVAC energy consumption by 10–40%. Recently, data-driven FDD methods have become the mainstream, due to the advancement of data science techniques and the increasing data availability in buildings. But there are still major challenges for achieving high data-driven FDD performances in practice, such as the class imbalance between fault data and normal data. Previous studies have evaluated the impact of this issue on FDD performances, and also proposed solutions to this practical issue. However, the influence of another data imbalance issue, the fault data intra-class imbalance under different climate conditions, has not yet been explored and addressed. The absence of fault data under certain climatic condition(s) may also affect FDD performance under these climatic condition(s).

This thesis assesses the impacts of HVAC single and multiple faults under different climatic conditions, and also evaluates HVAC FDD performances under different fault data seasonal intra-class imbalanced scenarios. To conduct the fault impact analyses under different climatic conditions, a series of typical variable air volume (VAV) system faults were modelled using Python scripts and a widely-used building performance simulation tool, EnergyPlus. The fault impact simulations were performed by introducing developed fault models into a verified prototype office building model. Based on fault simulations, the fault impacts under different seasons and climate periods were analysed. The results show that the fault impacts are affected by climatic condition variations and also present seasonal patterns.

Based on the insights acquired from fault impact analyses, the FDD performance for a combined convolutional and recurrent neural network classifier was assessed under fault data seasonal intra-class imbalanced scenarios, and an ideal fault data scenario covering climatic conditions from all the seasons. The developed FDD model was trained using four simulated data scenarios representing fault data seasonal intra-class imbalanced scenarios, and an ideal scenario. The trained model under each scenario was tested using each season's dataset. Moreover, further experiments based on real fault datasets were used to validate the findings from the experiments based on simulated fault datasets. As a result, the fault data experiments

revealed two key findings: (1) The FDD architectures trained on sufficient fault data under a certain season(s) show poor generalization ability to detect and identify the faults under other unseen seasons; (2) The coverage of fault data under different seasons is more crucial in enhancing the FDD performances than the amount of fault data under a certain season(s).

To address this practical issue leading to FDD performance degradation under FDD model's unseen seasons, a multi-label conditional Wasserstein generative adversarial network (ML-CWGAN) by incorporating both fault and season labels into its framework, was developed. The ML-CWGAN was integrated with the baseline FDD classifier to evaluate the quality of synthetic fault data. The FDD performances led by different GAN variants, i.e., multi-label conditional GAN (ML-CGAN), single-label CWGAN and the proposed ML-CWGAN, were evaluated and compared under 14 fault data seasonal imbalanced scenarios and 2 seasonal balanced but insufficient scenarios. The proposed ML-CWGAN demonstrated its ability to adapt to a wider range of seasonal imbalanced scenarios and was more robust in generating high quality fault samples. The results indicated that the ML-CWGAN considerably improved the baseline FDD performances by up to 5% in terms of F1 score, and was able to enhance FDD performances under all imbalanced scenarios. Compared with the SL-CWGAN, the ML-CWGAN-aided FDD outperformed the SL-CWGAN-aided FDD by an average F1 score of 1% and the SL-CWGAN performed poorly in improving FDD performances under some scenarios.

The future directions can be explored in two aspects, i.e., fault modeling and impact analysis aspect and FDD aspect. Future studies can explore HVAC fault impacts under different climate zones, such as hot climates, to compare the differences in fault impacts. The fault impacts from other types of HVAC systems, apart from VAV systems, are worth exploring. Besides, evolving fault intensity and fault occurrence frequency can be considered in fault modeling for simulation-based fault impact analyses. As for the FDD aspect, evaluating the impact of the fault data intra-class imbalance among different working conditions on HVAC FDD performances can be considered as a future direction. The integration of a high-quality synthetic fault sample selection protocol into the proposed framework is required to filter out high-quality fault samples. Moreover, Other variants of GANs, such as auxiliary classifier GAN, can be attempted to generate synthetic fault samples in future.

Table of Contents

Acknowledgements	i
Abstract	ii
Table of Contents	iv
List of Figures	x
List of Tables	xvii
List of Publications	xx
Nomenclature	xxii
Abbreviations	xxvi
Chapter 1 Introduction	1
1.1 Research Background	1
1.2 Fault Detection and Diagnosis Challenges	4
1.2.1 Data-side Challenges	4
1.2.2 Method-side Challenges	6
1.3 Motivation and Research Questions	6
1.4 Aims and Objectives	7
1.5 Thesis Structure	8
Chapter 2 Literature Review	10
2.1 Introduction	10
2.1.1 Literature Search Method	10
2.2 Fault Modeling Review	11
2.2.1 Fault Modeling Approach Category	12
2.2.2 Fault Modeling Platforms	13
2.2.3 Fault Characteristics	15
2.2.4 Fault Model Descriptions	20
2.2.4.1 Control Faults	21
2.2.4.2 Sensor Faults	23
2.2.4.3 Chiller Faults	29

2.2.4.4 RTU and Split Air Conditioner Faults-----	36
2.2.4.5 AHU Faults-----	39
2.2.4.6 VAV Terminal Faults-----	45
2.2.4.7 Other Faults -----	46
2.3 Fault Impact Analysis Review -----	51
2.3.1 Single Fault Impact Analysis -----	51
2.3.1.1 Simulation-based Studies-----	52
2.3.1.2 Experiment-based Studies-----	56
2.3.2 Multiple Faults Impact Analysis-----	59
2.3.2.1 Simulation-based Studies-----	59
2.3.2.2 Experimental-based Studies-----	62
2.3.3 Summary and Research Gaps -----	62
2.4 Data-driven Fault Detection and Diagnosis (FDD) Review-----	63
2.4.1 Data-driven FDD Fundamentals-----	64
2.4.2 Data-driven FDD Method Perspective-----	64
2.4.2.1 Supervised Learning-based FDD Methods -----	65
2.4.2.1.1 Supervised Classification Methods -----	65
2.4.2.1.2 Supervised Regression Methods -----	71
2.4.2.2 Semi-supervised Learning-based FDD Methods -----	73
2.4.2.3 Unsupervised Learning-based FDD Methods -----	75
2.4.3 Data-driven FDD Practical Challenges and Potential Solutions -----	78
2.4.3.1 Labeled Fault Data Insufficiency -----	79
2.4.3.2 Fault and Normal Data Class Imbalance -----	80
2.4.3.3 Poor Cross-Operation-Condition Generalization -----	83
2.4.4 Summary and Research Gaps -----	84
2.5 Research Gaps Summary -----	85
Chapter 3 Fault Modeling and Simulation -----	87
3.1 General Workflow-----	87
3.2 Fault Model Development -----	88
3.2.1 Control Faults -----	90
3.2.2 Sensor Faults -----	91
3.2.3 Packaged Air Conditioner Faults -----	92

3.2.4 VAV Terminal Unit Faults -----	93
3.2.5 Other Faults-----	94
3.3 Case Study Reference Building Model-----	94
3.3.1 Building and HVAC Systems Descriptions-----	95
3.3.2 Model Verification -----	97
3.4 Contemporary and Future Weather Data Generation -----	98
3.4.1 IPCC Climate Scenarios -----	99
3.4.2 UKCP09 Weather Generator Introduction -----	100
3.4.3 Meteonorm Introduction-----	101
3.4.4 Meteonorm Future Climate Change Data Generation -----	101
3.5 Fault Simulation -----	104
3.5.1 Fault Simulation Inputs and Settings -----	104
3.5.2 Fault Simulation Outputs -----	105
Chapter 4 Fault Impact Analysis under Changing Climatic Conditions-----	107
4.1 Fault Impact Analysis Overview -----	107
4.2 Fault Impact Analysis across Different Climate Periods -----	107
4.2.1 System Energy and Thermal Comfort Impact Ranking Indicators -----	108
4.2.2 Single Fault Impact and Rankings -----	109
4.2.2.1 Cause and Impact Analysis for Prioritized Single Faults -----	111
4.2.3 Multiple Faults Impact and Rankings-----	116
4.2.3.1 Multiple Faults Synergetic Effect -----	119
4.2.3.2 Multiple Faults Antagonistic Effect -----	122
4.2.4 Fault Impact Variations across Different Climate Periods-----	127
4.2.4.1 Single Fault Impact Variations -----	127
4.2.4.2 Multiple Fault Combinations Impact Variations-----	131
4.2.5 Comparisons With Findings From Previous Studies -----	133
4.2.6 Key Findings and Potential Future Work -----	135
4.3 Climate Change-driven Fault Impact Seasonal Pattern Variations -----	136
4.3.1 Enhanced Fault Modeling and Simulation Workflow -----	137
4.3.2 Season Onset and End Dates Predictions -----	137
4.3.3 Seasonal Fault Impact Indicators -----	138

4.3.4 Fault Impact Seasonal Patterns Analysis -----	139
4.3.5 Influence of Shifting Seasons on Fault Impact Seasonal Patterns -----	140
4.3.6 Potential Applications of Seasonal Fault Impact Analysis -----	142
4.3.7 Key Findings and Potential Future Work -----	143
4.4 Fault Impact Analysis Summary -----	143
Chapter 5 Data-driven Fault Detection and Diagnosis Model Development -----	145
5.1 General Workflow -----	145
5.2 Fault Data Acquisition -----	146
5.2.1 Description of Simulated Fault Dataset -----	146
5.2.2 Description of Real Fault Datasets -----	147
5.3 Data Pre-processing Techniques -----	149
5.3.1 Preliminary Feature Selection from Raw Data -----	150
5.3.2 Outlier Detection -----	152
5.3.3 Data Normalization -----	153
5.4 Sliding Window -----	154
5.5 Dataset Splitting -----	155
5.6 Model Architecture and Hyperparameters -----	155
5.6.1 Preliminary Theory for the Architecture -----	156
5.6.1.1 Convolutional Neural Network -----	156
5.6.1.2 Recurrent Neural Network -----	158
5.6.2 Proposed Architecture Structure and Hyperparameters -----	159
5.7 Model Evaluation -----	162
Chapter 6 Fault Detection and Diagnosis Evaluation Experiments -----	164
6.1 Introduction -----	164
6.2 Potential Practical Fault Data Scenarios -----	164
6.3 Evaluation Experiments Setup -----	165
6.3.1 Initial Data Experiments for Tuning General Hyperparameters -----	166
6.3.2 Practical Data Scenarios using Simulated Dataset -----	166
6.3.3 Verification Experiments using Real Building Datasets -----	167

6.4 General Hyperparameter Analysis Results	168
6.5 Evaluation Experiments Results based on Simulated Dataset	170
6.5.1 Experiment 1 Results: One-Season Dataset Evaluation	173
6.5.2 Experiment 2 Results: Dual-Seasons Dataset Evaluation	175
6.5.3 Experiment 3 Results: Full-Season Quarter-Size Dataset Evaluation	177
6.5.4 Experiment 4 Results: Full-Season Full-Size Dataset Evaluation	178
6.6 Practical Issue Verification Experiments Results	179
6.7 Fault Detection and Diagnosis Evaluation Summary	183
Chapter 7 Integration of Generative Modeling with Data-driven Fault Detection and Diagnosis	185
7.1 Introduction	185
7.2 Generative Adversarial Network Introduction	185
7.2.1 Conditional Generative Adversarial Network (CGAN)	187
7.2.2 Wasserstein Generative Adversarial Network (WGAN)	187
7.2.3 Conditional Wasserstein Generative Adversarial Network (CWGAN)	188
7.3 Generative Adversarial Network (GAN)-aided FDD Framework	188
7.4 Generative Adversarial Network Evaluation Experiments Setup	190
7.4.1 Fault Data Amount and Imbalance Ratio Selection	190
7.4.2 Baseline FDD Performance Evaluation Experiments	191
7.4.2.1 Baseline FDD Evaluation Data Scenarios	191
7.4.2.2 Baseline FDD Evaluation Experiments Setup	192
7.4.3 GAN-aided FDD Evaluation Experiments	193
7.4.3.1 GAN-aided Data Scenarios	193
7.4.3.2 Generative Adversarial Network Structure	194
7.4.3.3 Generative Adversarial Network Hyperparameters Tuning	195
7.4.3.4 GAN-aided FDD Evaluation Experiment Setup	196
7.5 Baseline FDD Evaluation Experiments Results	197
7.6 GAN-aided FDD Performance Results	200
7.6.1 ML-CWGAN-aided FDD Performance	201
7.6.2 FDD Performance Enhancement: ML-CGAN vs ML-CWGAN	203

7.6.3 FDD Performance Enhancement: ML-CWGAN vs SL-CWGAN -----	204
7.6.4 FDD Performance Gap: Aggregated Real and Synthetic Samples vs Real Samples -----	205
7.7 Key Findings and Potential Future Work-----	206
Chapter 8 Conclusions -----	208
8.1 Summary of Main Contributions and Key Findings-----	208
8.2 Research Limitations -----	210
8.3 Future Work-----	211
References -----	213

List of Figures

Figure 1.1 Global final energy consumption by sector in 2021 (IEA, 2022). -----	1
Figure 1.2 Energy end usages for (a) commercial (EIA, 2016) and (b) residential buildings (EIA, 2018) according to EIA's surveys. -----	2
Figure 1.3 Typical HVAC faults in a variable air volume (VAV) system (Zhang and Hong, 2017). -----	3
Figure 1.4 FDD classification scheme proposed by Katipamula and Brambley (2015). --	3
Figure 1.5 An illustration of FDD classification under (a) sufficiently balanced data; (b) imbalanced data; (c) insufficient data (the red line(s) in each case represent the separation boundaries for different data classes predicted by the FDD model). -----	5
Figure 1.6 Research workflow diagram of the present thesis. -----	9
Figure 2.1 Fault modeling structures on HVACSIM+ (Li and Wen, 2010).-----	14
Figure 2.2 An example of fault evolution from Kim et al. (2019). -----	16
Figure 2.3 An example of fault occurrence rate distribution for an HVAC system (Otto et al., 2012). -----	17
Figure 2.4 Comparison between HVAC setback faults and normal case (Kim et al., 2019).-----	23
Figure 2.5 Diagram of the main components and working mechanisms of a typical water- cooled chiller (Xu, Xiao and Wang, 2008). -----	30
Figure 2.6 Diagram of components in a typical AHU (Yu, Woradechjumroen and Yu, 2014).-----	39
Figure 2.7 Fan and duct characteristics curves at normal and fouling condition (Cheung and Braun, 2015). -----	47
Figure 2.8 The influence of dirty air filter on variable speed fan performance with the airflow rate maintained (Zhang and Hong, 2017). -----	50
Figure 2.9 The influence of dirty air filter on variable speed fan performance with the reduced airflow rate (Zhang and Hong, 2017). -----	50

Figure 2.10 The influence of dirty air filter on constant speed fan performance (Zhang and Hong, 2017).	51
Figure 2.11 Distributions of the papers about single fault impact analysis among various simulation platforms and the experiment approach.	52
Figure 2.12 Distributions of the papers about multiple faults impact analysis among various simulation platforms and the experiment approach.	59
Figure 2.13 Diagram of a SVM-based classification. (Zhao et al., 2019)	66
Figure 2.14 An example of a decision tree algorithm, i.e., CART. (Li et al., 2018)	67
Figure 2.15 Diagram of a FDD method using ANN. (Zhao et al., 2019)	68
Figure 2.16 Diagram of a RNN architecture. (Taheri et al., 2021)	69
Figure 2.17 Diagram of a 1D CNN for sensor signal classification. (Cheng et al., 2021)	70
Figure 2.18 Diagram of a 2D CNN for image data classification (Fan et al., 2022)	70
Figure 2.19 Diagram of a GAN architecture. (Li et al., 2021a)	75
Figure 2.20 PCA-based fault detection demonstration. (Zhao et al., 2019)	77
Figure 2.21 A clustering-based FDD method. (Chen et al., 2022)	78
Figure 3.1 (a) The workflow for evaluating HVAC fault impacts; (b) Detailed fault modeling and simulation steps.	88
Figure 3.2 (a) Case study medium office building model; (b) One core and four perimeter zones on the floor plan; (c) A typical VAV system layout with reheat coils and outdoor air economizer (IES, 2016).	97
Figure 3.3. CV(RMSE) and NMBE for system total and end energy usages of the present original model and the modified model, compared with the DOE source results.	98
Figure 3.4 Schematic diagram of SRES scenarios (IPCC, 2000).	100
Figure 3.5 Weather station selection for future climate change data generation.	101
Figure 3.6 Baseline climate database selection for future climate change data generation.	102
Figure 3.7 Calculation settings for future climate change data generation.	103
Figure 3.8 Weather data output format for future climate change data generation.	103

Figure 3.9 Results export for future climate change data generation. -----	104
Figure 4.1. Single fault impacts under the current climate period: (a) System energy usage deviations; (b) PPD deviations for the top-ranked faults. -----	113
Figure 4.2. Single fault impacts under the 2030s climate conditions: (a) System energy usage deviations; (b) PPD deviations for the top-ranked faults. -----	114
Figure 4.3. Single fault impacts under the 2050s climate period: (a) System energy usage deviations; (b) PPD deviations for the top-ranked faults. -----	115
Figure 4.4. Causes for the impacts of single faults: (a) Supply fan stuck at maximum speed; (b) Supply fan stuck at minimum speed; (c) No overnight setback and OA damper stuck fully closed; (d) Cooling coil SA temperature sensor positive bias; (e-f) Thermostat positive offset. -----	116
Figure 4.5. Multiple faults interactions for the top 5 combinations under the current climate period: (a) Difference between actual combined effect on (a) system energy consumptions; (b) PPD, and the algebraic sums of each single-fault impact. -----	124
Figure 4.6. Causes for multiple faults (a-e) synergetic effects and (f-h) antagonistic effects.-----	125
Figure 4.7. Variations across climate periods in the difference between actual combined effect and algebraic sum (‘Difference’ in the figures) of (a) air system total electric energy; (b) heating coil gas energy; (c) reheating coil electric energy; (d) cooling coil electric energy; (e) fan electric energy; (f) PPD for the top 5 fault combinations; (g) heating coil gas energy (h) reheating coil electric energy; (i) cooling coil electric energy; (j) air system total electric energy for the fault combinations with severe antagonistic effects.-----	127
Figure 4.8. The antagonistic effects of multiple faults combinations on system energy consumptions under the current climate period.-----	127
Figure 4.9. (a-f) Variations in air system total electric energy, cooling coil electric energy, heating coil gas energy, reheating coil electric energy, fan electric energy and PPD for single faults from the current to the 2030s period (in solid fill & black labels) and from the current to the 2050s period (in dash-dot line & red labels). (First label for each bar: the variation amount, second label: the variation percentage); Variations in	

OA dry-bulb temperature and relative humidity between (g) the current and 2030s periods; (h) the current and 2050s periods. -----	130
Figure 4.10. (a-f) Variations in air system total electric energy, cooling coil electric energy, heating coil gas energy, reheating coil electric energy, fan electric energy and PPD for multiple faults combinations from the current period to the 2030s period (in solid fill & black labels) and from the current period to the 2050s period (in dash-dot line & red labels). (First label for each bar: the variation amount, second label: the variation percentage). -----	133
Figure 4.11. The workflow for the proposed method including four main parts: (1) fault implementation to the case study building (orange block), (2) contemporary and future weather data acquisition (grey block), (3) future season onsets and ends predictions (green block) and (4) fault simulations (red block). -----	137
Figure 4.12 An example of the third-degree polynomial fit (red dot line) of the raw daily temperature observations and intersections with summer (green dash line) and winter (orange dash line) thresholds. -----	138
Figure 4.13 (a) System total electric energy usage; (b) System total gas energy usage; (c) Zone average PPD for all the faults under seasons of current, 2030s and 2050s climate periods. (the label on each grid represents the exact amount of the parameter rather than the variation).-----	141
Figure 4.14 (a) System electricity and (b) system gas cost distributions for all the faults under seasons of current, 2030s and 2050s climate periods (blue dash lines indicate the baseline cost).-----	142
Figure 5.1 Data-driven fault detection and diagnosis model development workflow.--	146
Figure 5.2 (a) The layout of Iowa Energy Resource Station; (b) a 3D model of the FRP (Buckberry and Bhandari, 2012).-----	148
Figure 5.3 The max-min normalization method. -----	154
Figure 5.4 The 1D-sliding window method for time-series data with a dimension of $n \times m$. -----	155
Figure 5.5 The proposed combined CNN-RNN architecture (the recurrent layers were selected from LSTM and GRU). -----	156

Figure 5.6 An example of 1D convolution operation. -----	157
Figure 5.7 (a) A typical recurrent layer; (b) The basic structures for vanilla RNN, GRU and LSTM. -----	159
Figure 6.1 Potential examples of FDD classification for practical data scenarios (a-c) and the ideal data scenario (d) (the red dash line in each case represents the possible separation boundary predicted by the FDD model to differentiate the normal data from the fault A data). -----	165
Figure 6.2 The proposed data experiments under the simulated fault dataset: (a) Training data: a one-season dataset; testing data: each season’s dataset (spring as an example here). (b) Training data: a mixed dataset of two seasons of half size; testing data: each season’s dataset (a mixed dataset of spring and summer as an example here). (c) Training data: a mixed dataset of four seasons of quarter size; testing data: each season’s dataset. (d) Training data: a mixed dataset of four seasons of full size; testing data: each season’s dataset. -----	167
Figure 6.3. The proposed data experiments under the ASHRAE RP1312 fault dataset: (a) Training data: a one-season dataset; testing data: each season’s dataset (spring as an example here). (b) Training data: a mixed dataset of two seasons of half size; testing data: each season’s dataset (a mixed dataset of spring and summer as an example here). (c) Training data: a mixed dataset of three seasons of quarter size; testing data: each season’s dataset. (d) Training data: a mixed dataset of three seasons of full size; testing data: each season’s dataset. -----	168
Figure 6.4. The proposed data experiments under the NREL’s fault dataset: (a) Training data: the summer dataset; testing data: each season’s dataset. (b) Training data: the winter dataset; testing data: each season’s dataset. (c) Training data: a mixed dataset of summer and winter of half size from each; testing data: each season’s dataset. -----	168
Figure 6.5. FDD performance comparisons among different (a) window lengths; (b) training data size ratios; (c) batch sizes. -----	170
Figure 6.6. Highest F1 scores between architectures trained on (a) one-season, (b) dual-seasons and full-seasons datasets; highest CK scores between architectures trained on (c) one-season, (d) dual-seasons and full-seasons datasets (‘All 1/4 size’ and ‘All full size’ represent the full-season quarter size and full-season full size scenarios, respectively).	171

Figure 6.7 Fault detection, fault diagnosis, false negative and false positive rates tested under each season for experiments 1–4 using the simulated dataset (‘All 1/4’ represents the full-season quarter size scenario; ‘All full’ represents the full-season full size scenario). -----	173
Figure 6.8 Scatterplots for the hidden features extracted by the best architecture CNN-LSTM tested under (a) spring; (b) summer; (c) autumn; (d) winter; in experiment 1 using the t-SNE method. -----	175
Figure 6.9 Scatterplots for the hidden features extracted by the best architecture CNN-GRU tested under (a) spring; (b) summer; (c) autumn; (d) winter; in experiment 2 using the t-SNE method. -----	176
Figure 6.10 Scatterplots for the hidden features extracted by the best architecture CNN-LSTM tested under (a) spring; (b) summer; (c) autumn; (d) winter; in experiment 3 using the t-SNE method. -----	178
Figure 6.11 Scatterplots for the hidden features extracted by the best architecture CNN-GRU tested under (a) spring; (b) summer; (c) autumn; (d) winter; in experiment 4 using the t-SNE method. -----	179
Figure 6.12 Fault detection rate, fault diagnosis rate, false negative rate and false positive rate tested under each season for experiments 1–4 using the ASHRAE RP1312 dataset (‘All 1/4’ represents the full-season quarter size scenario; ‘All full’ represents the full-season full size scenario). -----	182
Figure 6.13 Fault detection rate, fault diagnosis rate, false negative rate and false positive rate tested under each season for experiments 1–4 using the NREL’s dataset.	183
Figure 7.1 The diagram of a typical original GAN framework (Yan et al., 2020). -----	186
Figure 7.2 The diagram of a CGAN framework (Yan et al., 2020). -----	187
Figure 7.3 (a) The diagram of a Multi-label CWGAN (ML-CWGAN) framework; (b) The diagram of the entire GAN-aided FDD Framework. -----	189
Figure 7.4 The GAN structure for the proposed ML-CWGAN and other GAN-based architectures. -----	195
Figure 7.5 The test F1 score per season for (a) CNN-GRU classifier; (b) CNN-LSTM classifier under each data scenario. -----	199

Figure 7.6 Season-average F1 scores for CNN-GRU and CNN-LSTM classifiers under each data scenario. -----	200
Figure 7.7 (a) The each season's F1 score variation between the ML-CWGAN-aided performance and the baseline performance under each data scenario; (b) The F1 score per season for the ML-CWGAN-aided FDD under each data scenario. -----	203
Figure 7.8 Season-average F1 scores for ML-CGAN-aided FDD and ML-CWGAN-aided FDD under each data scenario. -----	204
Figure 7.9 (a) Season-average F1 scores for ML-CWGAN -aided FDD and SL-CWGAN-aided FDD under each data scenario; (b) The each season's F1 score variation between the SL-CWGAN-aided performance and the baseline performance under each data scenario. -----	205
Figure 7.10 The F1 score gap between the ML-CWGAN -aided FDD performance and the ideal benchmark performance based on real fault samples only.-----	206

List of Tables

Table 2.1 Literature Search Keyword List. -----	11
Table 2.2 Fault occurrence probability for a VAV system within a year. -----	20
Table 2.3 Fault incidence rates over 15 years (Heinemeier and Cherniack, 2012). -----	20
Table 2.4 Fault models and corresponding fault characteristics for HVAC control.-----	21
Table 2.5 Fault models and corresponding fault characteristics for sensors. -----	24
Table 2.6 Fault models and corresponding fault characteristics for chillers. -----	31
Table 2.7 Fault models and corresponding fault characteristics for RTUs and split air conditioners. -----	36
Table 2.8 Fault models and corresponding fault characteristics for AHUs and VAV terminals.-----	40
Table 2.9 Fault models and corresponding fault characteristics for other HVAC faults. -----	46
Table 2.10 Existing studies for single fault impact analysis using simulations-----	55
Table 2.11 Existing studies for single fault impact analysis using lab or field experiments -----	58
Table 2.12 Existing studies for multiple faults impact analysis using simulations. -----	61
Table 2.13 Existing studies for multiple faults impact analysis using lab or field experiments -----	62
Table 3.1 Descriptions of fault characteristics (Cheung and Braun, 2015; Kim et al., 2019).-----	89
Table 3.2 Building geometry and HVAC system specifications. -----	96
Table 3.3. Fault intensity input for each fault model. -----	105
Table 3.4 Fault simulation EnergyPlus settings.-----	105
Table 4.1 System energy and thermal comfort impact rankings for each single fault under the current, 2030s and 2050s periods. -----	110

Table 4.2. System energy and thermal comfort impact rankings for each multiple faults combination under the current, 2030s and 2050s periods. -----	118
Table 4.3. Predicted season onset and end dates and season lengths (date format: day/month).-----	138
Table 5.1 Simulated HVAC fault and normal data descriptions.-----	147
Table 5.2. Real HVAC fault and normal data descriptions. -----	149
Table 5.3 Descriptions of selected features from the simulated fault dataset. -----	150
Table 5.4 Descriptions of selected features from the ASHRAE RP-1312 dataset. -----	151
Table 5.5 Descriptions of selected features from the NREL’s dataset. -----	152
Table 5.6 Sliding window specifications for the simulated dataset. -----	154
Table 5.7 Proposed FDD architecture hyperparameters for fault data experiments.--	161
Table 6.1 Baseline architecture configurations used in the general hyperparameter analysis.-----	169
Table 6.2 FDD highest test metrics scores under each season for the architectures trained on each dataset. (The values in italics are the larger scores between the two architectures, and the values in bold are the highest scores under this test season for each scenario, respectively).-----	171
Table 6.3 Average fault detection, fault diagnosis, false positive and false negative rates among ten cross-validation times for the best architecture in experiments 1–4 using the simulated dataset. -----	172
Table 6.4 Average fault detection, fault diagnosis, false positive and false negative rates among ten cross-validation times for the best architecture in experiments 5–8 using the ASHRAE RP-1312 dataset. -----	181
Table 6.5 Average fault detection, fault diagnosis, false positive and false negative rates among ten cross-validation times for the best architecture in experiments 9–11 using the NREL’s dataset. -----	182
Table 7.1 Initial experiments for assessing the influences of the data amount per fault class and the normal to fault data imbalance ratio on the baseline FDD performance.	191

Table 7.2 The seasonal-imbalanced and seasonal-balanced but insufficient data scenarios used for evaluating the baseline FDD performances.-----	192
Table 7.3 The optimal set of hyperparameters for CNN-LSTM and CNN-GRU, respectively.-----	193
Table 7.4 The GAN-aided data scenarios by aggregating the real fault samples with the synthetic fault samples (in round bracket) for evaluating GAN-aided FDD performances. -----	194
Table 7.5 CGAN and CWGAN architecture hyperparameters. -----	196
Table 7.6 The candidates for the fault data generation module and the FDD classifier module. -----	197
Table 7.7 Test F1 score per season and season-average F1 score for the baseline FDD classifiers under each data scenario. -----	198
Table 7.8 Comparisons of season-average F1 score difference between each GAN-aided FDD performance and the baseline performance under each data scenario. -----	201
Table 7.9 ML-CWGAN architecture hyperparameters.-----	202

List of Publications

Journal Papers

1. **Zhong, F.**, Calautit, J.K., and Wu, Y. (2023) ‘Fault data seasonal imbalance and insufficiency impacts on data-driven heating, ventilation and air-conditioning fault detection and diagnosis performances for energy-efficient building operations’, *Energy*, 282, 128180. The chapters 5 and 6 are based on this paper.

2. **Zhong, F.**, Calautit, J.K., and Wu, Y. (2022) ‘Assessment of HVAC system operational fault impacts and multiple faults interactions under climate change’, *Energy*, 258, 124762. The chapters 3 and section 4.2 are based on this paper.

Building Performance Simulation-Relevant Papers

3. **Zhong, F.**, Chaudhry, H.N. and Calautit, J.K. (2021) ‘Effect of roof cooling and air curtain gates on thermal and wind conditions in stadiums for hot climates’, *Energies*, 14(13), 3941. An EnergyPlus stadium model was developed to assess energy consumptions for the proposed stadium cooling configurations, coupling with computational fluid dynamics simulations.

4. **Zhong, F.** and Calautit, J.K. (2020) ‘An integrated cooling jet and air curtain system for stadiums in hot climates’, *Atmosphere*, 11(5), 546. Building performance simulations were used to assess energy consumptions for the proposed stadium cooling configurations, coupling with computational fluid dynamics simulations.

5. **Zhong, F.**, Calautit, J.K. and Hughes, B.R. (2020) ‘Analysis of the influence of cooling jets on the wind and thermal environment in football stadiums in hot climates’, *Building Services Engineering Research and Technology*, 41(5), pp. 561-585. Building performance simulations were used to assess energy consumptions for the proposed stadium cooling configurations, coupling with computational fluid dynamics simulations.

Conference Papers

1. **Zhong, F.**, Calautit, J.K., and Wu, Y. (2022) ‘Evaluation of the effect of changing climate conditions on data-driven HVAC fault detection and diagnosis performances for energy-efficient building operations’, 17th Conference on Sustainable Development of Energy, Water and Environment Systems. Paphos, Cyprus, 6-10 November 2022.

2. **Zhong, F.**, Calautit, J.K., and Wu, Y. (2021) ‘Evaluation of HVAC system operational fault impacts under climate change scenarios’, 13th International Conference on Applied Energy (ICAE 2021). Bangkok, Thailand (virtual), 29 November – 2 December 2021.
3. **Zhong, F.**, Calautit, J.K., and Wu, Y. (2021) ‘Effect of climate change on HVAC system operational fault impacts’, 16th Conference on Sustainable Development of Energy, Water and Environment Systems. Dubrovnik, Croatia, 10-15 October 2021.
4. **Zhong, F.**, Calautit, J.K., and Wu, Y. (2021) ‘Analysis of HVAC system operational fault impacts on energy demand under current and future climate scenarios’, 2021 Applied Energy Symposium: MIT A+B. Cambridge, Massachusetts, United States, 11-13 August 2021.
5. **Zhong, F.** and Calautit, J.K. (2020) ‘Analysis of the effect of roof cooling and air curtain gates on thermal and wind conditions in football stadiums for hot climates’, 4th South East European Conference on Sustainable Development of Energy, Water and Environment Systems. Sarajevo, Bosnia and Herzegovina, 28 June – 2 July 2020.

Nomenclature

Symbol	Term	Unit
R	Reliability of an HVAC component	N/A
λ	Constant failure rate of the component	h^{-1}
t	Operation hour	h
z	Relative fault bias	N/A
x	Faulty value of a model parameter	N/A
x_0	Fault free value of a model parameter	N/A
pdf	Probability density function	N/A
b	The upper limit for the fault bias	N/A
a	The lower limit for the fault bias	N/A
$T_{evap-o,f}$	The actual evaporator outlet water temperature	$^{\circ}\text{C}$
$T_{evap-o,ff}$	The biased sensor reading of the evaporator outlet water temperature	$^{\circ}\text{C}$
$T_{tower-o,f}$	The actual condenser supply water temperature	$^{\circ}\text{C}$
$T_{tower-o,ff}$	The biased sensor reading of the condenser supply water temperature	$^{\circ}\text{C}$
$T_{coil-o,f}$	The actual coil supply air temperature	$^{\circ}\text{C}$
$T_{coil-o,ff}$	The biased sensor reading of the coil supply air temperature	$^{\circ}\text{C}$
T_f	The faulty reading of the economizer temperature sensor	$^{\circ}\text{C}$
T_{ff}	The fault free value of the temperature	$^{\circ}\text{C}$
RH_f	The faulty reading of the economizer relative humidity sensor	N/A
RH_{ff}	The fault free value of the relative humidity	N/A
h_f	The faulty reading of the economizer enthalpy sensor	J
h_{ff}	The fault free value of the air enthalpy	J
$T_{stat,f}$	The faulty reading of thermostat	$^{\circ}\text{C}$
$T_{stat,ff}$	The fault free value of actual temperature	$(^{\circ}\text{C}$
$RH_{stat,f}$	The faulty reading of humidistat	N/A
$RH_{stat,ff}$	The fault free value of actual relative humidity	N/A
T_{real}	The real-time value of zone air temperature	$^{\circ}\text{C}$
W_f	The humidity ratio corresponding to $T_{real} \pm \Delta T$ and $RH_{stat,ff}$	$\text{kg}_{\text{water}}/\text{kg}_{\text{dryair}}$
$T_{evap,out}$	The water temperature at the evaporator outlet	K
$T_{cond,in}$	The water temperature at the condenser inlet	K
$\dot{Q}_{evap}$	The evaporator heat removal rate	kW
$\dot{Q}_{evap,rat}$	The rated evaporator heat removal rate	kW
$\dot{W}_{chiller,f}$	The faulty chiller power consumption	kW

$\dot{W}_{chiller}$	The fault free chiller power consumption	kW
$Q_{chiller,f}$	The faulty chiller cooling capacity	kW
$Q_{chiller,ff}$	The fault free chiller cooling capacity	kW
$P_{r,liquidline,out,f}$	The faulty pressure at the outlet of the liquid line	Pa
$P_{r,liquidline,out,ff}$	The fault free pressure at the outlet of the liquid line	Pa
$P_{r,liquidline,in}$	The pressure at the inlet of the liquid line	Pa
$P_{r,evap,in}$	The pressure at the inlet of evaporator	Pa
ΔP_f	The pressure drop across the restriction	Pa
ΔP_{ff}	The fault free pressure drop across	Pa
NC	The non-condensable gas level	N/A
M_{NC}	The mass of non-condensable gas	kg
ρ_{NC}	The density of non-condensable gas	kg/ m ³
V_{total}	The total volume of the refrigerant circuit	m ³
$P_{p,NC}$	The partial pressure of non-condensable gas	kPa
R_{NC}	The gas constant of non-condensable gas	kJ/kgK
$T_{comp,out}$	The compressor outlet temperature	K
V_{cond}	The volume of the condenser	m ³
$w_{sh,cond}$	The area ratio of the superheated vapor in the condenser	N/A
$w_{tp,cond}$	The area ratio of the two phase flow in the condenser	N/A
$\bar{y}_{r,cond,tp}$	The void fraction of the two phase flow in the condenser	N/A
V_{hg}	The volume of hot gas line	m ³
$P_{p,r}$	The partial pressure of refrigerant	kPa
P_{total}	The total pressure	kPa
$\dot{m}_{r,comp,f}$	The refrigerant mass flow rate in the compressor under the valve leakage	kg/s
$\dot{m}_{r,comp,ff}$	The refrigerant mass flow rate in the compressor without the valve leakage	kg/s
$h_{r,comp,in}$	The refrigerant enthalpy at the compressor inlet	kJ
$h_{r,suction,out}$	The refrigerant enthalpy at the outlet of the suction line	kJ
$h_{r,comp,out}$	The refrigerant enthalpy at the compressor outlet	kJ
$h_{r,comp,f}$	The refrigerant enthalpy at the compressor under the valve leakage	kg/s
$P_{r,in}$	The pressure at the compressor inlet	Pa
$P_{r,out}$	The pressure at the compressor outlet	Pa
$\dot{Q}_{cool,dx,F}$	The cooling capacity of the faulty DX cooling coil	kW
$\dot{Q}_{cool,dx}$	The cooling capacity of the fault free DX cooling coil	kW
$T_{wb,ent}$	The wet-bulb temperature at the evaporator inlet	K
$T_{db,amb}$	The dry-bulb temperature at the ambient environment	K
$\dot{Q}_{cool,dx,rat}$	The rated cooling capacity of the DX cooling coil	kW

$\dot{W}_{dx, rat}$	The rated electrical power consumption of the DX cooling coil	kW
UA_F	The overall heat transfer coefficient of the faulty DX cooling coil	kW/K
UA_{rat}	The rated overall heat transfer coefficient of the fault free DX cooling coil	kW/K
$\dot{W}_{fan}$	The condenser fan power consumption	kJ
$\dot{W}_{cool}$	The compressor power consumption	kJ
m_s	The actual mass flow rate of supply air	kg/s
f_f	The flow fraction	N/A
f_{pl}	The part load factor	N/A
δP	The design pressure difference	Pa
e_{tot}	The fan total efficiency	N/A
ρ_{air}	The density of air	kg/m ³
Q_p	The fan total power	W
$V_{min, f}$	The actual minimum outdoor air flow rate	m ³ /s
V_{min}	The normal minimum outdoor air flow rate	m ³ /s
k	The schedule factor adjusting the actual minimum outdoor air flow rate based on the stuck minimum damper position	N/A
Q_{act}	The actual outdoor/return air flow rate	m ³ /s
Q_f	The faulty outdoor/return air flow rate	m ³ /s
Q_{ff}	The fault free outdoor/return air flow rate	m ³ /s
m_c	The water flow rate through the coil	kg/s
K_v	The valve capacity	m ³ /h
x_v	The valve position	N/A
R_{coil}	The coil resistance	kg · m
r_{air}	The air side fouling factor	m ² K/W
r_{water}	The water side fouling factor	m ² K/W
A_{air}	The surface area of air side coil	m ²
A_{water}	The surface area of water side coil	m ²
UA_{air}	The heat transfer coefficient of the air side coil	W/K
UA_{water}	The heat transfer coefficient of the water side coil	W/K
R_{foul}	The overall fouling factor	K/W
ΔT_A	The temperature difference between the air outlet and the water inlet	K
ΔT_B	The temperature difference between the water outlet and the air inlet	K
Q	The coil overall heat transfer rate	W
$UA_{tower, f}$	The overall heat transfer coefficient of the fouled cooling tower	W/K

$UA_{tower,ff}$	The overall heat transfer coefficient of the fault free cooling tower	W/K
F_{UA}	The UA reduction factor due to fouling	N/A
$Q_{boiler,f}$	The nominal capacity of the fouled boiler	kW
$Q_{boiler,ff}$	The nominal capacity of the fault free boiler	kW
m_d	The design mass flow rate	kg/s
Q_{fan}	The fan power	W
Q_{shaft}	The fan shaft power	W
e_{motor}	The motor efficiency	N/A
Q_{toair}	The power entering into the air	W
f_{toair}	The ratio of motor power entering into the air	N/A
h_{out}	The specific enthalpy of the air at the fan outlet	kJ/kg
h_{in}	The specific enthalpy of the air at the fan inlet	kJ/kg
η_{total}	The fan total efficiency	N/A
η_{motor}	The motor efficiency	N/A
$\eta_{impeller}$	The impeller efficiency	N/A
$\dot{m}_f$	The faulty fixed fan air mass flow rate	kg/s
θ	The fan speed stuck percentage	N/A
S_n	The number of data samples within the data sequence	N/A
W_l	The window length	N/A
W_s	The window sliding stride	N/A

Abbreviations

IEA	International Energy Agency
AANN	Auto-associative neural network
AHU	Air handling unit
ANN	Artificial neural network
BMS	Building management system
BN	Batch normalization
BPNN	Back-propagation neural network
BPS	Building performance simulation
CART	Classification and regression tree
CAV	Constant air volume
CDF	Cumulative distribution function
CFM	Cubic feet per minute
CGAN	Conditional generative adversarial network
CK	Cohen's Kappa
CNN	Convolution neural network
COP	Coefficient of performance
CV(RMSE)	Coefficient of variation of root mean square error
CWGAN	Conditional Wasserstein generative adversarial network
DANN	Domain adversarial neural network
DaNN	Domain-adaptive neural network
DBSCAN	Density-based spatial clustering of applications with noise
DOE	Department of Energy
DRL	Deep reinforcement learning
DX	Direct expansion
EEMD	Ensemble empirical mode decomposition
EIA	Energy Information Administration
EIR	Energy input ratio
ELM	Extreme learning machine
EMD	Empirical mode decomposition
EMS	Energy management system

EWMA	Exponentially-weighted moving average
FDD	Fault detection and diagnosis
FIR	Fault impact ratio
FRP	Flexible research platform
FT	Fine-tuning
GA	Genetic algorithm
GADG	Generative adversarial domain generalization
GAN	Generative adversarial network
GCN	Graph convolutional networks
GESD	Generalised extreme studentised deviate
GP	Gradient penalty
GRNN	General regression neural-network
GRU	Gated recurrent unit
HVAC	Heating, ventilation and air-conditioning
IDF	Input data file
IPCC	Intergovernmental Panel on Climate Change
IQR	Interquartile range
JAA	Joint angle analysis
KNN	K-nearest neighbors
LMTD	Log mean temperature difference
LSTM	Long short-term memory
MA	Mixed air
ML-CWGAN	Multi-label conditional Wasserstein generative adversarial network
MLR	Multiple linear regression
mRMR	Max-relevance and min-redundancy
NIST	National Institute of Standard and Technology
NMBE	Normalised mean bias error
NREL	National Renewable Energy Laboratory
OA	Outdoor air
PCA	Principal component analysis
PDF	Probability density function
PPD	Predicted percentage dissatisfied

PUDA	Partial unsupervised domain adaption
RA	Return air
RBF	Radial basis function
ReLU	Rectified linear activation function
RF	Random forest
RCP	Representative concentration pathways
RH	Relative humidity
RNN	Recurrent neural network
RTU	Roof top unit
SA	Supply air
SG	Satizky-Golay
SGD	Stochastic gradient descent
SL-CWGAN	Single-label conditional Wasserstein generative adversarial network
SMOTE	Synthetic minority over-sampling technique
SRES	Special report on emission scenarios
SVM	Support vector machine
SVR	Support vector regression
TARP	Thermal analysis research program
TCN	Temporal convolutional network
TL	Transfer learning
t-SNE	t-stochastic neighbour embedding
VAE	Variational auto-encoder
VAV	Variable air volume
VRF	Variable refrigerant flow
WD	Wavelet de-noising
WGAN	Wasserstein generative adversarial network

Chapter 1 Introduction

1.1 Research Background

Based on the World Energy Outlook 2022 (International Energy Agency (IEA), 2022), the building sector accounted for 30.1% of global final energy consumption (Fig. 1.1). Regarding the building energy end usages, heating, ventilation and air-conditioning (HVAC) systems are responsible for 40-50% of the total building energy usage. According to building energy end usage surveys implemented by United States Energy Information Administration (EIA), 44.2% of the total energy usage for commercial buildings (EIA, 2016), and 51.3% of the total energy usage for residential buildings (EIA, 2018) were attributed by HVAC systems (Fig. 1.2). With the goal of achieving net-zero carbon emissions by the middle of the century set in the 26th United Nations Climate Change Conference of the Parties (COP26) (United Nations Framework Convention on Climate Change, 2021), accelerating the energy transition within the building sector becomes a priority. Thus, maintaining energy-efficient HVAC operations needs to be highlighted considering the large portion of HVAC energy consumption in buildings.

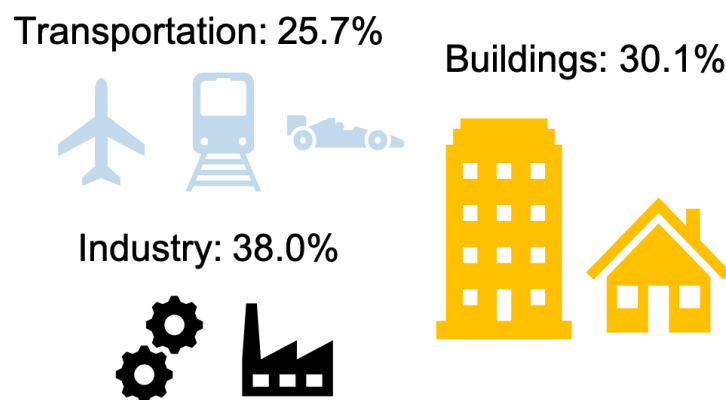


Figure 1.1 Global final energy consumption by sector in 2021 (IEA, 2022).

One of the major factors hindering HVAC systems from running under their desired operating conditions is the occurrences of faults. As stated in a survey covering more than 13000 air conditioning units in California (Downey and Proctor, 2002), faults were found in 65% of the residential air conditioning units and 71% of the commercial AC units. Typical HVAC faults contain control faults, sensor faults, component stuck, fouling and performance degradation faults (Zhang and Hong, 2017), as exemplified in Figure 1.3. The occurrences of some faults,

such as supply air (SA) temperature setpoint fault, have been reported over 50% of the monitoring time (Crowe et al., 2022). These operational faults can result in additional energy waste, occupant thermal discomfort, high system operating and maintenance costs, and low indoor air quality (Chen et al., 2023). The operational faults can lead to additional 20% or more energy waste for HVAC systems (Wu and Sun, 2011). Besides, the difference in thermally comfortable time under fault free operating condition and some faults, such as, the heating coil valve stuck at a closed position could be up to 70% (Rosato et al., 2021). As a result, the health wellbeing of occupants and work productivity of employees may also be degraded.

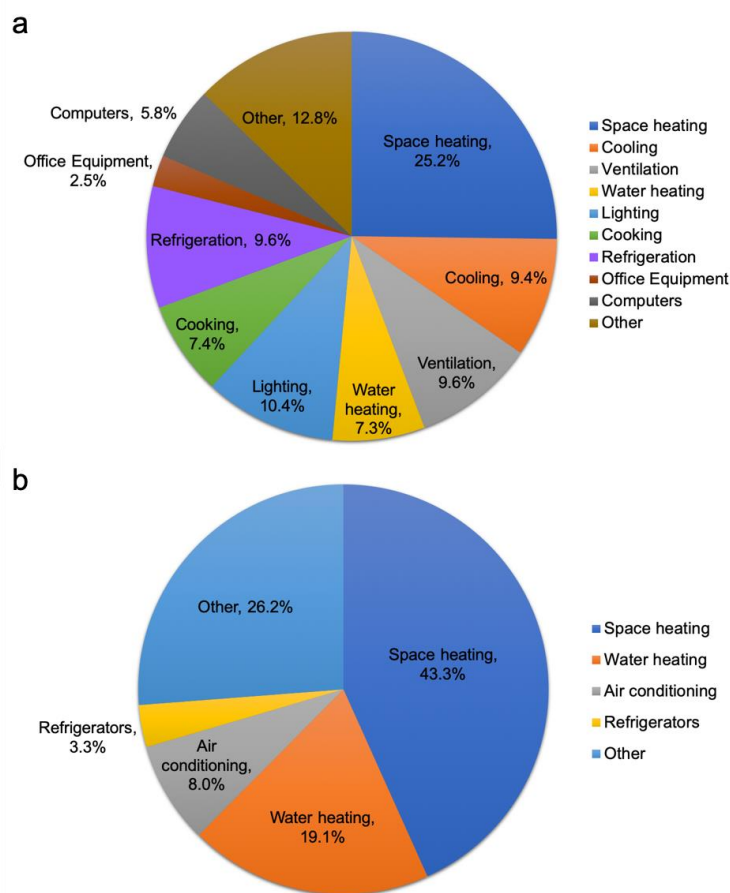


Figure 1.2 Energy end usages for (a) commercial (EIA, 2016) and (b) residential buildings (EIA, 2018) according to EIA's surveys.

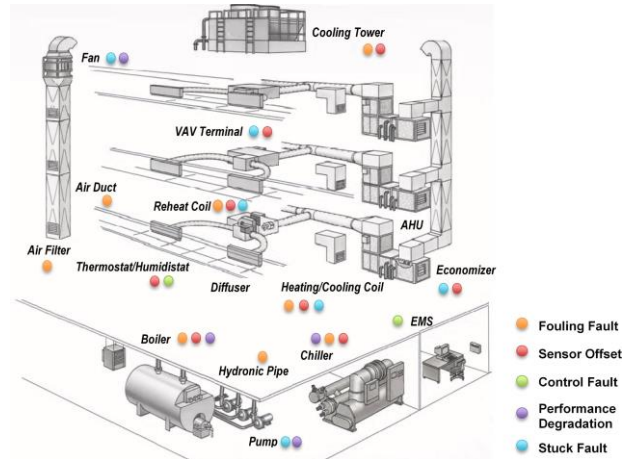


Figure 1.3 Typical HVAC faults in a variable air volume (VAV) system (Zhang and Hong, 2017).

A potential solution to occurrences of various faults in HVAC systems is to deploy automated fault detection and diagnosis (FDD) in building management systems (BMSs), as FDD tools have been indicated to save HVAC energy consumption by 10–40% (Schein et al., 2006). FDD is an application aiming to detect fault symptoms inside the industrial systems and diagnose the causes for these symptoms (Katipamula and Brambley, 2005). FDD consists of two components, namely, fault detection and fault diagnosis (Zhao et al., 2019). Fault detection is the process of examining whether any fault(s) exists in the system or not. Fault diagnosis is the process of identifying the fault type and the location of the faults. There are several classifications of FDD methods for HVAC systems. The earliest classification proposed by Katipamula and Brambley (2005) groups FDD methods into three categories: quantitative model-based, qualitative model-based and process history-based methods, as shown in Figure 1.4. However, this classification can be misleading as some methods are difficult to be assigned to one of the categories (Venkatasubramanian et al., 2003). For instance, some empirical physical models also require system data for model tuning (i.e., grey-box models), and thus they can also be grouped as process history-based methods.

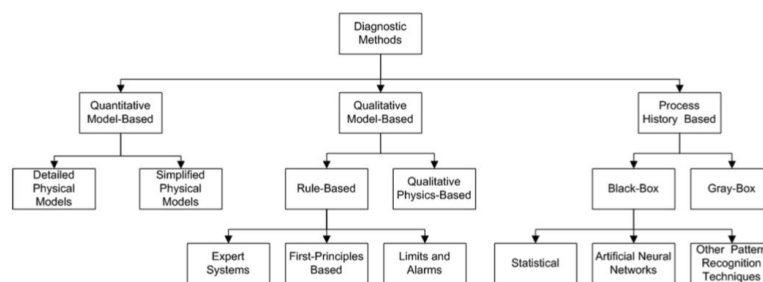


Figure 1.4 FDD classification scheme proposed by Katipamula and Brambley (2015).

Recently, another mainstream classification scheme has emerged, which is to divide the FDD methods into two categories: data-driven and knowledge-based methods (Zhao et al., 2019); or even into three categories: model-based, data-driven and knowledge-based methods (Mirnaghi and Haghighat, 2020). These classification schemes show more straightforward sense of each category, and thus bring up the focus of this thesis, i.e., data-driven FDD methods.

With the prevalence of BMSs in modern buildings, an increasing amount of data for HVAC systems are monitored in real-time. Due to the easiness of data acquisition from BMSs and the advancement of data science techniques, data-driven FDD methods for HVAC systems have been extensively explored in the recent decade. There are several benefits that data-driven FDD methods can offer to the building facility management (Chen et al., 2022): (1) They are able to automatically capture the underlying behaviours of faults through relevant features from HVAC system operational data; (2) They are less complex than model-based and knowledge-based methods as they don't require expert domain knowledge or physical models. However, there are also some weaknesses when implementing data-driven FDD methods: (1) They heavily rely on the amount of high-quality data to provide high detection and diagnosis accuracies (Fan et al., 2021a); (2) The data acquisition and fault labeling processes are laborious and time-consuming (Fan et al., 2021a); (3) The underlying FDD mechanisms of data-driven methods, particularly deep neural networks, are difficult to interpret (Chen et al., 2022).

1.2 Fault Detection and Diagnosis Challenges

Considering the weaknesses of data-driven FDD methods mentioned in section 1.1, several challenges were posed in data-driven FDD development. These challenges can be classified into two categories: data-side challenges and method side-challenges.

1.2.1 Data-side Challenges

HVAC operational data is a valuable asset for developing data-driven FDD. Although a variety of FDD methods have already demonstrated their robustness in detecting and diagnosing faults in simulated and experimental datasets (Fig. 1.5a), the capability of these methods dealing with real world fault scenarios remains unclear. One of the main reasons for this is the shortage of fault data from real-world HVAC systems (Fig. 1.5c). Chen et al.

(2022) indicated that the lower limit of FDD training data is hundreds of data, and more than tens of thousands of data are required by some methods.

In addition, the amount of fault operational data available in a HVAC system is typically lower than the amount of normal operational data due to the relatively low occurrences of faults in the system. The fault data insufficiency leads to imbalanced distribution between fault and normal data (known as inter-class imbalance from the machine learning perspective) (Fig. 1.5b), which imposes the biases towards the majority class, i.e., the normal class, on data-driven FDD. It should be noted that class imbalance refers to a phenomenon in a dataset where the ratio of the data amount of a class to that of other class(es) is highly uneven. As a result, the fault data samples are inclined to be classified as normal samples, which degrades FDD performances. Fan et al. (2021a) assessed the influences of the fault data amount and normal-fault imbalanced ratio on the fault diagnosis accuracy, and revealed that the fault diagnosis performance reduced with the increase in the imbalanced ratio and the reduction percentage was more significant when the amount of fault data sample was low than 60.

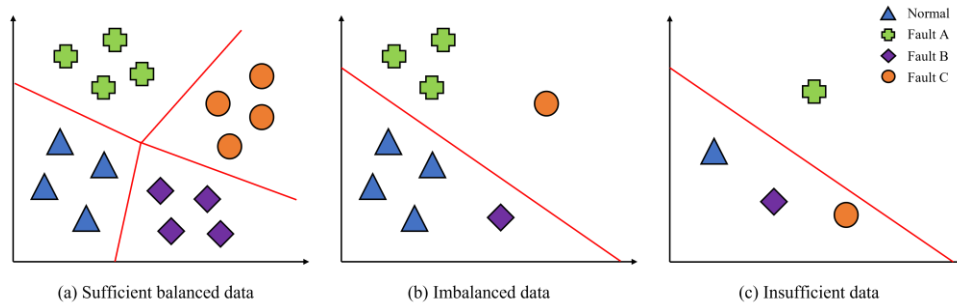


Figure 1.5 An illustration of FDD classification under (a) sufficiently balanced data; (b) imbalanced data; (c) insufficient data (the red line(s) in each case represent the separation boundaries for different data classes predicted by the FDD model).

Apart from these two obstacles to maintaining good FDD performances, dynamic HVAC system operating conditions also increase the variability in fault data patterns. Modern HVAC systems are designed to adapt to dynamic outdoor and indoor environment. The outdoor environment conditions, such as outdoor temperature, vary with the season transitions and global climate change. Therefore, HVAC systems operate at different working conditions in different climatic conditions, which may lead to fault data intra-class imbalance under

different climatic conditions and impose additional complexity on identifying the data pattern of each fault.

1.2.2 Method-side Challenges

As for challenges from the method perspective, the generality of data-driven FDD models to different HVAC systems is the major issue hindering the real-world applications of data-driven FDD. Most existing studies proposed FDD methods targeting a specific HVAC system or equipment due to the fault data availability of different HVAC systems or equipment, and the complexity and dissimilarity of different HVAC components. In addition, the inconsistency in building data and semantic representations induced difficulty for applying data-driven FDD, trained on fault data of a system/building, to another system/building.

Besides, the fact of simultaneous-occurred faults is a common situation in HVAC systems. Diagnosing multiple simultaneous faults requires FDD methods to accurately capture the data patterns of each fault. A survey conducted by Chen et al. (2022) indicated that only 15% of proposed data-driven methods incorporated the capability to identify multiple simultaneous faults. Apart from these, the interpretability of FDD results is also a major issue affect the commercialization of FDD products. The building facility engineers usually suffer from interpreting FDD processes and making decisions based on FDD results.

1.3 Motivation and Research Questions

Global climate change is an inevitable fact and the increase of global surface temperature from 1850-1900 has reached 1.09 °C in 2011–2020 (IPCC, 2021). Many studies have indicated that HVAC system performances vary with changing climate conditions. Bell et al. (2022) showed that the peak cooling demand and the unmet cooling time could increase by 24–35% and 41–189% for 2050 through building performance simulations (BPSs) of a typical office building under four Australian climates. Kharseh et al. (2014) found that the annual total HVAC energy usages intended to increase in hot climates and decrease in cold climates along with the changing climate. As a result, climate change may lead to HVAC fault impact variations across climate periods. Apart from fault impact variations caused by climate change, annually seasonal weather condition variations may also have an effect on fault impact patterns. Thus, this leads to the first research question of this thesis: how do changing climatic conditions affect HVAC fault impact patterns?

Due to the potential fault impact pattern variations caused by changing climatic conditions, the variability of fault data patterns among different climatic conditions may be increased. However, the coverage of fault data under different seasons is usually limited in experiment and real fault datasets, which leads to a fault data intra-class imbalance among different climatic conditions. The influence of the fault data intra-class imbalance on data-driven FDD performance remains unknown, which brings up the second research question: how much is the effect of the fault data intra-class imbalance among different climatic conditions on data-driven FDD performance?

Recently, with the emergence of various data augmentation techniques in the machine learning field, especially in computer vision (Shorten and Khoshgoftaar, 2019), incorporating synthetic data into the training set has enhanced the performances of machine learning tasks. In the field of HVAC FDD, generative modeling techniques, such as generative adversarial network (GAN) (Goodfellow, et al., 2014), have been adopted to resolve the fault and normal data imbalance issue (fault data inter-class imbalance). Therefore, this raises the third research question: can generative modeling techniques offer solutions to the fault data intra-class imbalance among different climatic conditions?

1.4 Aims and Objectives

The aim of this thesis is to evaluate the impacts of single and multiple faults under different climatic conditions, and also assess data-driven HVAC FDD performances under different fault data intra-class seasonal imbalanced scenarios. In addition, this thesis aims to propose a solution using generative modeling techniques to address the fault data intra-class imbalance challenge, and also incorporates the generative modeling solution into a data-driven FDD framework. The detailed objectives of this thesis are listed as below:

1. To model HVAC operational faults using Python scripts and a widely-used BPS tool, EnergyPlus.
2. To assess the impacts of single and multiple faults under different seasons of the current climate, and under different climate periods, in terms of system total and end energy usages, and occupant thermal comfort levels, through fault simulations using EnergyPlus, and to discover the correlations between fault impacts and climate periods/seasons.

3. To construct a fault dataset covering common HVAC faults under different seasons based on simulated data and to gather real fault public datasets.
4. To develop a baseline data-driven FDD classifier and to evaluate FDD performances under different fault data scenarios considering the fault data intra-class imbalance among different seasons.
5. To develop a generative modeling architecture to produce synthetic fault data for the season(s) when the fault data is absent or insufficient.
6. To integrate the generative modeling architecture with the baseline FDD classifier and to assess the effect of the entire framework on FDD performances under fault data intra-class imbalanced scenarios.

1.5 Thesis Structure

The entire thesis is organized in eight chapters and the links among main research components are demonstrated in Figure 1.6. Figure 1.6a explains the steps for achieving the objectives 1 and 2 and Figure 1.6b illustrates the workflow for achieving objectives 3 to 6, as mentioned in section 1.4. The purpose of each chapter is summarized as below:

Chapter 1 introduces the background of this research, the current challenges, the motivation and research questions, and the aims and objectives.

Chapter 2 reviews three sub-areas under data-driven HVAC FDD, i.e., fault modeling, fault impact analysis and data-driven FDD, and also provides a summary of the identified research gaps.

Chapter 3 explains the method for modeling the faults and simulating fault impacts under climatic conditions of different seasons and climate periods, which not only is useful in identifying correlations between fault impacts and climatic conditions in Chapter 4, but also formulates the foundation for the development of an FDD classifier in Chapter 5.

Chapter 4 presents fault impact analysis results across different climate periods and discusses fault impact seasonal patterns and their variations driven by climate change. The findings from this chapter stimulate the need for evaluating the developed FDD classifier's generalization ability under different seasons in Chapter 6.

Chapter 5 describes the processes for developing a data-driven FDD model using convolutional and recurrent neural networks, and the generalization ability of this developed FDD model is evaluated in Chapter 6.

Chapter 6 explains FDD evaluation experiments design based on the simulated fault dataset and two real fault datasets, and presents FDD evaluation experiments' results and key findings. Motivated by these findings, a fault data generation module is developed and the enhancement on the baseline FDD classifier's performance driven by this module is discussed in Chapter 7.

Chapter 7 integrates a fault data generation module using multi-label conditional Wasserstein generative adversarial network with the baseline FDD classifier in an entire framework and evaluates the GAN-aided FDD performances under fault data seasonal imbalanced scenarios for the proposed framework. In addition, the FDD performances led by the proposed framework are compared with other GAN-based FDD performances.

Chapter 8 summarizes the work from the above chapters and provides some recommendations for future research directions.

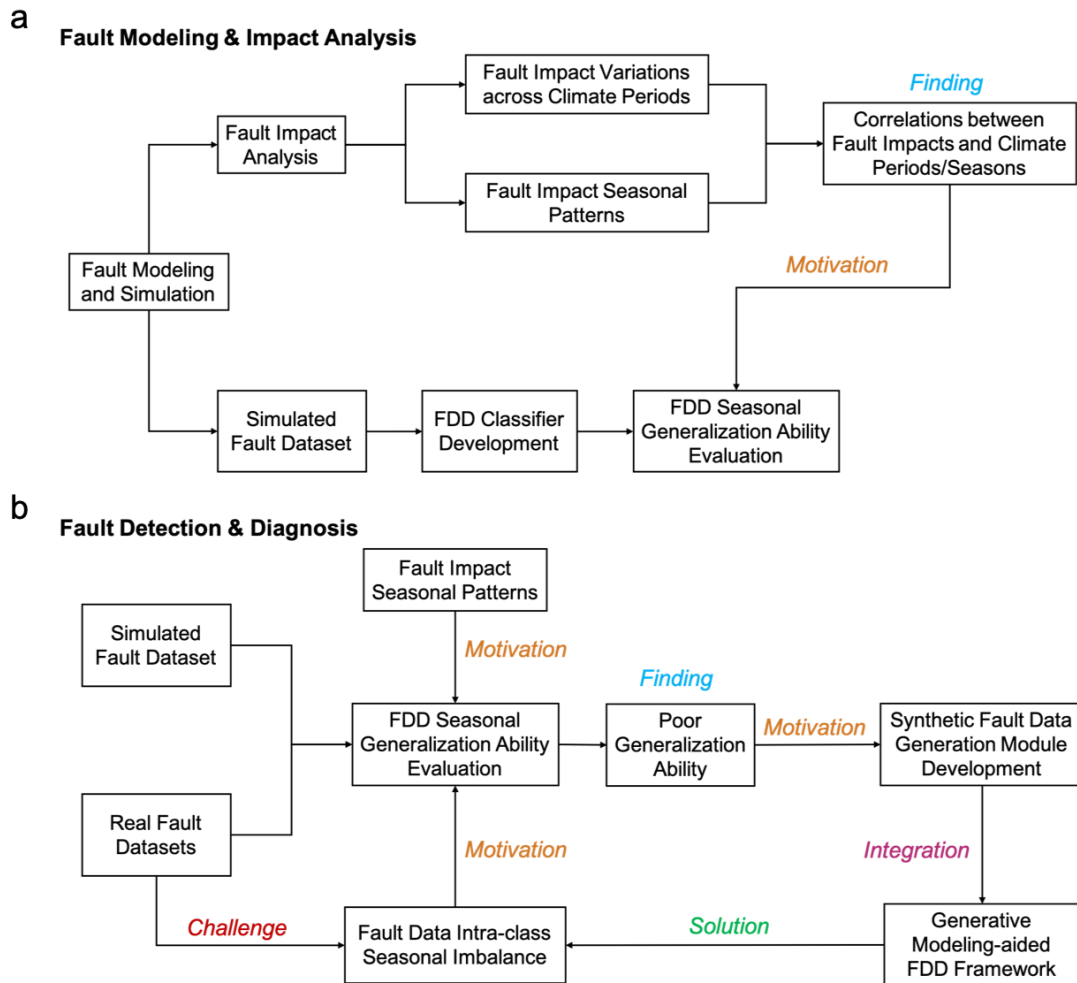


Figure 1.6 Research workflow diagram of the present thesis: (a) fault modeling (Chapter 3) and impact analysis (Chapter 4); (b) fault detection and diagnosis (Chapter 5-7).

Chapter 2 Literature Review

2.1 Introduction

This chapter separately reviews three main topics of HVAC operational faults from the existing literature: fault modeling, fault impact analysis and FDD. Fault modeling is a useful technique to create a fault model library to capture physical mechanisms of HVAC operational faults, in order to support the fault impact analysis and development of FDD frameworks. Fault impact analysis assesses the effects of single faults or multiple simultaneous faults cases on HVAC system energy consumptions, occupant thermal comfort and operating cost, etc. FDD aims to detect faults from HVAC system operating conditions and identify the fault type(s). The review for each aspect is presented in section 2.2, 2.3 and 2.4, respectively. Furthermore, a summary of the identified research gaps based on this chapter is shown in section 2.5.

2.1.1 Literature Search Method

The literature search was carried out thoroughly from the Scopus database using combinations of synonyms of keywords from categories including topic, target and method, as listed in Table 2.1. The fault modeling studies were searched by using keyword combinations of ‘fault modeling’ and a series of target systems, such as VAV system, and subsystems, such as air handling unit (AHU). The fault impact analysis literatures were gathered using the integration of ‘fault impact analysis’ and its synonyms ‘fault impact assessment’ and ‘fault impact evaluation’ with the target systems and subsystems. The FDD studies were collected using the combinations of topics ‘fault detection’ and ‘fault diagnosis’ with the target systems and subsystems, as well as the method, such as ‘machine learning’. Moreover, the ‘snow-ball’ method (Chen et al., 2022) was used to check references in the above reviewed papers to ensure that the literature search was comprehensively performed. It should be noted that all the gathered papers were manually scanned to only include papers related to the topics of fault modeling, fault impact analysis and FDD studies within the realm of HVAC systems published after 2000. For the fault modeling studies, 22 studies were retrieved from the literature search and all of them were selected after the manual screening. For the fault impact studies, 50 studies were gathered from the literature search and 46 of them were selected. For the FDD studies, 112 studies were obtained from the literature search and 101 of them were selected. It should be noted that the studies filtered out during the manual screening were due

to the repeated contributions of similar contents by same authors in journals and conference papers.

Table 2.1 Literature Search Keyword List.

Fault modeling	
<i>Topic</i>	Fault modeling
<i>Target</i>	HVAC, building systems, VAV, variable refrigerant flow (VRF), fan coil, AHU, chiller, heat pump, roof top unit (RTU), vapor compression system
Fault impact analysis	
<i>Topic</i>	Fault impact analysis/assessment/evaluation
<i>Target</i>	HVAC, building systems, VAV, VRF, fan coil, AHU, chiller, heat pump, RTU, vapor compression system
FDD	
<i>Topic</i>	Fault detection, fault diagnosis
<i>Target</i>	HVAC, building systems, VAV, VRF, fan coil, AHU, chiller, heat pump, RTU, vapor compression system
<i>Method</i>	Data-driven, machine learning, deep learning, neural network

2.2 Fault Modeling Review

HVAC Operational data is crucial for the FDD training and evaluation, and fault impact analysis. As for data-driven FDD methods using machine learning and deep learning algorithms, it is essential that a sufficient amount of high-quality fault data can be provided for training FDD models. However, in reality, it is common that there is a lack of large high-quality fault datasets due to the high cost to collect fault data from lab experiments and field tests, which may require a long-term monitoring of systems and sufficient types of sensors installed in buildings to capture HVAC system fault behaviours (Frank et al., 2016; Kim et al., 2019). Fault modeling can be an alternative means to generate fault datasets as faults are simulated based on calibrated building models. Compared with collecting fault data from field tests or lab experiments, fault modeling can offer several strengths (Kim et al., 2019; Frank et al., 2016): (1) the ground truth of faults is known and hence reliable; (2) the data quality issues such as missing time series data can be avoided; (3) the cost of generating fault datasets is lower; (4) the fault types from the simulation data are richer than those available from building monitoring data; (5) the impact of a given fault can be assessed with different building types and outdoor and indoor environmental conditions; (6) the multi-year fault data can be obtained under a short period of time. But there are still several limitations in using this method: (1) errors between the simulated data and actual measured data still exist

although the validity of models is verified; (2) fault simulations producing significant amount of data require high computational power.

This section presents the outcomes based on a comprehensive review on fault modeling studies. The different categories of fault modeling approaches are explained in subsection 2.2.1. The building simulation platforms that can be used for fault modeling are introduced in subsection 2.2.2. The fault characteristics incorporated in fault model development are described in subsection 2.2.3. The detailed approaches for modeling common HVAC faults are explained in subsection 2.2.4.

2.2.1 Fault Modeling Approach Category

Due to the complexity of HVAC systems and the sophisticated inherent connections among the HVAC subsystems and components, it is challenging to reflect the characteristics of HVAC operational faults through fault modeling. Haves (1997) suggested two ways of fault modeling: (1) altering the values of existing fault-associated parameters in a fault free HVAC system or building model; (2) explicitly introducing a new parameter describing the fault intensity into a fault free model on the current modeling platform. In the recent studies (Cheung and Braun, 2015; Kim et al., 2019), the fault modeling techniques can be further grouped into three categories: physical, semi-empirical and empirical modeling.

The physical approach is implemented by using building system physics (first principles) to model fault impacts and can be further grouped into detailed and simplified methods (Kim et al., 2019). The detailed physical modeling method holistically develops an HVAC system containing essential HVAC components, and simulates the faults by changing the values of fault-associated parameters in the corresponding components or at the system level. Thus, sufficient detailed simulation outputs of the HVAC system or components can be obtained, but high computational costs are required by this method. In contrast, the simplified method directly uses the physical parameters which have already been defined in BPS programs. However, the impacts of fault intensities on some relevant physical parameters may vary with different system operating conditions and are therefore difficult to quantify by a simplified physical model. For instance, the cooling coil performance degradation caused by the refrigerant undercharge may vary with evaporator inlet temperature (Kim et al., 2019) and the impact of this fault on cooling coil performance may not be included in a simplified physical model. In this circumstance, semi-empirical or empirical modeling methods are more

appropriate to be used since they are typically derived from measurements of real HVAC systems or components.

Empirical modeling is used to model faults that have complex influences on HVAC systems or components, such as condenser fouling in chillers. It directly applies the empirical correlations that capture the behaviors of faults to the HVAC systems or components, and thus requires training datasets to determine the empirical coefficients. Cheung and Braun (2015) employed this method to model faults for a vapor compression system. Empirical equations were derived for fault impact ratios which represent energy input ratio, cooling capacity and evaporator heat transfer conductance based on the values of these parameters at fault-free conditions. The coefficients in the empirical equations were estimated by linear regression from the training data generated from a detailed physical fault simulation conducted in their previous studies (Cheung and Braun, 2013a; 2013b).

Semi-empirical modeling is a combination of the physical and empirical method. It is typically performed by simply adding or modifying an empirical correlation for a physical operational parameter of an HVAC system, which further changes the fault impacts on the HVAC system based on system physics. Duct fouling is such an example that can be developed using semi-empirical method. Pressure drop and airflow rate are modified empirically based on correlations derived from training data generated by BPSs, and the associated fault impacts on the HVAC system are simulated on the basis of system physics equations embedded in the simulation program.

2.2.2 Fault Modeling Platforms

The emerging fault modeling approaches and the need for whole-building level fault impact analysis lead to the use of BPS programs to facilitate the fault modeling and simulation. Haves (1997) suggested the implementation of fault modeling can be achieved by either modifying the existing fault-associated input parameters in BPS programs or developing new fault-associated input parameters based on the algorithms of BPS programs. However, only a few of these programs can support HVAC fault modeling, including HVACSIM+, DOE-2, TRNSYS, EnergyPlus and Modelica buildings library.

HVACSIM+ is a hierarchical modular building system simulation program developed by National Institute of Standard and Technology (NIST) (Kelly et al., 1984). Bushby et al.

(2001) modeled a series of faults for a single-duct VAV air system including sensor offsets, air dampers stuck or leaking, and coil valve stuck or leaking in HVACSIM+. Apart from these, some faults occurring in typical AHUs were modeled in HVACSIM+ and validated by the field measurements (Li and Wen, 2010, Li et al., 2010). The fault modeling workflow used by Li and Wen (2010) is illustrated in Figure 2.1. Moreover, Pourarian et al. (2017) proposed a validated fan coil unit model in HVACSIM+ to facilitate the modeling of fan coil unit faults.

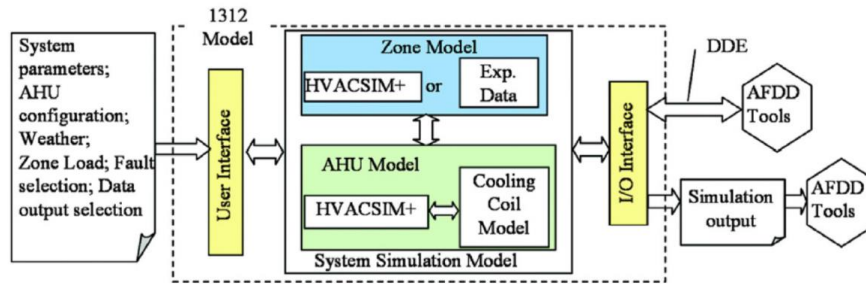


Figure 2.1 Fault modeling structures on HVACSIM+ (Li and Wen, 2010).

DOE-2 is a BPS tool developed by the Department of Energy (DOE) in the U.S. The fault modeling studies based on this tool are limited in the field. Katipamula et al. (1999) developed the fault models of OA temperature sensor offset and OA damper stuck in DOE-2 to assist in testing FDD algorithms.

TRNSYS is a flexible BPS program that provides the users with the flexibility of creating their own fault models based on its simulation engine. Otto et al. (2012) created fault models, including thermostat offset, OA damper leaking, heating coil valve stuck, etc. using TRNSYS and also introduced the fault occurrence probability into fault models. Khire and Trcka (2013) developed models for faults associated with OA dampers, fans, direct expansion (DX) coils, etc. on the TRNSYS platform.

EnergyPlus is one of the most widely used BPS programs which provides the users with powerful capability of fault modeling. Zhang and Hong (2017) introduced three methods for fault modeling in EnergyPlus: (1) directly modify the fault associated parameters or curves in EnergyPlus input data file (IDF); (2) overwrite the specified fault-related parameters using the energy management system (EMS); (3) create native fault objects by programming in EnergyPlus source codes by developers and therefore users can directly input required fault features into the IDF file. The authors indicated the third approach using the native fault

objects was more advantageous than the other two approaches since it is not only more user friendly for practitioners and researchers, but also highly capable and flexible for developers to create fault objects. Currently, HVAC operation faults containing air economizer sensor faults, thermostat/humidistat offset, heating/cooling coil fouling, dirty air filter, supply chilled water temperature sensor offset, condenser supply water temperature sensor offset, cooling tower fouling, coil SA temperature sensor offset, boiler fouling, water-cooled chiller fouling, evaporative cooler fouling are covered in the native fault objects of EnergyPlus (2020).

Modelica is an object-oriented modeling language for sophisticated dynamic systems. The Modelica buildings library (Wetter et al. 2014) is capable of modeling HVAC systems and components. Bonvini et al. (2014) modeled the pipe leaking fault using Modelica based on physical equations. Padilla and Choinière (2015) implemented sensor fault models in a VAV system on the basis of the Modelica platform. Sterling et al. (2014) created AHU fault models by introducing the fault-associated parameters into the fault free models to produce fault datasets for FDD performance evaluation.

2.2.3 Fault Characteristics

The characteristics of faults need to be carefully reviewed at the fault model design phase in order to comprehensively reflect the fault impacts on HVAC systems. The fault characteristics that have been investigated in the literature include the following: fault intensity, fault evolution, fault occurrence rate.

Fault Intensity

Fault intensity, also known as fault level or fault severity, is used to measure the fault numerically with a fault intensity range in the fault model (Cheung and Braun, 2015). The fault intensity range can be either binary or continuous. The details of fault intensity for each fault model will be discussed in Subsection 2.2.4.

Fault Evolution

Fault evolution refers to the case that the fault intensity of some faults may vary with time. For example, the economizer sensor bias may shift over time, and thus the fault evolution should be included in such a fault model to represent the actual fault behavior. The evolution of the fault intensity can be linear or nonlinear and may even be coupled with particular system operations. Kim et al. (2019) performed fault evolution for some faults based on an

assumption that the evolution is in a linear relationship with time, as illustrated in Figure 2.2. The nonlinear evolution of faults, containing discontinuous evolution which may occur when particular component is in operation, is still lacking in the field. But this kind of evolution is suggested to be included in fault models, for example, through manipulating the codes in the EMS module of EnergyPlus.

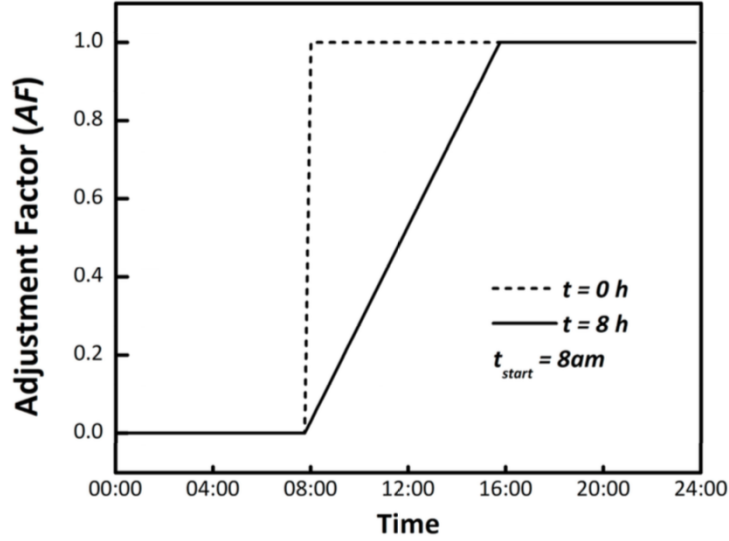


Figure 2.2 An example of fault evolution from Kim et al. (2019).

Fault Occurrence Rate

Different types of faults occur at different probabilities, and even the same fault with different fault intensities has different occurrence frequency. As a result, the fault occurrence rate has to be taken into account to realistically model the fault. To investigate the fault occurrence rate, knowledge of reliability theory was transferred from software engineering to the fault studies of HVAC systems. Myrefelt (2004) derived reliability functions for the HVAC system according to a large amount of fault occurrence data in HVAC systems. The reliability functions are proposed on the basis of an assumption that the faults are not affected by preventive maintenance. The reliability of an HVAC component can be determined by:

$$R(t) = e^{-\lambda t} \quad (2.1)$$

where λ is constant failure rate of the component (h^{-1}); t is the operation hour (h).

The reliability for the whole system can be calculated using the same formula, but the failure rate of the system needs to be determined from the mean time among failures of the system:

$$\lambda_{sys} = \frac{1}{\overline{\Delta t_{F,sys}}} = \frac{1}{\sum_{i=1}^n \lambda_i} = \frac{1}{\sum_{i=1}^n \frac{1}{\overline{\Delta t_{F,i}}}} \quad (2.2)$$

where $\overline{\Delta t_{F,sys}}$ is the mean time between system failures (o.h.); $\overline{\Delta t_{F,i}}$ is the mean time between failures of subsystem i (h); λ_{sys} is the failure rate of the system (number/h); λ_i is the failure rate of the subsystem i (number/h).

Note that Eqn. (2.2) is on the ground of the random fault occurrences and the independency among the faults. Hence, the reliability of the whole system is given by:

$$R_{sys}(t) = e^{-\lambda_{sys}t} \quad (2.3)$$

Otto et al. (2012) used a probability density function (PDF) to define fault occurrence distribution, as depicted in Figure 2.3. Based on this approach, the statistical distribution models can be applied to the PDF function depending on the actual fault behaviors.

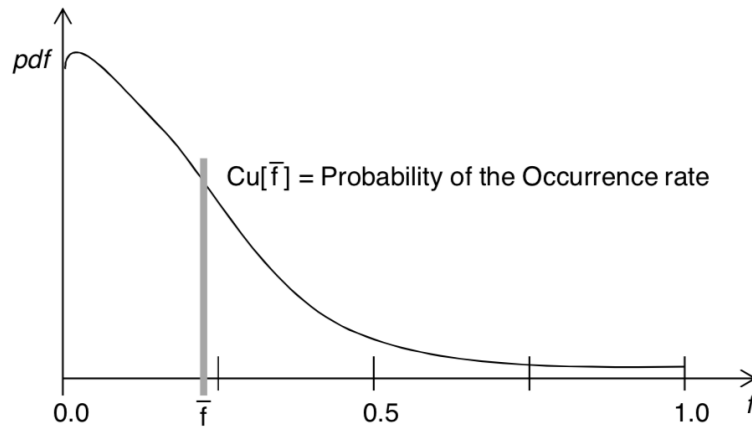


Figure 2.3 An example of fault occurrence rate distribution for an HVAC system (Otto et al., 2012).

Li and O'Neill (2019) further improved this approach and classified faults into four groups according to their occurrence distributions: Fault-by-Lacking, Fault-by-Excessing, Fault-by-Deviating and Fault-by-Uniform. To propose fault occurrence distribution equations for each type of faults, the relative fault bias is introduced in their work:

$$z = \frac{x - x_0}{x_0} \quad (2.4)$$

where x_0 is the fault free value of a model parameter; x is the faulty value of a model parameter; z is the relative fault bias.

Fault-by-Lacking

This type of faults describes the faults which result in the decrease of fault-associated parameters from the fault free value. Common instances of this fault type in HVAC systems contain the coefficient of performance (COP) degradation and the fan or pump efficiency degradation. The typical feature for this fault type is that the fault-associated parameter has a minimum value of zero. For instance, the efficiency of fan or pump cannot reach zero which indicates a shut-down status of the equipment. The occurrence probability distribution for this type of fault is derived from the Weibull distribution model (Bourguignon, Silva and Cordeiro, 2014; Cordeiro, Ortega and Ramires, 2015; Hassan et al., 2017) and represented as follows:

$$pdf = -3e^{z^3} z^2 \quad (2.5)$$

$$cdf = 1 - e^{z^3} \quad (2.6)$$

$$z \in [-1, 0] \quad (2.7)$$

where the PDF presents the fault occurrence probability distribution corresponding to the specified cumulative distribution function (CDF).

Fault-by-Excessing

On the contrary, the second type of faults represents the faults which lead to fault-associated parameter increasing from the fault free value. Typical example for this type is the minimum position of OA damper stuck on the maximum position. There is usually an upper limit for the fault-associated parameters of such faults, which means that the fault parameter can grow up to this upper limit while maintaining the system to be in operation. The occurrence probability equations for this type of faults are also obtained from the Weibull distribution (Bourguignon, Silva and Cordeiro, 2014; Cordeiro, Ortega and Ramires, 2015; Hassan et al., 2017):

$$pdf = 8e^{-\left(\frac{z+0.2}{0.5}\right)^2} \left(\frac{z+0.2}{0.5}\right) \quad (2.8)$$

$$cdf = 1 - e^{-\left(\frac{z+0.2}{0.5}\right)^2} \quad (2.9)$$

$$z \in [0, 1] \quad (2.10)$$

Fault-by-Deviating

The third type of faults is the faults which can cause the fault-associated parameters both increasing and decreasing from the fault free value. An example for this type of faults is the sensor reading offsets. The occurrence probability density function for this type follows the normal distribution:

$$pdf(z) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{z}{\sigma}\right)^2} \quad (2.11)$$

where σ is the standard deviation.

Fault-by-Uniform

The fourth type of faults follows a uniform occurrence distribution, which is uncommon in the HVAC system. For example, the program bug for a specific parameter in HVAC system control loops. Due to the randomness of the bug occurrence, any fault bias within the reasonable range is possible. The probability density function of such type of faults can be expressed by:

$$pdf(z) = \frac{1}{b-a} \quad (2.12)$$

where b and a are the upper and lower limit for the fault bias, respectively.

Apart from these generic fault occurrence functions, fault occurrence probability was also investigated using maintenance records or survey data. Dey et al. (2016) studied occurrence probability of faults in five types of components: OA mixing box, heating coil, cooling coil, controller and sensors with one-year maintenance records for a VAV system. The detailed probability for each fault from this study was listed in Table 2.2.

In addition, Heinemeier and Cherniack (2012) displayed fault incidence rates over 15 years in a research report, as shown in Table 2.3. It was found that the refrigerant undercharge had the

highest occurrence probability of 85%, and the air temperature sensor malfunction had the lowest incidence of 2%.

Table 2.2 Fault occurrence probability for a VAV system within a year.

Faults	Number of occurrences	Probability
Heating coil leaking	5	0.108696
Heating coil stuck	2	0.043478
Hot water temperature low	4	0.086957
Cooling coil leaking	3	0.065217
Cooling coil stuck	3	0.065217
Cooling coil fouled	3	0.065217
Chilled water temperature high	4	0.086957
Controller logic error	1	0.021739
Controller signal error	1	0.021739
SA temperature biased	4	0.086957
Mixing air (MA) temperature biased	4	0.086957
Return air (RA) temperature biased	4	0.086957
RA damper stuck	5	0.108696
Outdoor air (OA) damper stuck	3	0.065217

Table 2.3 Fault incidence rates over 15 years (Heinemeier and Cherniack, 2012).

Faults	Fault incidence
Air temperature sensor malfunction	2%
Refrigerant charge: 80% of nominal charge	85%
Compressor short cycling	30%
Refrigerant line restrictions/TXV problems	62%
Refrigerant line non-condensibles	50%
Low side HX problem including low airflow	74%
High side HX problem	48%
Not economizing when it should (high-limit setpoint 55F instead of 75F)	30%
Damper not modulating	24%
Excess outdoor air	24%

2.2.4 Fault Model Descriptions

HVAC systems consist of various subsystems and each subsystem also contains a variety of components. Each component is associated with different types of faults. Thus, the developed fault models in the existing studies are reviewed by the categories of HVAC components. Currently, seven categories of HVAC system faults have been modeled: control faults, sensor faults, AHU faults, chiller faults, RTU and split air conditioner faults, VAV box faults and

other faults. In this subsection, the available fault models under each category are described in detail with the fault symptoms and the modeling approach. The characteristics related to each fault model are summarized in Tables 2.4-2.9.

2.2.4.1 Control Faults

Control faults are faults that occur in the HVAC control and mainly associated with the heating and (or) cooling setpoints of the HVAC system. The currently available control fault models in the literature (Cheung and Braun, 2015; Kim et al., 2019) are: (1) HVAC setback fault: no overnight setback; (2) HVAC setback fault: extended morning setback; (3) HVAC setback fault: delayed evening setback; (4) Occupant manual changes of setpoint due to extreme outdoor temperatures. The descriptions for corresponding characteristics of these faults are shown in Table 2.4.

Table 2.4 Fault models and corresponding fault characteristics for HVAC control.

Fault models	Fault intensity definition	Fault intensity range	Fault evolution
HVAC setback fault: no overnight setback	Overnight setback is established or not	0 or 1 ^[1]	N ^[2]
HVAC setback fault: extended morning setpoints	Hours that the setpoint for occupied hours is extended forward in the morning	0 to 3 h	N
HVAC setback fault: delayed evening setpoints	Hours that the setpoint for occupied hours is delayed in the evening	0 to 3 h	N
Occupant manual changes of setpoint due to extreme outdoor temperatures	The critical indoor temperature occupants alter the thermostat setpoints and the degree of setpoints altered by occupants	User-defined	N/A

[1] For binary fault intensity range, 0 represents ‘off’ and 1 represents ‘on’.

[2] ‘Y’ represents the fault model has fault evolution feature; ‘N’ represents no fault evolution feature.

(1) HVAC setback fault: no overnight setback

Typically, thermostat setback is used to increase setpoints in the cooling season and decrease setpoints in the heating season for unoccupied hours, in order to reduce HVAC energy consumptions. This fault occurs when thermostat setpoint setback is not set for unoccupied

hours accidentally and thus a constant setpoint for occupied hours is used for all times, which leads to unnecessary energy consumption for maintaining the conditioned space at comfort level. This fault is modeled by altering the thermostat setpoint for unoccupied hours to a constant setpoint for occupied hours. If multiple thermostat setpoints exist during the occupied hours for a HVAC system, the thermostat setpoint immediately before switching to unoccupied hours will be selected as the setpoint for the unoccupied period (Cheung and Braun, 2015). The green dash line in Figure 2.4 presents an example of no overnight setback fault (Kim et al., 2019).

(2) HVAC setpoint setback fault: extended morning setback

This fault model assumes that a precooling or preheating schedule exists in the HVAC control, but the setpoint for this pre-operation time is mistakenly extended forward to unoccupied hours (Cheung and Braun, 2015). For instance, a precooling schedule is designed to be between 7 a.m. and 8 a.m., but the building operation team set the precooling schedule an hour earlier to 6 a.m. by mistake. As a result, this redundant one-hour precooling consumes extra energy without remarkably improving the occupant thermal comfort. This fault can be modeled by extending the initial change of setpoint in advance, and the initial change is defined as the time when precooling or preheating is switched on for HVAC systems with pre-operation settings or the time when the building opens for HVAC systems without pre-operation need. The fault intensity is defined as hours that the setpoint for occupied hours is extended forward in the morning and ranges from zero to three hours (Kim et al., 2019). It should be noted that this extension may reduce the original occupant discomfort at the beginning of occupied hours if the designed building setpoint schedule is not optimal and accordingly the extension cannot be regarded as a fault. The orange dash line in Figure 2.4 shows an example of an extended setpoint setback in the morning (Kim et al., 2019).

(3) HVAC setpoint setback fault: delayed evening setback

Similar to the fault of extended morning setback, this fault delays the last shift of thermostat setpoint from the value for occupied hours to that for unoccupied hours. This delay of the setpoint shift can be modeled by extending the setpoint change to the value of unoccupied hours to a later time (Cheung and Braun, 2015). The fault intensity is defined as hours that the setpoint for occupied hours is delayed in the evening and ranges from zero to three hours

(Kim et al., 2019). Figure 2.4 plots the delay of thermostat setpoint setback in red (Kim et al., 2019).

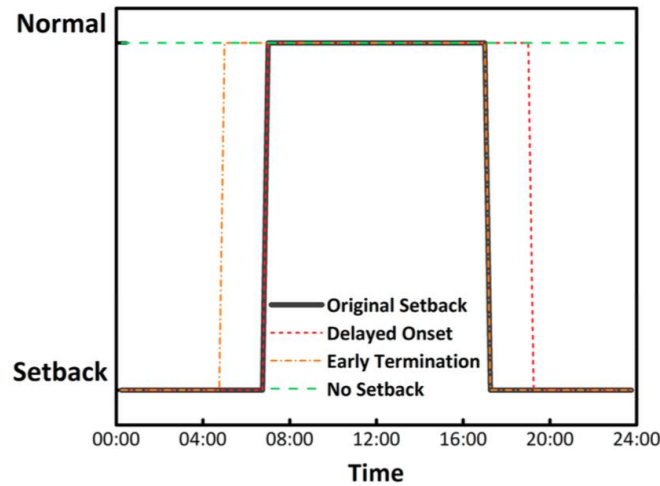


Figure 2.4 Comparison between HVAC setback faults and normal case (Kim et al., 2019).

(4) Occupant manual changes of setpoint due to extreme outdoor temperatures

During days with extreme outdoor temperatures, occupants may feel considerably hot or cold at the moment when they enter the conditioned space and require prompt cooler or warmer thermal sensation. Consequently, they may adjust the thermostat setpoint manually to be lower or higher and do not set the setpoint back to the designed values after feeling thermally comfortable. Since the modern HVAC control schedules are optimized to achieve the required conditions of the space with lowest energy usage, manual alteration of the setpoint can result in excessive energy consumption. Therefore, this manual alteration can be considered as a fault for the HVAC system (Cheung and Braun, 2015). In heating season, this fault is modeled by defining the upper limit for the outdoor temperature at which occupants will manually alter the heating setpoint and the resultant increase in the setpoint. In the cooling season, similarly, this fault is modeled by setting the lower limit for the outdoor temperature at which occupants will alter the cooling setpoint and the resultant decrease in the setpoint. However, the statistical data for the upper and lower limit of outdoor temperature and the resultant changes in the setpoints is currently not available. For this reason, the values of these fault intensity parameters are user-defined.

2.2.4.2 Sensor Faults

A wide range of sensors are installed in HVAC systems in order to achieve the monitoring and control of components inside the system. These sensors contain temperature sensors,

humidity sensors, enthalpy sensors, pressure sensors, flow rate sensors, etc. The sensor faults are caused by the aging of sensors, which leads to the drift of sensor measurements from the calibrated readings (Cheung and Braun, 2015). If the faulty sensor is not calibrated promptly, the fault may substantially affect system energy performance and occupant thermal comfort since the sensor presents the biased measurement from the actual system or environmental conditions. For some sensors such as air economizer sensors, the deviations from the true conditions may also influence other parameters such as the indoor air quality. It should be noted that the thermostat and humidistat offset are also included in this category due to the similar symptoms in the sensor component of thermostat and humidistat. The fault intensity of sensor faults is defined as the deviation of the sensor readings from the actual values of parameters that the sensor is recording. The descriptions for corresponding characteristics of these faults are shown in Table 2.5.

Table 2.5 Fault models and corresponding fault characteristics for sensors.

Fault models	Fault intensity definition	Fault intensity range	Fault evolution
Chiller supply water temperature sensor bias	Sensor offset in supply chilled water temperature	Within the range of 7°C to 20°C.	Y
Condenser supply water temperature sensor bias	Sensor offset in condenser supply water temperature	Within the user-defined range	Y
Coil SA temperature sensor bias	Sensor offset in coil SA temperature	$\pm 4^{\circ}\text{C}$	Y
Economizer OA/RA/MA temperature sensor bias	Sensor offset in OA/RA/MA temperature	$\pm 4^{\circ}\text{C}$	Y
Economizer OA/RA relative humidity sensor bias	Sensor offset in OA/RA relative humidity	$\pm 10\%$	Y
Economizer OA/RA enthalpy sensor bias	Sensor offset in OA/RA enthalpy	N/A	N/A
Thermostat offset	Thermostat offset in air temperature	$\pm 3^{\circ}\text{C}$	Y
Humidistat offset	Humidity offset in air relative humidity	N/A	Y

(1) Chiller supply water temperature sensor bias

The chiller supply water temperature sensor is used to control the temperature of supply chilled water to the AHU(s). The deviation of set supply chilled water temperature from the

true water temperature happens due to irregular calibration of temperature sensors at the chiller evaporator outlet. Since the sensor readings of supply chilled water temperature may affect multiple components along the HVAC water loop, it is inefficient to manually change the water temperature input for all the associated components. It is also invalid to modify the temperature reading record at the sensor location in the building performance simulation tools for the reason that it alters the real water temperature and results in inaccurate heat transfer and mass conservation calculations. Cheung and Braun (2015) proposed a way to model this fault by changing the temperature setpoint for this chilled water supply temperature sensor. For example, the temperature setpoint for the sensor increases by 2°C if the temperature sensor reading is 2K lower than the actual condition. In addition, the native fault object of the supply chilled water temperature sensor offset embedded in EnergyPlus (2020) simplifies the modifications of the water temperature input for all the associated components by modifying the evaporator outlet water temperature to its actual value in the native fault object. This actual water temperature at the evaporator outlet under the faulty case can be calculated by the following equation:

$$T_{evap-o,f} = T_{evap-o,ff} - \Delta T \quad (2.13)$$

where $T_{evap-o,f}$ is the actual evaporator outlet water temperature (°C); $T_{evap-o,ff}$ is the biased sensor reading of the evaporator outlet water temperature (°C); ΔT is the deviation of the sensor temperature reading from the actual temperature (°C). Note that if this fault results in a supply chilled water temperature out of the limit value range, the predefined upper or lower limit will replace the actual evaporator outlet water temperature $T_{evap-o,f}$. The supply chilled water temperature is typically within the range of 7°C to 20°C (Schein and Bushby, 2005).

(2) Condenser supply water temperature sensor bias

The condenser supply water temperature sensor is used to control the temperature of the supply cooling water from the cooling tower to the condenser. Similar to the supply chilled water temperature sensor bias, the condenser supply water temperature sensor bias occurs when the temperature reading of supply water from the cooling tower deviates from the actual water temperature at the condenser inlet. As a result, the fault may influence the actual operating performance of the condenser and the cooling tower. EnergyPlus (2020) also

included this fault in its native fault objects feature. The sensor offset in the condenser supply water temperature is modeled as follows:

$$T_{tower-o,f} = T_{tower-o,ff} - \Delta T \quad (2.14)$$

where $T_{tower-o,f}$ is the actual condenser supply water temperature (°C); $T_{tower-o,ff}$ is the biased sensor reading of the condenser supply water temperature (°C); ΔT is the deviation of the sensor temperature reading from the actual temperature (°C). Note that if this fault results in a condenser supply water temperature out of the limit value range, the predefined upper or lower limit will replace the actual condenser supply water temperature.

(3) Coil SA temperature sensor bias

The SA temperature sensor bias for the AHU coils can be either positive or negative. A positive bias will lead to a lower temperature of the SA from the coils than the SA temperature setpoint, which can reduce the required airflow rate in the cooling season and thus reduce the AHU fan energy consumption but may increase the chiller energy (Bushby et al., 2001). If the required airflow rate drops below the minimum flow rate of the AHU(s), the use of reheat coils will be increased and thus causes additional reheating energy and the zone temperature may still be lower than the setpoint if the maximum heating capacity of reheating coils is reached (Cheung and Braun, 2015). In the heating season, the required airflow rate will increase as a consequence, which may result in lower boiler energy and higher reheating energy. On the contrary, a negative bias will lead to a higher supply temperature from the coils with potential higher humidity, which increases the required cooling airflow rate and thus results in an increased AHU fan energy consumption but the chiller energy and the reheating energy may be reduced (Bushby et al., 2001). In heating season, extra boiler energy will be caused by the fault. Similar to the temperature sensor bias along the water loop, this fault can be modeled by (EnergyPlus, 2020):

$$T_{coil-o,f} = T_{coil-o,ff} - \Delta T \quad (2.15)$$

where $T_{coil-o,f}$ is the actual coil SA temperature (°C); $T_{coil-o,ff}$ is the biased sensor reading of the coil SA temperature (°C); ΔT is the deviation of the sensor temperature reading from

the actual temperature (°C). In terms of the upper and lower bounds for the fault intensity range, Bushby et al. indicated $\pm 4^\circ\text{C}$ for their fault models.

(4) Economizer temperature sensor bias

The air economizer is a mechanical device integrated into the AHU(s) and is used to control the airflow rate of OA into the conditioned space depending on the conditions of OA and RA. It can save mechanical cooling energy and may enhance the indoor air quality if the OA is of low pollution (Zhang and Hong, 2017). There are several types of economizer sensors, including temperature sensor, humidity sensor and enthalpy sensor, can be applied to the economizer to achieve an optimal balance between OA and RA. For the economizer temperature sensor, the sensor types also vary with the location of sensor for OA, RA and MA. As a result, the economizer temperature sensor bias can be classified by the sensor locations. The symptom of these faults is similar to the coil SA temperature sensor bias, which is the deviation of temperature sensor reading from the actual air temperature. For economizer OA temperature sensor bias, a positive bias may intend to close the economizer damper to its minimum OA position when the full OA position is needed. On the contrary, a negative bias may lead to the full OA position of economizer when the OA temperature is higher than the RA temperature. The economizer RA temperature sensor bias will result in the opposite impact caused by the OA temperature sensor. If the RA temperature reading is higher than the actual temperature, the full OA condition settings of the economizer may be triggered and thus cause additional chiller energy. On the contrary, the portion of OA in the MA will be reduced and may affect the indoor air quality and increase the chiller energy (Bushby et al., 2001). Kim et al. (2019) indicated that the economizer temperature sensor bias for OA and RA can be modeled by replacing the fault free sensor temperature reading with the biased sensor value. EnergyPlus (2020) also contained these two faults in the native fault objects. The effect of the sensor biases is represented by:

$$T_f = T_{ff} \pm \Delta T \quad (2.16)$$

where T_f is the faulty reading of the economizer temperature sensor (°C); T_{ff} is the fault free value of the temperature (°C); ΔT is the deviation of the biased temperature sensor reading from the actual value of the temperature (°C). The MA temperature sensor bias is slightly different from the above two sensor biases because this type of sensor is not used to directly

control the economizer but mostly used to modulate the MA dampers to keep the MA temperature at the setpoint (Bushby et al., 2001). Thus, the modeling approach of this fault is difference from the above two faults. It is modeled by altering the MA setpoint schedule (Kim et al., 2019). The fault intensity range of the three types of temperature sensor biases is all defined as $\pm 4^{\circ}\text{C}$ (Bushby et al., 2001).

(5) Economizer relative humidity sensor bias

Apart from the fixed economizer control scheme relying on only temperature sensor(s), the relative humidity sensor(s) is usually integrated with the temperature sensor(s) to control the OA flow rate based on the enthalpy of OA and RA (Cheung and Braun, 2015). If the OA relative humidity sensor reading is higher than the actual relative humidity, the enthalpy of OA calculated by the economizer controller will be increased. Therefore, the OA damper may close to its minimum position when the actual enthalpy of OA is lower than that of RA and leads to additional chiller energy. The effect of this fault can be modeled by modifying the actual value of relative humidity by the bias (EnergyPlus, 2020):

$$RH_f = RH_{ff} \pm \Delta RH \quad (2.17)$$

where RH_f is the faulty reading of the economizer relative humidity sensor; RH_{ff} is the fault free value of the relative humidity; ΔRH is the deviation of the biased relative humidity sensor reading from the actual value of the relative humidity.

(6) Economizer enthalpy sensor bias

The economizer enthalpy sensor is a device which combines temperature and humidity sensors as described above, and thus the impact of this type of sensor bias is not repeated. The fault can be modeled by the following equation (EnergyPlus, 2020):

$$h_f = h_{ff} \pm \Delta h \quad (2.18)$$

where h_f is the faulty reading of the economizer enthalpy sensor (J); h_{ff} is the fault free value of the air enthalpy (J); Δh is the deviation of the biased enthalpy sensor reading from the actual value of the air enthalpy (J).

(7) Thermostat/Humidistat offset

The thermostat and humidistat are crucial devices for HVAC control, which monitor the temperature or humidity of air in the HVAC system or conditioned space to ensure that the zone air temperature and humidity are maintained within the designed comfort range. There are multiple parts in the thermostat and humidistat, but most of the faults happen at the sensors. The fault symptom is typically the thermostat or humidity sensor reading drifts from the actual values of zone air temperature or humidity. The faults are modeled in a similar way as the other sensor offsets (Zhang and Hong, 2017):

$$T_{stat,f} = T_{stat,ff} \pm \Delta T_{stat} \quad (2.19)$$

$$RH_{stat,f} = RH_{stat,ff} \pm \Delta RH_{stat} \quad (2.20)$$

where $T_{stat,f}/RH_{stat,f}$ is the faulty reading of thermostat (°C) /humidistat; $T_{stat,ff}/RH_{stat,ff}$ is the fault free value of actual temperature (°C) /relative humidity; $\Delta T_{stat}/\Delta RH_{stat}$ is the deviation of the offset thermostat/humidity reading from the actual value of temperature (°C) /relative humidity.

For the humidistat offset fault, it can be further grouped into the thermostat dependent fault and thermostat independent fault. For the thermostat dependent fault, the thermostat offset will lead to the humidistat offset and thus the offset of relative humidity ΔRH_{stat} will be influenced by both the thermostat offset and the actual zone air conditions. This type of offset is given by:

$$\Delta RH_{stat} = RH_{stat,ff} - f(T_{real}, W_f) \quad (2.21)$$

where T_{real} is the real-time value of zone air temperature (°C); W_f is the humidity ratio corresponding to $T_{real} \pm \Delta T$ and $RH_{stat,ff}$ (kg_{water}/kg_{dry air}). As for the humidistat independent of the thermostat, the offset can be simply defined by a pre-defined schedule.

2.2.4.3 Chiller Faults

Chiller is a critical central plant component in the HVAC system to support the cooling demand of conditioned zones and it typically consist of four main components, namely, condenser, evaporator, expansion device and compressor, as illustrated in Figure 2.5. One of the widely used types of chillers is the vapor compression chiller. Cheung and Braun (2013a,

2013b) developed the component models of a typical vapor compression chiller, in order to simulate the impacts of common faults occurring in vapor compression system. In addition, Cheung and Braun (2015) explained the modeling method of some typical chiller faults in their recent work. Based on the review of previous studies, the chiller fault models have been developed in the current field are as follows: (1) Non-standard refrigerant charging; (2) Heat exchanger fouling; (3) Liquid line restriction; (4) Presence of non-condensable gas in refrigerant circuit; (5) Compressor valve leakage; (6) Excess oil in water-cooled chillers. The descriptions for corresponding characteristics of these faults are shown in Table 2.6. Cheung and Braun (2016) also developed a black box model to predict the water-cooled chiller fault impact ratio (FIR), as described by:

$$FIR = 1 + F \left(c_1 + c_2 T_{evap,out} + c_3 T_{cond,in} + c_4 \frac{\dot{Q}_{evap}}{\dot{Q}_{evap,rat}} \right)^2 + c_5 \left(\frac{\dot{Q}_{evap}}{\dot{Q}_{evap,rat}} \right) F + c_6 F \quad (2.22)$$

where F is the fault intensity, $T_{evap,out}$ is the water temperature at the evaporator outlet (K), $T_{cond,in}$ is the water temperature at the condenser inlet (K), $\dot{Q}_{evap}$ is the evaporator heat removal rate (kW), $\dot{Q}_{evap,rat}$ is the rated evaporator heat removal rate (kW), $c_1, c_2, c_3, c_4, c_5, c_6$ are the coefficients which depends on the fault type. The detailed values for these coefficients are given in the paper of Cheung and Braun (2016).

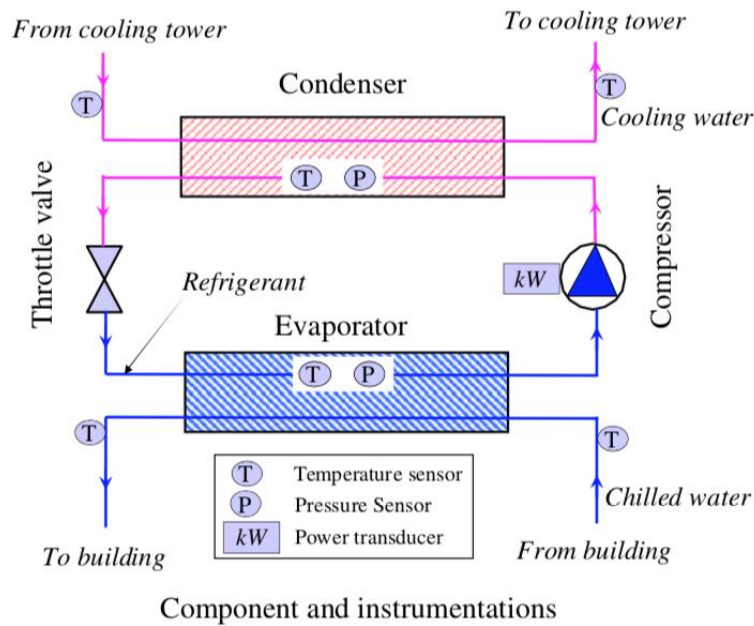


Figure 2.5 Diagram of the main components and working mechanisms of a typical water-cooled chiller (Xu, Xiao and Wang, 2008).

Table 2.6 Fault models and corresponding fault characteristics for chillers.

Fault models	Fault intensity definition	Fault intensity range	Fault evolution
Non-standard refrigerant charging	The percentage difference of the actual charge level from the level specified by the manufacturer	-30% to 40%	Y
Heat exchanger fouling	The percentage of water/air flow rate reduced due to fouling	0 to 50%	Y
Refrigerant liquid line restriction	The percentage of variation in the pressure difference between the condenser outlet and evaporator inlet due to the liquid line restriction	0 to 30%	Y
Presence of non-condensable gas in refrigerant circuit	The ratio of the volume of non-condensable gas in the refrigerant circuit to the maximum volume of non-condensable gas that can exist in the refrigerant circuit at atmospheric pressure	0 to 60%	Y
Compressor valve leakage	The ratio of the leaked refrigerant mass flow rate to the refrigerant mass flow rate without the valve leakage	N/A	N/A
Excess oil in water-cooled chillers	The ratio of the excessive mass of oil in the chiller to the mass of oil specified by the manufacturer	0 to 73%	N/A

(1) Non-standard refrigerant charging

The non-standard refrigerant charging fault occurs when the actual charged amount of refrigerant deviates from the amount of refrigerant specified by the manufacturer. This fault can be further classified as undercharging and overcharging refrigerant. Both of the two classifications will lead to the COP reduction of the chiller (Li and O'Neill, 2018). The faulty charging level is represented by the percentage difference of the actual charge level from the level specified by the manufacturer (Cheung and Braun, 2013b):

$$charge_{faulty} = \frac{actual\ charge\ level - specified\ level}{specified\ level} \quad (2.23)$$

Cheung and Braun (2015) performed an empirical approach to train the refrigerant overcharging model using experimental data with fault levels between 0 to 40%. The training was conducted based on a modified FIR equation of the chiller power consumption:

$$\frac{\dot{W}_{chiller,f}}{\dot{W}_{chiller}} = 1 + F^2 \left(c_1 + c_2 T_{evap,out} + c_3 T_{cond,in} + c_4 \frac{\dot{Q}_{evap}}{\dot{Q}_{evap,rat}} + c_5 \left(\frac{\dot{Q}_{evap}}{\dot{Q}_{evap,rat}} \right)^2 \right)^2 \quad (2.24)$$

where $\dot{W}_{chiller,f}$ is the faulty chiller power consumption (kW) and $\dot{W}_{chiller}$ is the fault free chiller power consumption (kW). The results showed that the overcharging refrigerant fault increased the building electricity usage by 0.7%.

(2) Heat exchanger fouling

Evaporator and condenser are both heat exchangers in a chiller. The heat exchanger fouling occurs due to the accumulation of dirt or deposits on the surfaces of the heat exchanger. The condenser fouling on the water side of a chiller can increase the resistance of water flow and decrease the water flow rate. It will also decrease the heat exchange area of the condenser. As a result, the heat rejection rate of the condenser and the chiller power efficiency will be reduced (Cheung and Braun, 2016). The evaporator fouling can be described in terms of the fouling level (Cheung and Braun, 2013b):

$$Evaporator \text{ fouling level} = 1 - \frac{Reduced \text{ air mass flow}}{Unfouled \text{ air mass flow}} \quad (2.25)$$

The condenser fouling level is defined as the percentage of water flow rate reduced due to fouling (Cheung and Braun, 2015):

$$Condenser \text{ fouling level} = 1 - \frac{Reduced \text{ water mass flow}}{Unfouled \text{ water mass flow}} \quad (2.26)$$

Based on the Eqn. (2.26), Cheung and Braun (2015) trained a condenser fouling model using data with fault level of 0 to 45%. EnergyPlus (2020) provided a native fault object of water-cooled chiller fouling which models the fault impact on the chiller cooling capacity and the chiller COP, as given by the following equations:

$$Q_{chiller,f} = Q_{chiller,ff} \times F_{chiller} \quad (2.27)$$

$$COP_{chiller,f} = COP_{chiller,ff} \times F_{chiller} \quad (2.28)$$

where $Q_{chiller,f}$ is the faulty chiller cooling capacity (kW); $Q_{chiller,ff}$ is the fault free chiller cooling capacity (kW); $COP_{chiller,f}$ is the faulty chiller COP; $COP_{chiller,ff}$ is the fault free chiller COP; $F_{chiller}$ is the fouling ratio used to reduce the chiller cooling capacity and COP.

(3) Liquid line restriction

The liquid line restriction is caused by the accumulation of deposits in the refrigerant filter that blocks the liquid line between the expansion valve and the condenser. This fault will increase the refrigerant flow resistance and the pressure difference across the compressor, which will further lead to higher compressor power consumption. Besides, this fault will reduce the mass flow rate of refrigerant in the evaporator, the evaporating temperature and the density of refrigerant. The decrease in evaporating temperature will lead to lower sensible heat ratio and the decrease in the density of refrigerant will result in lower refrigerant mass flow rate in compressor and evaporator heat removal rate (Cheung and Braun, 2015). This fault model was developed by introducing a liquid line restriction level (LL) to the fault free pressure at the liquid line outlet:

$$P_{r,liquidline,out,f} = P_{r,liquidline,out,ff} - LL(P_{r,liquidline,in} - P_{r,evap,in}) \quad (2.29)$$

where $P_{r,liquidline,out,f}$ is the faulty pressure at the outlet of the liquid line (Pa), $P_{r,liquidline,out,ff}$ is the fault free pressure at the outlet of the liquid line (Pa), LL is the liquid line restriction level, $P_{r,liquidline,in}$ is the pressure at the inlet of the liquid line (Pa), $P_{r,evap,in}$ is the pressure at the inlet of evaporator (Pa). It should be noted that $P_{r,liquidline,in}$ and $P_{r,evap,in}$ are obtained from the condenser outlet and the evaporator inlet under the fault free case.

The liquid line restriction level LL was defined by Breuker and Braun (1998) as follows:

$$restriction\ level = \frac{\Delta P_f}{\Delta P_{ff}} \times 100\% \quad (2.30)$$

where ΔP_f is the pressure drop across the restriction (Pa); ΔP_{ff} is the fault free pressure drop across (Pa). The fault level is given by the percentage of variation in the pressure difference

between the condenser outlet and evaporator inlet due to the liquid line restriction and the fault level range is indicated as 0 to 30% in the study of (Kim et al., 2019).

(4) Presence of non-condensable gas in the refrigerant circuit

The presence of non-condensable gas in the refrigerant circuit is a fault which occurs when the chiller is not fully evacuated prior to being charged with refrigerant. Consequently, the chiller operates with a mixture of non-condensable gas, typically air, and refrigerant rather than a pure refrigerant. Since the non-condensable gas keeps in the vapor state in the refrigerant circuit, it will be trapped in the superheated vapor downstream of the compressor. Thus, the discharge pressure of the compressor and the compressor power consumption will be increased (Cheung and Braun, 2016). The fault level is described by the ratio of the volume of non-condensable gas in the refrigerant circuit to the maximum volume of non-condensable gas that can exist in the refrigerant circuit at atmospheric pressure (Kim et al., 2009). Cheung and Braun (2013b) used this definition of fault level, as shown in Eqn. (2.31), to estimate the partial pressure of non-condensable gas and total pressure of the mixture of the non-condensable gas and the refrigerant by Eqn. (2.32) and Eqn. (2.33), respectively:

$$NC = \frac{M_{NC}}{\rho_{NC}(P = 101.325 \text{ kPa}, T = 26.7 \text{ }^\circ\text{C})V_{total}} \quad (2.31)$$

where NC is the non-condensable gas level, M_{NC} is the mass of non-condensable gas (kg), ρ_{NC} is the density of non-condensable gas at a pressure of 101.325 kPa and temperature of 26.7 °C (kg/ m³), V_{total} is the total volume of the refrigerant circuit (m³).

$$P_{p,NC} = \frac{M_{NC}R_{NC}T_{comp,out}}{V_{cond}(w_{sh,cond} + w_{tp,cond}\bar{y}_{r,cond,tp}) + V_{hg}} \quad (2.32)$$

$$P_{total} = P_{p,NC} + P_{p,r} \quad (2.33)$$

where $P_{p,NC}$ is the partial pressure of non-condensable gas (kPa), R_{NC} is the gas constant of non-condensable gas (kJ/kgK), $T_{comp,out}$ is the compressor outlet temperature (K), V_{cond} is the volume of the condenser (m³), $w_{sh,cond}$ is the area ratio of the superheated vapor in the condenser, $w_{tp,cond}$ is the area ratio of the two phase flow in the condenser, $\bar{y}_{r,cond,tp}$ is the

void fraction of the two phase flow in the condenser, V_{hg} is the volume of hot gas line (m^3), $P_{p,r}$ is the partial pressure of refrigerant (kPa), P_{total} is the total pressure (kPa).

(5) Compressor valve leakage

The compressor valve leakage includes the leakage of the discharge valve and the suction valve, which leads to a reverse flow of refrigerant from the compressor discharge to the suction. This fault is modeled by introducing a valve leakage ratio VL (Cheung and Braun, 2013b), as described in Eqn. (2.34):

$$VL = 1 - \frac{\text{Refrigerant mass flow rate under valve leakage}}{\text{Refrigerant mass flow rate without valve leakage}} \quad (2.34)$$

Hence, the fault model is established as follows (Cheung and Braun, 2013b):

$$\dot{m}_{r,comp,f} = (1 - VL)\dot{m}_{r,comp,ff} \quad (2.35)$$

$$h_{r,comp,in} = h_{r,suctionline,out}(1 - VL) + h_{r,comp,out}VL \quad (2.36)$$

$$h_{r,comp,f} = f_{comp}(P_{r,in}, P_{r,out}, h_{r,comp,in}) \quad (2.37)$$

where $\dot{m}_{r,comp,f}$ is the refrigerant mass flow rate in the compressor under the valve leakage (kg/s), $\dot{m}_{r,comp,ff}$ is the refrigerant mass flow rate in the compressor without the valve leakage (kg/s), $h_{r,comp,in}$ is the refrigerant enthalpy at the compressor inlet (kJ), $h_{r,suction,out}$ is the refrigerant enthalpy at the outlet of the suction line (kJ), $h_{r,comp,out}$ is the refrigerant enthalpy at the compressor outlet (kJ), $h_{r,comp,f}$ is the refrigerant enthalpy at the compressor under the valve leakage (kg/s), $P_{r,in}$ is the pressure at the compressor inlet (Pa), $P_{r,out}$ is the pressure at the compressor outlet (Pa).

(6) Excess oil in water-cooled chillers

The excessive oil in water-cooled chillers is caused by surplus lubricating oil injected into the refrigerant circuit of the chiller. It occurs when some refrigerant is absorbed by the excessive oil and thus the amount of refrigerant in the vapor compression cycle is decreased, which further induces a reduction in chiller efficiency (Cheung and Braun, 2016). The fault level is defined as the ratio of the excessive mass of oil in the chiller to the mass of oil specified by

the manufacturer. The chiller power consumption FIR model was trained by data with the fault level ranging from 0 to 73%, based on the study of Cheung and Braun (2015).

2.2.4.4 RTU and Split Air Conditioner Faults

RTU and split air conditioners are typically employed in small scale buildings. They operate on the vapor compression cycle, which is the same as most of the chillers. Therefore, the faults for these air conditioners are mostly the same as those that happened in chillers. Cheung and Braun (2015) listed five types of faults that are common in RTU and split air conditioners: (1) Non-standard refrigerant charging; (2) Heat exchanger fouling; (3) Condenser fan motor degradation; (4) Liquid line restriction; (5) Presence of non-condensable gas in the refrigerant circuit. The descriptions for corresponding characteristics of these faults are shown in Table 2.7.

Table 2.7 Fault models and corresponding fault characteristics for RTUs and split air conditioners.

Fault models	Fault intensity definition	Fault intensity range	Fault evolution
Non-standard refrigerant charging	The percentage difference of the actual charge level from the level specified by the manufacturer	-30% to 40%	Y
Heat exchanger fouling	The percentage of air flow rate reduced due to fouling	0 to 50%	Y
Condenser fan motor degradation	The ratio of fan motor efficiency reduction to the fault free fan motor efficiency	0 to 30%	Y
Refrigerant liquid line restriction	The percentage of variation in the pressure difference between the condenser outlet and evaporator inlet due to the liquid line restriction	0 to 30%	Y
Presence of non-condensable gas in refrigerant circuit	The ratio of the volume of non-condensable gas in the refrigerant circuit to the maximum volume of non-condensable gas that can exist in the refrigerant circuit at atmospheric pressure	0 to 80%	Y

Yuill, Cheung and Braun (2014) proposed the FIR for cooling capacity and COP as follows:

$$FIR_{capacity} = \frac{Capacity_{fault}}{Capacity_{faultfree}} \quad (2.38)$$

$$FIR_{COP} = \frac{COP_{fault}}{COP_{faultfree}} \quad (2.39)$$

Based on these FIRs (Yuill, Cheung and Braun, 2014), Cheung and Braun (2015) derived the specific FIRs of cooling capacity and energy input ratio as well as the sensible heat ratio, as described in following equations:

$$\begin{aligned} FIR_{\dot{Q}} &= \frac{\dot{Q}_{cool,dx,F}}{\dot{Q}_{cool,dx}} \\ &= 1 + F\{c_{F,0} + c_{F,1}\left(\frac{T_{wb,ent}}{273.15\text{ K}} + 1\right) + c_{F,2}\left(\frac{T_{db,amb}}{273.15\text{ K}} + 1\right) + c_{F,3}F + c_{F,4}F^2 \\ &\quad + c_{F,5}\frac{\dot{Q}_{cool,dx,rat}}{\dot{W}_{dx,rat}} \end{aligned} \quad (2.40)$$

$$\begin{aligned} FIR_{EIR} &= \frac{EIR_{dx,F}}{EIR_{dx}} \\ &= 1 + F\{c_{F,6} + c_{F,7}\left(\frac{T_{wb,ent}}{273.15\text{ K}} + 1\right) + c_{F,8}\left(\frac{T_{db,amb}}{273.15\text{ K}} + 1\right) + c_{F,9}F + c_{F,10}F^2 \\ &\quad + c_{F,11}\frac{\dot{Q}_{cool,dx,rat}}{\dot{W}_{dx,rat}} \end{aligned} \quad (2.41)$$

$$UA_F = (1 + c_{F,12}F)UA_{rat} \quad (2.42)$$

where $\dot{Q}_{cool,dx,F}$ is the cooling capacity of the faulty DX cooling coil (kW), $\dot{Q}_{cool,dx}$ is the cooling capacity of the fault free DX cooling coil (kW), $EIR_{dx,F}$ is the energy input ratio of the faulty DX cooling coil, EIR_{dx} is the energy input ratio of the fault free DX cooling coil, F is the fault level, $c_{F,0} \sim c_{F,12}$ are the coefficients depending on the fault type, $T_{wb,ent}$ is the wet-bulb temperature at the evaporator inlet (K), $T_{db,amb}$ is the dry-bulb temperature at the ambient environment (K), $\dot{Q}_{cool,dx,rat}$ is the rated cooling capacity of the DX cooling coil (kW), $\dot{W}_{dx,rat}$ is the rated electrical power consumption of the DX cooling coil (kW), UA_F is the overall heat transfer coefficient of the faulty DX cooling coil (kW/K), UA_{rat} is the rated overall heat transfer coefficient of the fault free DX cooling coil (kW/K).

(1) *Non-standard refrigerant charging*

The non-standard refrigerant charging of RTU and split air conditioners occurs in a similar manner as that of chillers, as presented in subsection 2.2.4.3.

(2) *Heat exchanger fouling*

The heat exchanger fouling of RTU and split air conditioners has similar fault characteristics as that of chillers and can be modeled in a similar way. The only difference between the two types of components is the condenser of RTU and split air conditioners are typically air cooled, and the condenser fans are usually single speed. Hence, the condenser fouling will lead to the decrease in condenser air flow rate and increase the condensing pressure and temperature, which further increases the compressor pressure difference and the power consumption (Cheung and Braun, 2015). The fault level is defined as the percentage of airflow rate reduced due to fouling and ranges from 0 to 50% based on the literature of Kim et al. (2019).

(3) *Condenser fan motor degradation*

The condenser fan motor degradation is similar to the AHU fan motor degradation, which is caused by a bearing fault or a stator winding fault. Both of the bearing fault and stator winding fault will lead to higher fan power consumption without affecting the fan airflow. To model this fault, the FIR for the energy input ratio (EIR) of the air conditioner was proposed by Cheung and Braun (2015):

$$FIR_{EIR} = \frac{EIR_F}{EIR} = 1 + \frac{\dot{W}_{fan}}{\dot{W}_{cool}} \left(\frac{F}{1 - F} \right) \quad (2.43)$$

where $\dot{W}_{fan}$ is the condenser fan power consumption (kJ), $\dot{W}_{cool}$ is the compressor power consumption (kJ), F is the fault level.

The condenser fan motor degradation can be modeled by altering the EIR of the air conditioner, as given in Eqn. (2.43). The fault level is defined as the ratio of fan motor efficiency reduction to the fault free fan motor efficiency and ranges from 0 to 30% based on the literature of Kim et al. (2019).

(4) Refrigerant liquid line restriction

The modeling approach and the characteristics of refrigerant liquid line restriction in RTU and split air conditioners are similar to those in chillers, as presented in subsection 2.2.4.3.

(5) Presence of non-condensable gas in the refrigerant circuit

The presence of non-condensable gas in the refrigerant circuit in RTU and split air conditioners is similar to that occurring in chillers. The fault model and fault characteristics can be found in subsection 2.2.4.3.

2.2.4.5 AHU Faults

AHU is a key component in an HVAC system, which serves the heating and cooling functions for conditioned zones. It is integrated with an air distribution system to meet the thermal comfort and indoor air quality requirements of occupant for each zone. A typical AHU consist of heating and cooling coils, OA and RA dampers, a fan and ducts, as shown in Figure 2.6. According to Yu, Woradechjumroen and Yu (2014), there are four main categories of AHU faults, including AHU equipment faults, AHU actuator faults, AHU sensor faults, AHU controller faults. However, only a portion of the AHU faults listed in the work of Yu, Woradechjumroen and Yu (2014) has been modeled in the field. Based on the review of previous fault modeling studies, there are overall nine types of AHU fault models which have been created in the field: (1) AHU fan motor degradation; (2) AHU fan stuck at a minimum speed; (3) OA/RA damper stuck; (4) OA/RA damper leakage; (5) Cooling/heating coil valve stuck; (6) Cooling/heating coil valve leakage; (7) Cooling/heating coil fouling; (8) SA duct leakage; (9) Return air duct leakage. The descriptions for corresponding characteristics of these faults are shown in Table 2.8.

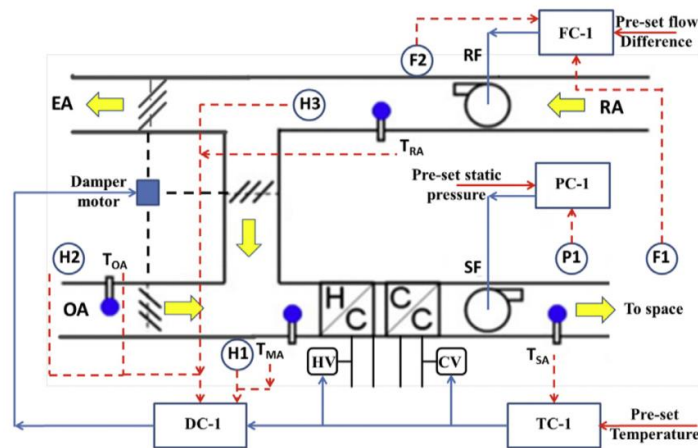


Figure 2.6 Diagram of components in a typical AHU (Yu, Woradechjumroen and Yu, 2014).

Table 2.8 Fault models and corresponding fault characteristics for AHUs and VAV terminals.

Fault models	Fault intensity definition	Fault intensity range	Fault evolution
AHU fan motor degradation	The percentage of fan motor efficiency reduction	0 to 30%	Y
AHU fan stuck at a minimum speed	The fan minimum flow fraction f_f	N/A	N/A
OA/RA damper stuck	The opening percentage that the damper stuck at	0 to 100%	N
OA/RA damper leakage	The leakage percentage of the full OA/RA flow rate	0 to 40%	Y
Cooling/heating coil valve stuck	The valve stuck position	0 to 100%	N
Cooling/heating coil valve leakage	The change of valve capacity K_v	N/A	Y
Cooling/heating coil fouling	Reduced overall heat transfer coefficient (UA)	N/A	Y
SA duct leakage	The percentage of the leaking flow rate to the SA flow rate	0 to 30%	N
RA duct leakage	The percentage of unconditioned air entering the return duct to the RA flow rate	0 to 30%	Y
VAV box stuck damper	The VAV box damper stuck position	N/A	N/A
VAV box stuck reheat coil valve	The valve stuck position	N/A	N/A

(1) AHU fan motor efficiency degradation

The fan motor efficiency degradation is a typical fault symptom for AHUs due to bearing or (and) stator winding faults. The causes for the bearing and stator winding faults have been explained in condenser fan motor efficiency degradation of chillers faults and thus not repeated in this part. Since these faults only affect the fan power consumption, the fan motor efficiency degradation can be modeled by simply altering the fan motor efficiency (Kim et al., 2019). The fault intensity of this fault is defined as the percentage of fan motor efficiency reduction, and the limit for the efficiency degradation percentage is 30% based on.

(2) AHU fan stuck at minimum speed

There are generally two types of AHU systems: constant air volume (CAV) system and VAV system. For the VAV system, the AHU supply fan will adjust the SA flow rate on the basis of

the zone load. This fault occurs when describes the supply fan is stuck at minimum speed which is higher than the nominal minimum speed and is caused by either inoperative fan controller or incorrect control signal. Li and O'Neill (2016) modeled this fault using the embedded equations of fan power consumption and air flow rate in EnergyPlus (2020) as follows:

$$f_f = m_s / m_{s,max} \quad (2.44)$$

$$f_{pl} = c_1 + c_2 f_f + c_3 f_f^2 + c_4 f_f^3 + c_5 f_f^4 \quad (2.45)$$

$$Q_p = f_{pl} m_{s,max} \delta P / (e_{tot} \rho_{air}) \quad (2.46)$$

where m_s is the actual mass flow rate of SA (kg/s), $m_{s,max}$ is the maximum or design mass flow rate of SA (kg/s), c_1, c_2, c_3, c_4, c_5 are the fan power coefficients, f_f is the flow fraction, f_{pl} is the part load factor, δP is the design pressure difference (Pa), e_{tot} is the fan total efficiency, ρ_{air} is the density of air (kg/m³) and Q_p is the fan total power (W).

According to Eqn. (2.44) to (2.46), this fault will lead to a higher flow fraction f_f and a higher part load factor f_{pl} when the AHU fan is required to deliver an airflow of minimum flow rate. Consequently, the fan total power will be increased under a higher minimum flow rate than the nominal minimum flow rate.

(3) OA/RA damper stuck

Economizer damper stuck can be caused by a variety of factors, such as economizer control faults or seized actuators. This fault can significantly influence the HVAC system energy consumption and the thermal comfort of conditioned zones especially when the damper is stuck at either fully open or fully closed. Kim et al. (2019) proposed that the modeling of this fault can be implemented by setting the damper at a constant opening percentage in the algorithm of economized control. The fault intensity is defined as the opening percentage that the damper is stuck at. Li and O'Neill (2016) modeled the OA damper with a stuck minimum opening percentage by proposing a schedule factor for adjusting the minimum OA flow rate in EnergyPlus/Openstudio. The actual minimum OA flow rate under the damper stuck is given by:

$$V_{min,f} = V_{min} \times k \quad (2.47)$$

where $V_{min,f}$ is actual minimum OA flow rate (m^3/s); V_{min} is the normal minimum OA flow rate (m^3/s); k is the schedule factor adjusting the actual minimum OA flow rate based on the stuck minimum damper position.

(4) OA/RA damper leakage

The economizer damper leakage will result in excess OA or RA entering AHU(s) in the case that the required OA/RA flow rate is less than the leaking amount of OA/RA. Thus, it causes an increase of heating or cooling energy. Basarkar et al. (2011) modeled this fault by simply modifying the actual OA/RA flow rate by the maximum value between the required OA/RA flow rate and the leaking OA/RA flow rate, as given by Eqn. (2.48):

$$Q_{act} = \max(Q_f, Q_{ff}) \quad (2.48)$$

where Q_{act} is the actual OA/RA flow rate (m^3/s); Q_f is the faulty (leaking) OA/RA flow rate (m^3/s); Q_{ff} is the fault free (required) OA/RA flow rate (m^3/s). The fault intensity is defined as the leakage percentage of the full OA/RA flow rate. Bushby et al. (2001) modeled this fault with intensity between 0 to 40%.

(5) A stuck cooling/heating coil valve

The stuck cooling/heating coil valve is one of the common faults occurring in cooling/heating coils. For a stuck cooling coil valve, the zone temperature setpoint will not be able to be achieved, and thus the required zone temperature will not be achieved. But the cooling energy may be reduced since the SA is not fully cooled to its temperature setpoint. For a stuck heating coil valve, the insufficient heating may be compromised by the reheat coils in VAV boxes. Similar to the stuck economizer damper, the stuck valve can also be modeled by modifying the valve opening percentage. The fault intensity is defined as the opening position that the valve is stuck at. Pourarian et al. (2017) developed a coil model to simulate this fault, as described by:

$$\frac{1}{m_c^2} = \frac{1296}{\Delta P K_v^2} (ax_v + b)^{-2} + \frac{R_{coil}}{\Delta P} \quad (2.49)$$

where m_c is the water flow rate through the coil (kg/s), ΔP is pressure drop across the coil and valve (kPa), K_v is the valve capacity (m^3/h), a, b are the coefficients of valve depending

on the valve positions, x_v is the valve position in $[0, 1]$, R_{coil} is the coil resistance ($0.001 \text{ kg} \cdot \text{m}$). The stuck coil valve can be then modeled by modifying the valve position x_v , and the fault intensity is defined as the valve stuck position.

(6) Cooling/heating coil valve leakage

In addition to the stuck coil valve, cooling/heating coil valve leakage is another major fault occurring in cooling/heating coils. In the heating season, the cooling coil valve leakage will require additional heating to maintain the SA at the setpoint. In the cooling season, the cooling coil valve leakage will lead to excess cooling of the SA, and the heating coil valve leakage will result in additional cooling to maintain the SA at the setpoint. The coil model developed by Pourarian et al. (2017) can also be used to model this fault by modifying the valve capacity K_v in Eqn. (2.49) (Li and O'Neill, 2018).

(7) Cooling/heating coil fouling

The cooling/heating coil fouling is also a common fault for cooling/heating coils. This fault usually happens when the dirt accumulates at the coil. It will cause degraded overall heat transfer coefficient and lower coil cooling/heating capacity, which may lead to the increase in water flow rate and the decrease in temperature difference at the water side. On the basis of literature review, there are generally three approaches to model this fault.

The first approach is proposed by Zhang and Hong (2017) and is also included in EnergyPlus (2020). This approach gives two ways to model the degraded overall heat transfer coefficient UA : Fouled UA Rated and Fouling Factor. For the Fouled UA Rated method, the fouled overall heat transfer coefficient (UA_f) is defined by the users directly. For the Fouling Factor method, the air/water side fouling factor is specified by the users and the fouled overall heat transfer coefficient (UA_f) is calculated according to the following equations:

$$R_{foul} = \frac{r_{air}}{A_{air}} + \frac{r_{water}}{A_{water}} \quad (2.50)$$

$$UA_f = [(UA_{air})^{-1} + R_{foul} + (UA_{water})^{-1}]^{-1} \quad (2.51)$$

where r_{air} is the air side fouling factor ($\text{m}^2\text{K/W}$), r_{water} is the water side fouling factor ($\text{m}^2\text{K/W}$), A_{air} is the surface area of air side coil (m^2), A_{water} is the surface area of water side

coil (m^2), UA_{air} is the heat transfer coefficient of the air side coil (W/K), UA_{water} is the heat transfer coefficient of the water side coil (W/K), R_{foul} is the overall fouling factor (K/W), UA_f is the fouled overall heat transfer coefficient of the coil (W/K).

Yi and O'Neill (2016) proposed the second approach by adding a UA degrading factor a to the normal UA_0 for the unfouled coil:

$$UA = UA_0(1 - a) \quad (2.52)$$

where UA_0 is the normal overall heat transfer coefficient (W/K), a is the normal overall heat transfer coefficient degrading factor under a range of $[0, 1]$, UA is the fouled overall heat transfer coefficient (W/K).

The third modeling approach is to develop this fault by introducing the log mean temperature difference (LMTD) under the coil fouling condition (Basarkar et al., 2011):

$$LMTD = \frac{\Delta T_A - \Delta T_B}{\ln\left(\frac{\Delta T_A}{\Delta T_B}\right)} \quad (2.53)$$

$$UA_f = \frac{Q}{LMTD} \quad (2.54)$$

where ΔT_A is the temperature difference between the air outlet and the water inlet (K); ΔT_B is the temperature difference between the water outlet and the air inlet (K); $LMTD$ is the log mean temperature difference (K); Q is the coil overall heat transfer rate (W); UA_f is the fouled overall heat transfer coefficient of the coil (W/K).

(8) Supply/Return air duct leakage

Air duct is a key component that is connected with the AHU to deliver the SA from the AHU to the conditioned zones. Duct leakage is a common fault which can be led by broken duct wrap, duct disconnections, temperature and pressure cycling, etc (Kim et al., 2019). The SA duct leakage will increase the heating or cooling energy and also increase fan power consumption for VAV system since the SA is leaked into the unconditioned space. The return air duct leakage will lead to the same impact as the SA duct leakage for the reason that the

unconditioned air is introduced into the duct and mixed with the return air. To model these faults, Bonvini et al. (2014) developed an air duct model as given by the following equations:

$$\dot{m}_a + \dot{m}_l = k\sqrt{\Delta p} \quad (2.55)$$

where $\dot{m}_a$ is the actual mass flow rate after the leaking position of the duct (kg/s); $\dot{m}_l$ is the leaking flow rate (kg/s); k is the friction factor; Δp is the pressure drop (Pa).

Apart from this, Kim et al. (2019) indicated that the return air duct leakage could be modeled by changing the OA flow rate entering the AHU, but the modeling approach for the SA duct leakage was not mentioned. The fault intensities for supply and return air duct leakage are defined as the percentage of the leaking flow rate to the SA flow rate and the percentage of unconditioned air entering the return duct to the return air flow rate, respectively.

2.2.4.6 VAV Terminal Faults

VAV terminal is a critical component in the VAV system which is used to adjust the airflow rate based on the cooling/heating demand of each zone. There are mainly two types of faults in the VAV terminal: stuck VAV terminal damper and stuck reheat coil valve. The descriptions for corresponding characteristics of these faults are shown in Table 2.8.

(1) Stuck VAV terminal damper

The stuck VAV terminal damper will lead to the SA flow to the served zone is not able to adjust and thus affect the occupant thermal comfort and indoor air quality in the served zone. Bushby et al. (2001) modeled this fault by setting the damper position for the VAV terminal to a constant value.

(2) Stuck reheat coil valve

The reheat coil valve stuck will have different fault impacts depending on the cooling mode or heating mode of the VAV system. In cooling mode, a stuck reheat coil valve at its fully closed position may have no influence on the zone occupant thermal comfort and energy consumption. A stuck coil valve at an open position may increase the VAV fan energy consumption and affect the zone occupant thermal comfort. In the heating season, a stuck fully closed coil valve may cause the lower SA temperature than the setpoint and thus

decrease the zone thermal comfort level. If the valve is stuck at an open position, additional heating energy may be wasted. This fault is modeled by setting the valve position in the valve actuator to a constant value Bushby et al. (2001).

2.2.4.7 Other Faults

Apart from the above six categories of faults, there are still some uncategorized faults: (1) Duct fouling; (2) Blower motor efficiency degradation; (3) Cooling tower fouling; (4) Hot water boiler fouling; (5) Evaporative cooler fouling; (6) Dirty air filter. The descriptions for corresponding characteristics of these faults are shown in Table 2.9.

Table 2.9 Fault models and corresponding fault characteristics for other HVAC faults.

Fault models	Fault intensity definition	Fault intensity range	Fault evolution
Duct fouling	The ratio of the increased pressure difference across the fan caused by duct fouling to the pressure difference across the fan at the rated speed	N/A	N
Blower motor efficiency degradation	The ratio of fan motor efficiency reduction to the fault free fan motor efficiency	N/A	Y
Cooling tower fouling	The UA reduction factor	N/A	N/A
Boiler fouling	The reduction factor for the boiler capacity and thermal efficiency	N/A	N/A
Evaporative cooler fouling	The reduction factor for the indirect stage efficiency of the evaporator cooler	N/A	N/A
Dirty air filter	The faulty fan curve and pressure rise schedule	N/A	N/A

(1) Duct fouling

Duct fouling occurs when the accumulation of dust at the duct filters or the fins of heat exchangers blocks the airflow. As a result, the flow resistance of the duct and the pressure difference across the duct are both increased. Figure 2.7 shows the variations of fan and duct operating states due to duct fouling (Cheung and Braun, 2015). For a single speed fan, the normal operating state of the fan can be found at point A in Figure 2.7. The duct fouling will result in the upward shift of the duct curve, and hence the fan operating state will move from

point A to point B. For a variable speed fan, the normal operating state with a lower speed of the fan can be found at point C. The fouled duct will cause an increase in fan speed from speed 2 to speed 1 to keep the volumetric air flow rate at V_3 . This fault can be model by changing the pressure difference across the fan from the value at the normal duct curve to that at the fouled duct curve. The fault intensity is defined as the ratio of the increased pressure difference across the fan caused by duct fouling to the pressure difference across the fan at the rated speed.

(2) Blower motor efficiency degradation

The blower motor efficiency degradation has the same mechanism as the condenser fan motor degradation in the RTU and split air conditioner faults. Thus, it also results in a higher motor power consumption without changing the airflow across the blower. This fault can be simulated using the same fault model as given by Eqn. (43) in subsection 2.2.4.4. The fault intensity is also defined as the ratio of fan motor efficiency reduction to the fault free fan motor efficiency.

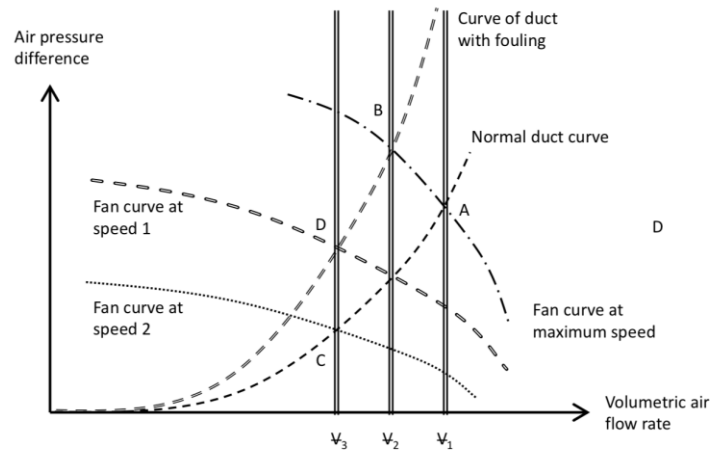


Figure 2.7 Fan and duct characteristics curves at normal and fouling condition (Cheung and Braun, 2015).

(3) Cooling tower fouling

Fouling is also a common fault occurring in the cooling tower due to the accumulation of deposits from water. It will decrease the overall heat transfer coefficient UA of the cooling tower, which further reduces the energy efficiency of the tower. EnergyPlus (2020) provided a native fault object of cooling tower fouling by the following equation:

$$UA_{tower,f} = UA_{tower,ff} \times F_{UA} \quad (2.56)$$

where $UA_{tower,f}$ is the overall heat transfer coefficient of the fouled cooling tower (W/K); $UA_{tower,ff}$ is the overall heat transfer coefficient of the fault free cooling tower (W/K); F_{UA} is the UA reduction factor due to cooling tower fouling.

(4) Boiler fouling

The boiler fouling is also caused by the accumulation of deposits from water and decreases the capacity and thermal efficiency of the boiler. EnergyPlus (2020) also contained the fault model of boiler fouling as follows:

$$Q_{boiler,f} = Q_{boiler,ff} \times F_{boiler} \quad (2.57)$$

$$Eff_{boiler,f} = Eff_{boiler,ff} \times F_{boiler} \quad (2.58)$$

where $Q_{boiler,f}$ is the nominal capacity of the fouled boiler (kW); $Q_{boiler,ff}$ is the nominal capacity of the fault free boiler (kW); $Eff_{boiler,f}$ is the nominal thermal efficiency of the fouled boiler; $Eff_{boiler,ff}$ is the nominal thermal efficiency of the fault free boiler; F_{boiler} is the reduction factor for the capacity and thermal efficiency of the boiler due to fouling.

(5) Evaporative cooler fouling

The evaporative cooler fouling only happens at the indirect wet-coil evaporator coolers due to the accumulation of dust from the air. It will degrade the energy efficiency of the evaporative cooler. The fault model of the evaporative cooler fouling can be found in EnergyPlus (2020):

$$Eff_{EvapCooler,f} = Eff_{EvapCooler,ff} \times F_{EvapCooler} \quad (2.59)$$

where $Eff_{EvapCooler,f}$ is the indirect stage efficiency of the fouled evaporator cooler; $Eff_{EvapCooler,ff}$ is the indirect stage efficiency of the fault free evaporator cooler; $F_{EvapCooler}$ is the reduction factor for the indirect stage efficiency of the evaporator cooler.

(6) Dirty air filters

Air filters are typically located at the inlets of OA or inside the duct in order to prevent the contaminants from entering the air loop system. The dirty air filters are caused by the

accumulation of dust or other blockages at the filter. It will increase the airflow resistance and thus alter the air system curve. In detail, it will increase the fan pressure, fan power consumption and/or the fan outlet air enthalpy. Besides, it may result in a decrease in the fan airflow rate and hence influences the operation of other components in the air system. The dirty air filters are modeled by introducing faulty fan curves and pressure rise schedules (Zhang and Hong, 2017). Three cases of fan operations with dirty air filters are considered by Zhang and Hong (2017): (a) The variable speed fan operating at a higher speed is able to keep the desired airflow rate; (b) The variable speed fan is unable to keep the desired airflow rate even with a higher speed; (c) The constant speed fan is unable to keep the design airflow rate.

(a) The variable speed fan operating at a higher speed is able to keep the desired airflow rate

In this case, the required airflow rate is maintained by a variable speed fan at a higher speed (from point A to point B in Fig. 2.8). As a result, the fan pressure rise, fan power and power entering the air are increased and the specific enthalpy of the air at the fan outlet is affected. These fault impacts are modeled by the following equations:

$$f_f = m/m_d \quad (2.60)$$

$$f_{pl} = c_1 + c_2 f_f + c_3 f_f^2 + c_4 f_f^3 + c_5 f_f^4 \quad (2.61)$$

$$Q_{tot} = f_{pl} m_d \Delta P_f / (e_{tot} \rho_{air}) \quad (2.62)$$

$$Q_{shaft} = e_{motor} Q_{fan} \quad (2.63)$$

$$Q_{toair} = Q_{shaft} + (Q_{fan} - Q_{shaft}) f_{toair} \quad (2.64)$$

$$h_{out} = h_{in} + Q_{toair}/m \quad (2.65)$$

where f_f is the flow fraction, m is the actual mass flow rate (kg/s), m_d is the design mass flow rate (kg/s), f_{pl} is the part load factor, $c_1 \sim c_5$ are the coefficients for the fan performance curve, Q_{fan} is the fan power (W), ΔP_f is the faulty fan pressure rise (Pa), e_{tot} is the fan total efficiency, ρ_{air} is the air density (kg/m³), Q_{shaft} is the fan shaft power (W), e_{motor} is the motor efficiency, Q_{toair} is the power entering into the air (W), f_{toair} is the ratio of motor power entering into the air, h_{out} is the specific enthalpy of the air at the fan outlet (kJ/kg), h_{in} is the specific enthalpy of the air at the fan inlet (kJ/kg).

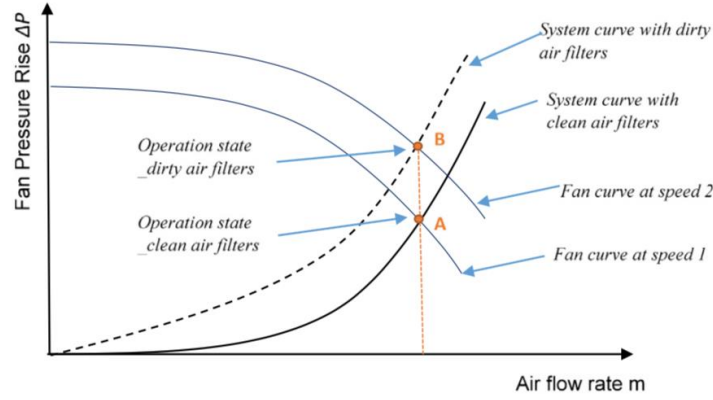


Figure 2.8 The influence of dirty air filter on variable speed fan performance with the airflow rate maintained (Zhang and Hong, 2017).

(b) *The variable speed fan is unable to keep the desired airflow rate even with a higher speed*
 In this case, the actual airflow rate m is decreased to the design airflow rate $m_{design,f}$ and the fan pressure is increased (from point A to point B in Fig. 2.9). Similar to the case (a), the fan power, power entering the air and the specific enthalpy of the air at the fan outlet are all influenced. The fault model can follow the Eqn. (2.60)-(2.65) but the flow fraction $f_{flow,f}$ is changed to 1.

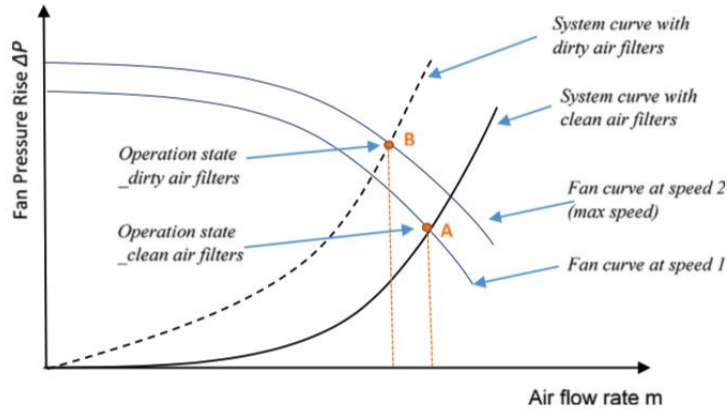


Figure 2.9 The influence of dirty air filter on variable speed fan performance with the reduced airflow rate (Zhang and Hong, 2017).

(c) *The constant speed fan is unable to keep the design airflow rate.*

The case (c) has the similar fault impacts as the case (b) but the actual air flow rate m is decreased to an airflow rate at operation state with dirty air filters m_f (from point A to point B in Fig. 2.10). Hence, the Eqn. (2.62) and Eqn. (2.65) are modified to the following expressions:

$$Q_{tot,f} = m_{df} \Delta P_f / (e_{tot} \rho_{air}) \quad (2.66)$$

$$h_{out,f} = h_{in} + Q_{toair,f} / m_f \quad (2.67)$$

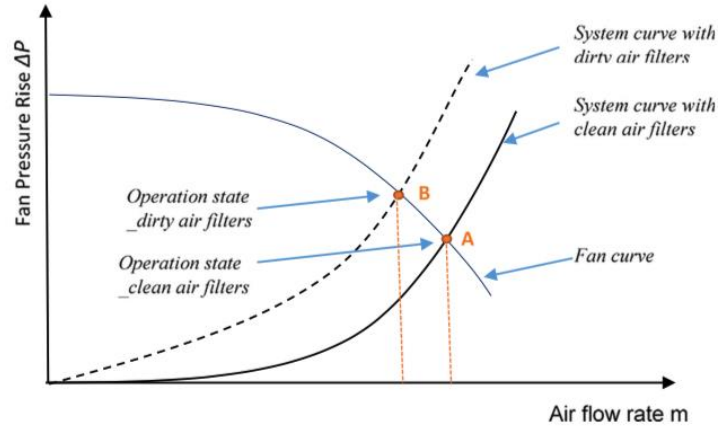


Figure 2.10 The influence of dirty air filter on constant speed fan performance (Zhang and Hong, 2017).

2.3 Fault Impact Analysis Review

Fault simulation outputs are crucial for evaluating fault impacts, in terms of HVAC system energy consumptions and occupants' thermal comfort. The current fault impact assessment studies in the field can be mainly divided into two groups: (1) single fault impact analysis; (2) multiple faults impact analysis. The single fault impact analysis evaluates the impacts of each fault on system energy usages and indoor space conditions separately and the relevant work will be reviewed in subsection 2.3.1. The multiple fault impact analysis investigates the interactions among multiple faults and compares combined effects of multiple faults combinations with the individual effect of each fault involved in each combination. The corresponding studies for multiple fault analysis will be discussed in subsection 2.3.2. Based on the literature search method indicated in subsection 2.1.1, 46 fault impact analysis studies have been reviewed in this thesis. Among these studies, 32 studies only considered single fault impact and 14 studies took into account the impacts of both single and multiple faults.

2.3.1 Single Fault Impact Analysis

32 studies in total investigated the impacts of single faults for a variety of HVAC systems and components. 20 of the 32 literature employed simulation platforms to evaluate the impacts of single faults and 12 of the 32 literature conducted experiments for assessing single fault's impacts, as presented in Figure 2.11. The detailed descriptions, including target system or component type, fault types and fault impact indicators involved in each study, climate zone

used for simulations, and simulation platform type, of each literature using fault simulation approach were provided in Table 2.10. The target system or component type, fault types and fault impact indicators involved in each study for each work based on fault assessment experiment were described in Table 2.11.

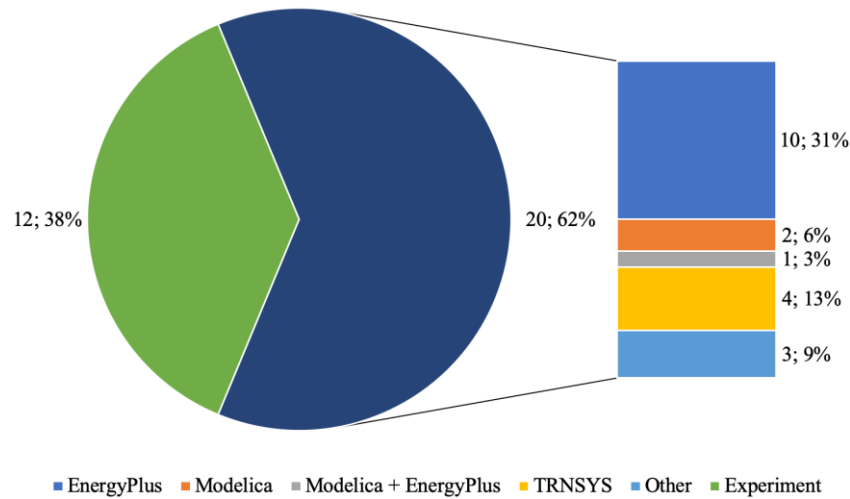


Figure 2.11 Distributions of the papers about single fault impact analysis among various simulation platforms and the experiment approach.

2.3.1.1 Simulation-based Studies

With respect to single fault impact analysis, most studies implemented fault simulations which leverage the flexibility of fault types and climatic conditions provided by fault simulation platforms. As shown in Figure 2.11, EnergyPlus is the most popular BPS platform used for assessing single fault impacts due to its robust and feasible fault modeling capacity. In particular, sensor and actuator faults are the most common type of faults evaluated using EnergyPlus in which incorporates embedded fault objects with respect to these faults. Kim et al. (2019) evaluated the impacts of a series of faults on the readings of sensors and energy meters under the mixed humid climate. Basarkar et al. (2011) simulated the impacts of four single faults, including room air temperature sensor bias and OA damper leakage, on the VAV system cooling, heating and pump energy under the mixed humid and cold climates. Zhang and Hong (2017) investigated the effect of sensor biases, such as economizer sensor bias and thermostat bias, on HVAC system heating and cooling energy and occupant thermal comfort under the hot humid, marine and cold climates. Gunay et al. (2018) explored the reduction potential for energy usages and thermal discomfort levels by prompt maintenance of

AHU faults, incorporating OA damper stuck and OA temperature sensor bias, and zone-level faults, containing indoor air temperature sensor bias and VAV terminal damper stuck under the cold climate. Lee and Yik (2010) presented the influences of sensor biases, damper and valve stuck, and valve leakage on VAV system energy consumptions under the humid subtropical climate. Li et al. (2023a) investigated the influences of a group of sensor incipient faults, such as VAV terminal SA temperature and flow rate sensor biases, on system total and end energy usages under the mixed humid climate. Li et al. (2022a) assessed the impact of thermostat sensor bias on system energy end usages, thermal comfort levels and system COP for three difference HVAC systems, including ground source heat pump system, VRF system, and chiller and boiler system, under the humid subtropical climate.

Beyond the assessments of sensor and actuator single faults through EnergyPlus simulations, Cheung and Braun (2016) assessed the influences of four refrigerant-cycle related faults on energy consumptions of chiller and cooling tower under the hot humid, and hot dry climates. Mattera et al. (2020) explored the effects of heat exchanger efficiency and heating effect degradation, fan efficiency degradation and duct fouling on heating and fan energy consumptions under the temperate oceanic climate. Gunay et al. (2023) simulated the heating and cooling energy usages under HVAC control system faults in the cold climate.

In addition, TRNSYS and Modelica are the other commonly-used platforms for single fault impact evaluations. Ji et al. (2023) applied TRNSYS into the impact analyses of AHU sensor bias and valve faults as well as indoor air temperature sensor bias on the chilled water pump energy consumption and indoor air temperature under the climate of Hongkong. Gao et al. (2016) carried out TRNSYS simulations for exploring the influence of AHU coil and plate heat exchanger performance degradations on chilled pump energy usage under the humid subtropical climate. Otto et al. (2012) evaluated the impacts of a group of VAV system faults, including VAV damper stuck, AHU OA damper stuck, etc., on whole building energy consumption by implementing TRNSYS simulations. Khire and Trcka (2013) also conducted TRNSYS simulations for studying the effect of VAV and AHU faults on whole building energy consumption.

As for the studies using Modelica simulations, Lu et al. (2021) employed Modelica to analyse the effects of 359 fault cases, including sensor biases, damper stuck and leakage, on various

impact indicators, including energy usages, thermal comfort levels, operational cost, ventilation metrics, etc., under the cold climate. Huang et al. (2021) integrated Modelica with EnergyPlus simulations for the annual energy evaluation of chiller and boiler plants and thermal unmet hours caused by 13 control-associated faults containing chilled and condensor water temperature and pressure sensor biases. Marigo et al. (2023) explored the impacts of AHU faults on thermal and electrical energy consumptions, thermal comfort levels and indoor air quality under the marine climate using Modelica simulations.

Apart from these fault simulation platforms, MATLAB, HVACSIM+ and IDA Indoor Climate and Energy (Björnsell, 1999) have also been used in a few studies. Pelella, Viscito and Mauro (2022a) simulated the impacts of refrigerant leakage, condenser and evaporator fouling on seasonal COP and total equivalent warming impact using MATLAB under three different climate zones, including Mediterranean, hot humid and humid subtropical climate. Chen et al. (2022) used HVACSIM+ to assess the effects of fan coil sensor and actuator faults and control faults on energy consumption and zone thermal comfort under the cold climate. Andersen et al. (2020) simulated demand-controlled ventilation system sensor biases using IDA Indoor Climate and Energy, and analysed their impacts on system energy end usages, thermal comfort and indoor air quality.

In summary, the majority of simulation-based single fault impact analysis studies focused on the fault impacts on system energy total and end usages, and thermal comfort and only a few of work incorporated other impact indicators, such as indoor air quality level and COP. In addition, the fault categories considered in these studies emphasized on control, sensor, actuator faults, and other VAV and AHU faults. This finding is correlated to the target system or component investigated in each study as the most work employed VAV, AHU, RTU or a combination of some of them as the case studies. In terms of simulation platforms used for evaluating single fault impacts, EnergyPlus, TRNSYS and Modelica are the mainstreams. As for the climatic data used for fault simulations, diverse climate zones were considered while only several studies explored fault impacts under different climate zones. However, none of the studies took into account the influence of changing climatic conditions from current climate scenarios to future climate scenarios on single fault impacts.

Table 2.10 Existing studies for single fault impact analysis using simulations

Ref.	System/ component type	Fault types	Fault impact indicators	Climate zone	Simulation software
Kim et al. (2019)	VAV and RTU	Control, sensor, RTU and other faults	Heating, cooling, fan, system total energy usages	Mixed humid	EnergyPlus
Basarkar et al. (2011)	VAV	Chiller, actuator, sensor faults	Cooling, heating, pump energy usages	Mixed humid and cold	EnergyPlus
Zhang and Hong (2017)	VAV	Sensor, actuator, other faults	Heating, cooling, fan energy usages and thermal comfort	Hot humid, marine and cold	EnergyPlus
Gunay et al. (2018)	VAV and AHU	AHU/VAV sensor and actuator faults	Energy use intensity and thermal comfort	Cold	EnergyPlus
Lee and Yik (2010)	VAV, AHU and chiller	AHU/VAV sensor and actuator faults	Chiller, fan, pump, heating, system total energy usages	Humid subtropical	EnergyPlus
Li et al. (2023a)	VAV and RTU	Sensor faults	Site, cooling, heating and fan energy usages	Mixed humid	EnergyPlus
Li et al. (2022a)	Heat pump/VRF/chiller and boiler	Sensor faults	Heating, cooling, pump, fan, heat recovery energy usages, system COP, thermal comfort	Humid subtropical	EnergyPlus
Cheung and Braun (2016)	VAV/CAV and chiller	Chiller faults	Chiller, cooling tower, building total electricity usages	Hot humid and hot dry	EnergyPlus
Mattera et al. (2020)	Ventilation unit	AHU and other faults	Heating and fan energy usage	N/A	EnergyPlus
Gunay et al. (2023)	VAV and AHU	Control faults	Heating and cooling energy usages	Cold	EnergyPlus
Huang et al. (2021)	Chiller and boiler plants	Control, sensor and actuator faults	Chiller and boiler annual energy usage and thermal discomfort	N/A	Modelica and EnergyPlus

Table 2.10 Continue

Lu et al. (2021)	VAV	Control, sensor, AHU, chiller and VAV faults	Operational cost, source/site energy, thermal comfort and indoor air quality	Cold	Modelica
Ji et al. (2023)	AHU and chiller	Sensor and actuator faults	Chilled water pump energy usage	Humid subtropical	TRNSYS
Gao et al. (2016)	AHU	AHU faults	Chilled water pump energy consumption	Humid subtropical	TRNSYS
Otto et al. (2012)	VAV	VAV and AHU faults	Building energy consumption	N/A	TRNSYS
Khire and Trcka (2013)	VAV and AHU	VAV and AHU faults	Building energy consumption	N/A	TRNSYS
Marigo et al. (2023)	CAV and AHU	AHU and other faults	Thermal and electrical energy usages, thermal comfort and indoor air quality	Marine	Modelica
Pelella, Viscito and Mauro (2022a)	Heat pump	Split air-conditioner faults	Seasonal COP and total equivalent warming impact	Mediterranean, hot humid and humid subtropical	MATLAB
Chen et al. (2022)	Fan coil unit	Control, sensor, actuator and other faults	System energy consumption and zone comfort	Cold	HVACSIM+
Andersen et al. (2020)	Demand-controlled ventilation system	Sensor faults	Fan, pump, heating and cooling energy usages, thermal comfort, indoor air quality	Humid continental	IDA Indoor Climate and Energy

2.3.1.2 Experiment-based Studies

Apart from fault simulations, laboratory or field experiment is another approach for single fault impact analysis. As listed in Table 2.11, the majority of experiment-based fault impact analysis studies focused on the impacts of refrigerant-cycle related faults occurring in split system air conditioner (Hu and Yuill, 2022a), heat pump (Hu et al., 2021a), RTU (Kim and Lee, 2021) and VRF system (Gou et al., 2023). Hu and Yuill (2022a) evaluated the influences

of refrigerant undercharge and overcharge, improper evaporator airflow rate, liquid line restriction, and non-condensable gas in the refrigerant on cooling capacity and COP for a split air conditioner. Similarly, Hu et al. (2021a) investigated the impacts of same faults on cooling capacity and COP for a heat pump. Guo, Li and Li (2023) also studied the impacts of a variety of faults, including refrigerant leakage and undercharge, on system operating power for a air-source heat pump. Gou et al. (2023) assessed the effects of refrigerant overcharge and undercharge, indoor and outdoor unit fouling, indoor unit expansion valve failure, etc. on heating capacity, total system power, and COP for a VRF system. Bellanco et al. (2022) explored the impacts of evaporator fouling, compressor valve leakage, liquid line restriction and refrigerant overcharge on COP, compressor power consumption and condenser heat duty for a heat pump. Kim and Lee (2021) analysed the impacts of condenser fouling, refrigerant overcharge and undercharge on cooling capacity, COP, seasonal energy efficiency ratio, annual operating cost for a RTU. Mehrabi and Yuill (2018) assessed the influences of liquid line restriction, compressor valve leakage, non-condensable gas in the refrigerant, evaporator airflow reduction and condenser heat transfer reduction on COP, system heating and cooling capacity for both heat pump and split air conditioner. Du, Domanski and Payne (2016) conducted an experimental assessment for common faults of two vapour compression systems, split system and RTU. The impacts of the faults, containing refrigerant charge faults, evaporator low indoor airflow, compressor valve leakage, liquid line restriction, etc., on cooling capacity, COP and sensible heat ratio were quantified. Catrini and Piacentino (2023) implemented an experimental evaluation for the evaporator fouling's impact on cooling capacity and RTU power consumption for a RTU. Qureshi and Zubair (2014) carried out a fault impact analysis of condenser fouling and evaluated its impact on compressor power consumption and COP for a split system.

Apart from the impact assessments for refrigerant-cycle related faults, the influences of AHU coil and humidifier valve faults (Rosato et al., 2022a), and AHU sensor and fan faults (Rosato et al., 2022b) on thermal and hygrometric indoor comfort time were explored using experiments.

To conclude, most experiment-based single fault impact analysis work focused on vapour compression systems, including heat pumps, split air conditioners, VRF system and RTUs. As

a result, most work assessed the effects of refrigerant-cycle related faults on COP, heating or cooling capacity and system or component power consumption.

Table 2.11 Existing studies for single fault impact analysis using lab or field experiments

Ref.	System/ component type	Fault types	Fault impact indicators
Hu and Yuill (2022a)	Split air conditioner	Refrigerant-cycle related faults	Cooling capacity and COP
Hu et al. (2021a)	Heat pump	Refrigerant-cycle related faults	Cooling capacity and COP
Guo, Li and Li (2023)	Heat pump	Split air-conditioner faults including refrigerant-cycle related faults	System operating power
Gou et al. (2023)	VRF	Refrigerant-cycle related faults, actuator fault and other faults	Heating capacity, total system power and COP
Catrini and Piacentino (2023)	RTU	Evaporator fouling	Cooling capacity and RTU power consumption
Rosato et al. (2022a)	CAV and AHU	AHU actuator faults	Thermal/hygrometric indoor comfort time
Rosato et al. (2022b)	CAV and AHU	AHU sensor and actuator faults	Thermal/hygrometric indoor comfort time
Bellanco et al. (2022)	Heat pump	Refrigerant-cycle related faults	COP, compressor power consumption and condenser heat duty
Kim and Lee (2021)	RTU	Refrigerant-cycle related faults	Cooling capacity, COP, seasonal energy efficiency ratio, annual operating cost
Mehrabi and Yuill (2018)	Heat pump and split air conditioner	Refrigerant-cycle related faults	System heating/cooling capacity and COP
Du, Domanski and Payne (2016)	Split system and RTU	Refrigerant-cycle related faults	Cooling capacity, COP and sensible heat ratio
Qureshi and Zubair (2014)	Split system	Condenser fouling	Compressor power consumption and COP

2.3.2 Multiple Faults Impact Analysis

Overall, 14 existing studies assessed the impacts of HVAC multiple faults. Among the 14 studies, 10 studies applied fault simulations for evaluating the impacts of multiple faults while only four studies analysed multiple faults' impacts based on fault experiments (Fig. 2.12). The detailed descriptions, including target system or component type, fault types and fault impact indicators involved in each study, climate zone used for simulations, and simulation platform type, of each literature using fault simulation approach were provided in Table 2.12. The target system or component type, fault types and fault impact indicators involved in each study for each work based on fault assessment experiment were described in Table 2.13.

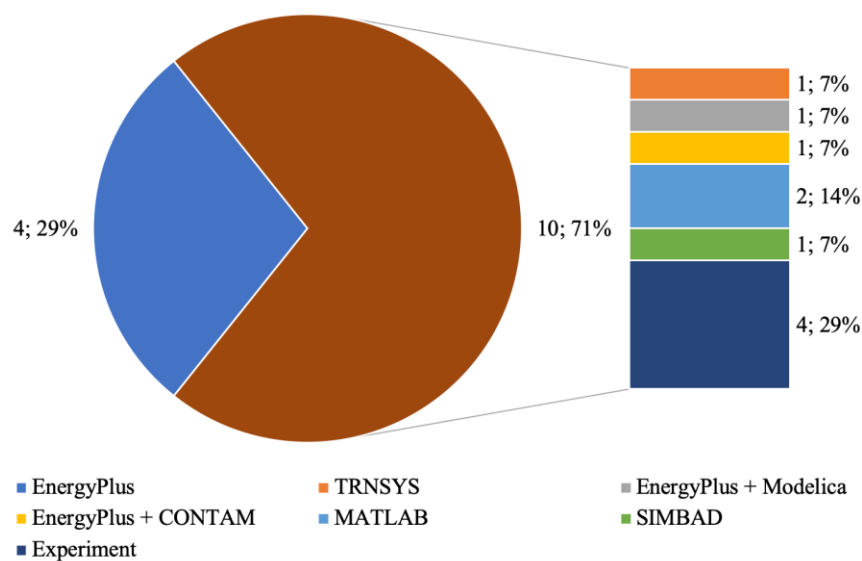


Figure 2.12 Distributions of the papers about multiple faults impact analysis among various simulation platforms and the experiment approach.

2.3.2.1 Simulation-based Studies

As for the studies concerning both single and multiple faults, fault simulation is still the mainstream approach for assessing the impacts of fault combinations. EnergyPlus is the most widely-used tool for assessing multiple faults' impacts under different climate zones. Regarding the fault impact analysis work using EnergyPlus, Wang and Hong (2013) assessed the influences of thirteen single faults including temperature sensor offset, OA damper leakage and stuck, etc., and two multiple faults combinations on system energy end usages for a VAV system under the cold climate. Li and O'Neill (2019) evaluated the effects of 129 fault scenarios, including both single and multiple faults cases, on VAV system energy usage and occupant thermal comfort under the mixed humid, hot humid, marine and cold climates.

Winkler et al. (2020) investigated the national-scale influence of two installation faults for residential heat pumps and air conditioners on the annual increases in energy and utility costs under the climate zones of very hot, hot, warm, mixed, cool, cold, and very cold. Yoon et al. (2019) studied the influences of AHU and chiller sensors on system end-use energy, COP and thermal comfort under the marine, cool dry, hot humid, cold and mixed humid climates. In addition, some works coupled EnergyPlus with other building simulation tools to facilitate fault impact analysis. Lu et al. (2020) carried out coupled EnergyPlus and CONTAM simulations for assessing the impacts on HVAC energy usage and ventilation performance induced by individual and multiple faults of CO₂ and airflow sensors for a demand-controlled ventilation system. Chen, Huang and Vrabie (2018) integrated Modelica modeling with EnergyPlus simulations to quantify the effects of OA and SA temperature sensor biases, cooling coil valve stuck and the combination of them on the fan, heating and total electric energy usages under the cold climate.

Apart from EnergyPlus simulations, Verhelst et al. (2017) used TRNSYS simulations to estimate the economic impact of sensor and actuator faults for a concrete core activated office building under marine climate. Besides, MATLAB is also a popular tool to model single and multiple faults' impacts. Mauro, Pelella and Viscito (2023) used MATLAB to model the fault impacts of refrigerant leakage, condenser fouling and evaporator fouling on heating capacity and COP of an air-to-air heat pump under the climate of mixed humid, humid continental and marine. Pelella, Viscito and Mauro (2022b) evaluated the effects of refrigerant leakage, condenser and evaporator fouling on cooling capacity and COP for a residential heat pump. Ginestet, Marchio and Morisot (2008) employed the SIMBAD Simulink library to explore the effects of individual faults from valve, damper and sensors, and combinations of them on heating coil energy usage and indoor air quality.

To recap, the existing simulation-based multiple faults' impact assessment work is inclined to investigate the fault impacts for VAV systems and AHUs. However, compared with those simulation-based single fault impact analysis work, there are some work exploring the impacts of refrigerant-cycle related faults on COP and heating or cooling capacity for heat pumps using fault simulations. Regarding simulation platforms, EnergyPlus is still the top candidate for carrying out fault simulations. As for the work mentioning the climate zone specified for fault simulations, comparisons for fault impacts under multiple climate zones are

popular. But there is still a lack of work considering the effect of changing climate conditions from current climate scenarios to future climate scenarios on multiple faults' impacts.

Table 2.12 Existing studies for multiple faults impact analysis using simulations.

Ref.	System/ component type	Fault types	Fault impact indicators	Climate zone	Simulation software
Wang and Hong (2013)	VAV and chiller	VAV and chiller faults	Heating, cooling, fan, pump and cooling tower energy usages	Cold	EnergyPlus
Li and O'Neill (2019)	VAV and RTU	VAV and RTU faults	System energy consumption and thermal comfort	Mixed humid, hot humid, marine and cold	EnergyPlus
Winkler et al. (2020)	Central air conditioners and air-source heat pumps	Refrigerant-cycle related faults	Site energy usage and utility cost	Very hot, hot, warm, mixed, cool, cold, very cold	EnergyPlus
Yoon et al. (2019)	VAV and chiller	Sensor faults	Heating, cooling, pump, fan energy usages, COP and thermal comfort	Marine, cool dry, hot humid, cold, mixed-humid	EnergyPlus
Lu et al. (2020)	VAV and AHU	Sensor faults	System energy and indoor air quality	N/A	EnergyPlus and CONTAM
Chen, Huang and Vrabie (2018)	VAV	Sensor and actuator faults	Fan, system total electric and gas heating energy	Cold	Modelica and EnergyPlus
Verhelst et al. (2017)	CAV, AHU and chiller	Sensor and actuator faults	Energy cost, thermal discomfort cost and maintenance cost	Marine	TRNSYS
Mauro, Pelella and Viscito (2023)	Air-to-air heat pump	Refrigerant-cycle related faults	Heating capacity and COP	Mixed humid, humid continental and marine	MATLAB
Ginestet, Marchio and Morisot (2008)	AHU	Sensor and actuator faults	Heating energy usage and indoor air quality	N/A	Simulink SIMBAD library
Pelella, Viscito and Mauro (2022b)	Residential heat pump	Refrigerant-cycle related faults	Cooling capacity and COP	N/A	MATLAB

2.3.2.2 Experimental-based Studies

In addition to these studies assessed by fault simulations, Zhou et al. (2022) carried out an experimental investigation into the effects of single and multiple faults of refrigerant-related faults, including refrigerant charge faults, indoor and outdoor unit fouling and non-condensable gas in the refrigerant, on system power, COP and cooling capacity for a VRF system. Similarly, Hu et al. (2021b) experimented to explore the impacts of single and multiple faults of refrigerant-related faults, consisting of refrigerant charge faults, improper evaporator airflow rate, non-condensable gas in the refrigerant and liquid line restriction, on COP and cooling capacity for a split system heat pump. Hu, Chen and Yuill (2023) assessed the effects of single and multiple faults same to Hu et al.'s (2021b) on the temperature difference across the liquid line for both split air-conditioner and heat pump. Hu and Yuill (2022b) evaluated the influences of the same single and multiple faults as Hu et al.'s (2021b) on cooling capacity and COP for a split air-conditioner. To sum up, there is only a limited number of studies investigating the multiple faults' impacts based on fault experiments, and all of them emphasized on the effects of refrigerant-cycle related faults for vapor compression systems containing VRF system, heat pumps and split air conditioners.

Table 2.13 Existing studies for multiple faults impact analysis using lab or field experiments

Ref.	System/ component type	Fault types	Fault impact indicators
Zhou et al. (2022)	VRF	Refrigerant-cycle related faults	Cooling capacity, system power, and COP
Hu et al. (2021b)	Heat pump	Refrigerant-cycle related faults	Cooling capacity and COP
Hu, Chen and Yuill (2023)	Split air conditioner and heat pump	Refrigerant-cycle related faults	Temperature difference across the insulated liquid line
Hu and Yuill (2022b)	Split air conditioner	Refrigerant-cycle related faults	Cooling capacity and COP

2.3.3 Summary and Research Gaps

The target systems or components for simulation-based fault impact analysis work were typically VAV systems, AHUs and RTUs, and therefore these work focused on the effects of control, sensor, actuator faults, and other VAV and AHU faults on system energy total and end usages, and thermal comfort. Besides, EnergyPlus is the most widely-used simulation

platform for assessing both single and multiple faults' impacts. Most simulation-based single fault impact analyses were implemented under a certain climate zone and only a few studies considered single fault impact under multiple climate zones. In contrast, the majority of multiple fault impact analyses were carried out under multiple climate zones.

On the other hand, the experiment-based fault impact analysis studies mainly targeted at refrigerant-cycle related faults for vapor compression systems, including heat pumps, split air conditioners, VRF system and RTUs. The assessed fault impacts consisted of COP, heating or cooling capacity and system or component power consumption.

Based on the exhaustive search and comprehensive review of literature, to the author's best knowledge, there are still several research gaps in the field:

1. Apart from the work investigating HVAC fault impacts under climatic conditions of different climate zones, there is a lack of work exploring the fault impacts under changing climatic conditions from current to future climate scenarios, and across different seasons.
2. There is still a lack of literature investigating the interactions among a wider array of multiple faults, such as the interactions among sensor and actuator faults.

2.4 Data-driven Fault Detection and Diagnosis (FDD) Review

Apart from fault impact analysis studies, another major aspect in HVAC operational fault research is FDD for HVAC systems. With the advancement of data science techniques and the prevalence of sensors and BMSs in modern buildings, data-driven methods lead to the latest trend in FDD for HVAC systems. Various review studies focusing on data-driven FDD methods for HVAC systems have emerged in recent years. Zhao et al. (2019) introduced data-driven FDD methods in terms of three categories, i.e., classification-based, regression-based and unsupervised-based methods and discussed the applications of different category of methods into existing studies. Mirnaghi and Haghighat (2020) divided data-driven FDD methods into supervised, semi-supervised and unsupervised methods, and explained the advantages and challenge for each group of FDD methods. Chen et al. (2022) reviewed data-driven FDD methods using similar categories to previous work, i.e., supervised classification, unsupervised, and regression and statistics-based methods. In addition, Chen et al. (2023) provided an overview of the essential procedures for implementing data-driven FDD methods, and discussed the target systems and data sources for existing data-driven FDD studies.

Moreover, Zhang, Saeed and Sadeghian (2023) reviewed the work incorporating a state-of-the-art category of data-driven FDD methods, deep learning-based methods, and carried out a meta analysis for the robustness of these methods.

Based on these review studies, this section will firstly explain the basic concepts of data-driven FDD in subsection 2.4.1. Subsection 2.4.2 will discuss the method perspective of data-driven FDD by supervised, semi-supervised and unsupervised methods, and further introduce the sub-categories and popular candidates under each sub-category. Subsection 2.4.3 will list the practical challenges for implementing data-driven FDD methods and potential solutions to these challenges. Subsection 2.4.4 will conclude the research gaps still existing in the field.

2.4.1 Data-driven FDD Fundamentals

FDD is regarded as a vital process for ensuring the steady and safe operation for an industrial system (Shi and O'Brien, 2019). Conventional FDD methods, such as model-based and knowledge-driven methods, rely on deep domain knowledge, and thus require heavy experts' efforts into the development of these methods. However, data-driven FDD methods require no or minimal prior system knowledge and automatically capture the underlying fault behaviours through relevant features from HVAC system operational data. Therefore, the FDD performances of data-driven methods are significantly affected by the amount of training data and the data quality. Under the term of data-driven FDD, several basic FDD concepts are introduced as follows:

- (1) Fault: a fault is an abnormal system operation behaviour that leads to the malfunctions or performance degradations of the system.
- (2) Fault detection: the detection of faults occurring during system operation. The output of fault detection is binary, i.e., faulty or normal.
- (3) Fault diagnosis: the identification of a fault to its corresponding fault type. The output of fault diagnosis can be multiple fault classes.

2.4.2 Data-driven FDD Method Perspective

According to Mirnaghi and Haghighat (2020), data-driven FDD methods can be categorised into supervised, semi-supervised and unsupervised methods. This subsection will follow this classification and review the existing studies employing the popular specific methods under each category.

2.4.2.1 Supervised Learning-based FDD Methods

Supervised learning is a group of machine learning methods widely-used in data-driven FDD methods for HVAC systems. Supervised learning refers to the process of learning the underlying mechanisms of input data to produce outputs that match the actual data labels.

Through capturing fault behaviours from fault data fed into a supervised learning FDD model, the FDD model is capable of detecting and diagnosing the faults based on its unseen data.

Supervised-learning FDD methods can be grouped into two major categories: classification and regression methods.

2.4.2.1.1 Supervised Classification Methods

Classification is a machine learning task that identifies the class type each data sample belongs to. Under the FDD field, fault detection is treated as a binary classification task to determine whether the patterns of HVAC operational data are similar to those of fault data or not. Fault diagnosis is the process of identifying the exact fault class based on HVAC faulty operational data and is also treated as a multi-class classification task.

Support vector machine (SVM) is a popular classification algorithm used for both fault detection and fault diagnosis tasks. As shown in Figure 2.13, the algorithm searches for the optimal hyperplane to separate the data samples of a fault class from those of others at the maximum margin. Liang and Du (2007) proposed a multi-layer SVM-based classifier to detect and identify chiller faults and the classifier was tested to be robust in FDD for return air damper stuck, cooling coil fouling and supply fan speed degradation. Namburu et al. (2007) compared the FDD performance of three classification algorithms, SVM, principal component analysis (PCA) and partial least squares, and SVM indicated the highest fault classification rates among the three algorithms. Han et al. (2011a) adopted the genetic algorithm (GA)-based feature selection wrapper with SVM to diagnose seven typical chiller faults and the proposed method achieved a classification correct rate of 99.5%. Han et al. (2011b) implemented SVM with multi-label technique for the detection and diagnosis of multiple simultaneous chiller faults in another study. The developed multi-label SVM method resulted in a classification correct rate over 99.9%. Sun et al. (2016) integrated the method of SVM with wavelet de-noising (WD) into max-relevance and min-redundancy (mRMR) algorithm to diagnose refrigerant charge faults for VRF systems. WD was used to preprocess the measured VRF system data and mRMR was adopted to select the most important features and remove

the redundant variables for training the SVM model. The WD-mRMR-SVM model with fewer features was demonstrated to be capable of reducing the computational time with loss of fault diagnosis accuracy of 2.14%. Ebrahimifakhar, Kabirikopaei and Yuill (2020) combined a synthetic data generation technique, synthetic minority over-sampling technique, with nine classifiers, including logistic regression, linear discriminant analysis, quadratic discriminant analysis, k-nearest neighbors, bagging, random forest (RF), AdaBoost, XGBoost and SVM, to detect and identify seven typical RTU faults. The nine combinations were tested on a simulated RTU fault dataset and SVM indicated the highest classification accuracy of 96.2%.

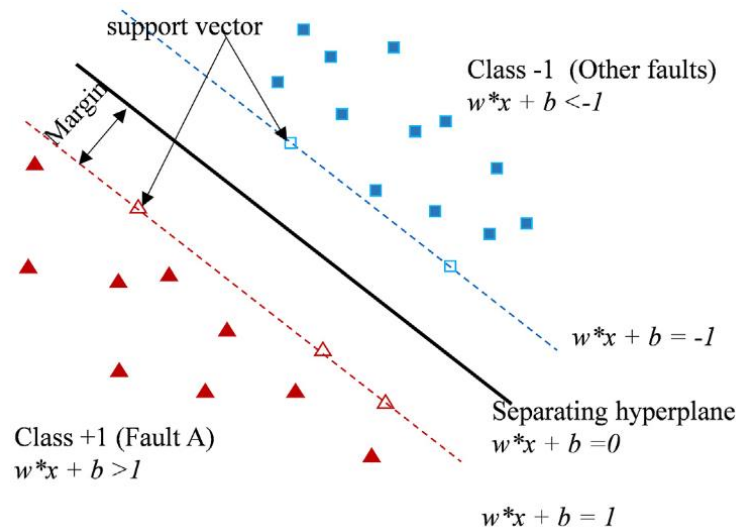


Figure 2.13 Diagram of a SVM-based classification. (Zhao et al., 2019)

In addition, decision tree-based methods have also been widely used in a number of data-driven FDD work. Yan et al. (2016) applied a decision tree algorithm, i.e., classification and regression tree (CART) algorithm (Fig. 2.14) for diagnosing AHU faults. The FDD method was tested under ASHRAE research project 1312 (RP-1312) (Wen and Li, 2011) and indicated an excellent FDD performance with an average F1 score of 0.97 (F1 score is a metric suitable for evaluating machine learning model performance under class-imbalanced data distributions). Liu et al. (2016) implemented the CART algorithm to diagnose refrigerant undercharge and overcharge faults for VRF systems. The validation results showed that the refrigerant undercharge fault was accurately identified by the CART algorithm with a diagnosis accuracy of 93.8%, but the amount of refrigerant overcharge fault data was insufficient to be identified by the algorithm. Using the same algorithm, Li et al. (2018) introduced an enhanced CART-based FDD method to detect and identify condenser air-side

fouling, refrigerant overcharge and undercharge faults for VRF systems. The enhanced method integrated the CART algorithm with virtual sensor-based fault indicators, and demonstrated better fault diagnosis accuracy, compared with other three tree-based algorithms, i.e., vanilla CART, RF and generalized boosted regression. Besides, another decision tree-based algorithm, C5.0 decision tree, was performed with expert knowledge by Yu et al. (2018) to identify refrigerant charge faults for VRF systems. Compared with the C5.0 decision tree-based method without expert knowledge, the expert-modified C5.0 decision tree led to an increase of 10% in fault diagnosis rate. Apart from these, RF is an ensemble classification model incorporating multiple decision trees, which demonstrates higher FDD performance and generalization ability than single decision tree model. Wang, Feng and Liu (2021) proposed a two-layer RF-based FDD method for VAV terminals and the proposed method was capable of effectively identifying both single and multiple faults of VAV terminals. Zhu et al. (2019) developed a combined decoupling features and RF classifier method and the test results showed that the FDD model achieved correct diagnosis rate of 94.2% under unseen system operation conditions.

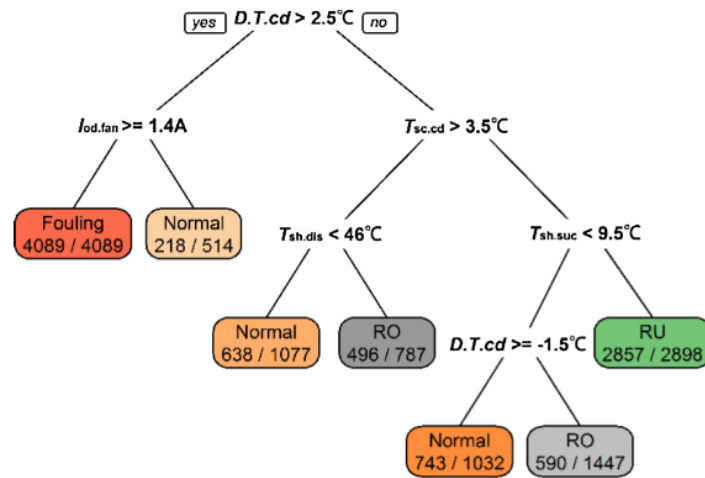


Figure 2.14 An example of a decision tree algorithm, i.e., CART. (Li et al., 2018)

Beyond these methods, artificial neural network (ANN) has demonstrated its robust power in FDD for HVAC systems. As illustrated in Figure 2.15, an ANN contains a input layer, one or multiple hidden layers and a output layer. Each layer consists of several neurons that are connected to other neurons in the adjacent layers. The connection of a neuron to other neuron in the adjacent layers represents a weight which is constantly adjusted during ANN model training, in order to minimize the losses between the model outputs and actual targets. Most

studies integrated ANN with other statistics-based techniques, such as wavelet analysis (Du, Jin and Yang, 2009), to enhance FDD performance. Du, Jin and Yang (2009) proposed a wavelet neural network to isolate the faults of temperature, pressure and flow rate sensor offsets for a VAV system. The wavelet analysis was employed to capture the essential characteristics of the VAV system operation, in order to reduce the false alarm rate. Allen, Rubaai and Chawla (2016) utilized fuzzy logic to detect VAV faults and to produce fault symptoms for different fault types, and employed ANN to identify the fault type based on the fault symptoms. Zhou, Wang and Xiao (2009) also applied fuzzy models to quantifying the residuals of important performance indices for chillers, and used ANN to diagnose faults according to the fuzzy models' outputs. Shi et al. (2017) used ReliefF filter for feature selection and employed a Bayesian ANN to identify the refrigerant charge faults for a VRF system. The proposed method was able to achieve correct diagnosis rates higher than 90%. Shi et al. (2018) further introduced the PCA method into the previous work, and utilized two Bayesian ANNs to identify refrigerant charge faults with different charge levels, respectively. The proposed method improved diagnosis accuracy by over 9% for refrigerant overcharge and undercharge faults. Sun et al. (2017) performed independent component analysis for fault detection and used back-propagation neural network (BPNN) for fault diagnosis. The hybrid model enhanced the diagnosis accuracy for refrigerant overcharge fault by 11.1%, compared with individual BPNN model. Guo et al. (2017) adopted the association rule mining technique to optimize the feature selection process for developing a BPNN-based fault classifier. The introduction of association rule mining technique facilitated an increase of 7.7% in fault diagnosis rate for the BPNN-based fault classifier.

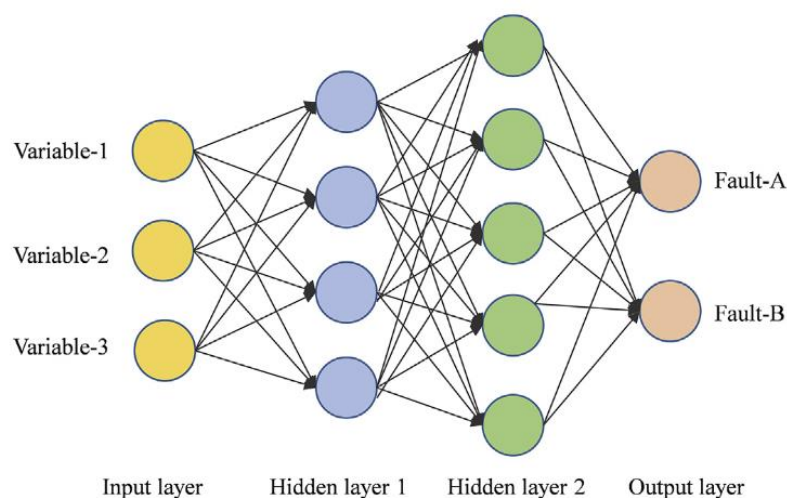


Figure 2.15 Diagram of an FDD method using ANN. (Zhao et al., 2019)

Moreover, deep learning, as a group of state-of-the-art machine learning techniques, has caught substantial attention in the recent years, and shows a significant potential in achieving more accurate and robust FDD performances than conventional supervised classification methods. Recurrent neural network (RNN) (Fig. 2.16) has advantage in capturing the temporal patterns in high-dimensional datasets, such as HVAC operational time-series datasets. Therefore, the vanilla RNN and its variants, such as Long Short-Term Memory (LSTM) (Hochreiter and Schmidhuber, 1997) and Gated Recurrent Unit (GRU) (Cho, van Merriënboer and Bahdanau, 2014), have been broadly utilized in detecting and diagnosing HVAC operational faults. Shahnazari et al. (2019) implemented RNN for identifying multiple simultaneous faults cases for VAV systems, and the developed FDD model outperformed FDD approaches using subspace identification techniques. Taheri et al. (2021) proposed different configurations of deep RNNs using LSTMs to diagnose AHU faults, and the optimal configuration indicated substantially higher performance metrics than RF and gradient boosting algorithms. Tian et al. (2021) integrated a feature extraction technique, mRMR algorithm with the LSTM model to diagnose chiller faults. The optimal LSTM model could achieve a diagnosis accuracy of 98.92%, and outperformed other machine learning methods, including vanilla RNN, BPNN and SVM. Zhang et al. (2022b) developed a novel FDD method using the integration of slow feature analysis and LSTM, which outperformed other machine learning-based methods including SVM, temporal convolutional network, convolution neural network (CNN) and deep belief network. Han et al. (2022a) combined three feature selection techniques, Correlation Analysis, Information Entropy, and Gini Impurity, with an attention-based BiLSTM model to identify four typical VRF faults. The proposed model was able to reach a fault diagnosis accuracy up to 98.3% under system working conditions included in the model training, and could also achieve an average fault diagnosis accuracy of 86% under unseen system working conditions for the model.

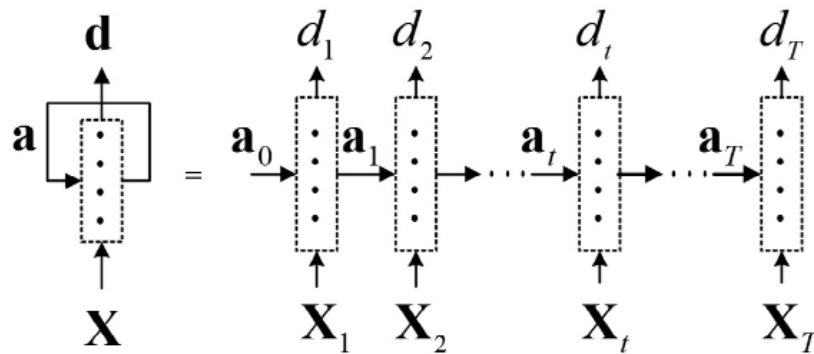


Figure 2.16 Diagram of an RNN architecture. (Taheri et al., 2021)

CNN is powerful in extracting discriminative features from either one-dimensional (1D) sequential data (Fig. 2.17) or two-dimensional (2D) matrix data (Fig. 2.18). As a result, CNNs can be either employed as a feature extractor to reduce data dimensions for the connecting FDD classifier or directly utilized as an end-to-end FDD classifier (Zhang, Saeed and Sadeghian, 2023). Regarding FDD studies using 1D-CNN algorithm, Eom et al. (2019) employed 1D-CNN to address the limitation of requiring expert knowledge for feature extraction and trained a CNN-based classifier to reach an average diagnosis accuracy of 99.9% for refrigerant charge faults of heat pump systems. Zhou et al. (2021) proposed a 1D-CNN-based FDD classifier to identify multiple coupled-faults scenarios for VRF systems and the model diagnosis accuracies for all faults were over 97.5%. Cheng et al. (2020) compared the FDD performance of ensemble 1D-CNN models with a individual 1D-CNN model and the results showed ensemble 1D-CNN models achieved a diagnosis accuracy of 97.4%, higher than that of 95.0% for the individual 1D-CNN model. Cheng et al. (2021) developed a ensemble multi-scale 1D-CNN architecture for AHU FDD, which used multiple convolutional kernel sizes to extract crucial features from multiscale AHU sensor signals. The developed architecture had superior FDD performance than other FDD methods using fully-connected neural networks, SVM, combined PCA and SVM approach, wavelet neural networks and extreme learning machine (ELM).

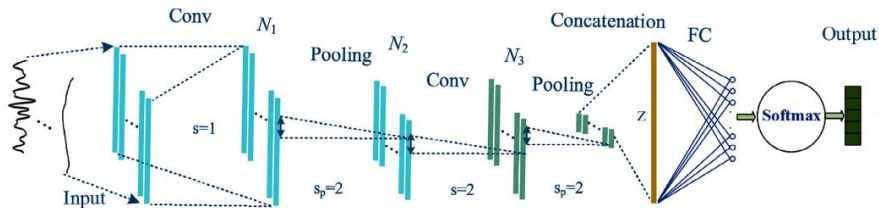


Figure 2.17 Diagram of a 1D CNN for sensor signal classification. (Cheng et al., 2021)

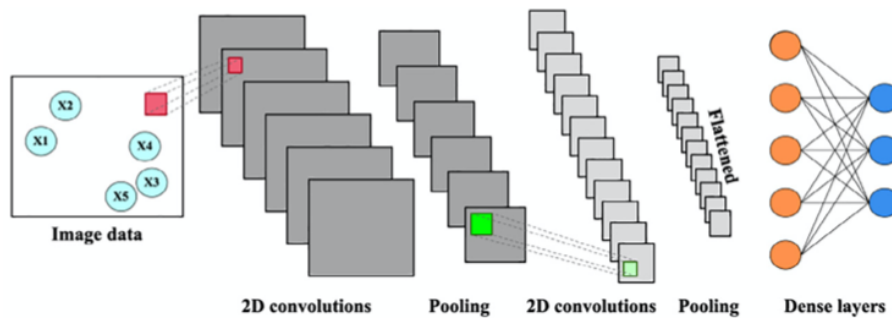


Figure 2.18 Diagram of a 2D CNN for image data classification (Fan et al., 2022)

As for 2D-CNN-based FDD methods, Elnour and Meskin (2020) transformed 1D HVAC sensor time series data into 2D matrix data and employed 2D CNN to diagnose VAV actuator faults. The fault diagnosis accuracy of the proposed method could reach up to 99.7% and outperformed shallow neural network and SVM. Gao et al. (2021) proposed a FDD model using data self-production and 2D-CNN for identifying chiller faults. The data self-production technique converted 1D sequential chiller operational data into 2D matrix, and contributed to an improvement in fault diagnosis accuracy of 1.63% compared with the CNN without using data self-production technique. Tun, Wong and Ling (2023) employed Gramian angular field (GAF) to transform 1D HVAC time series data to 2D data and developed a 2D CNN architecture to identify AHU faults. The developed GAF-2DCNN model could reach a diagnosis accuracy of 97% and indicated superior performance than SVM, RF, combined RF-SVM, 1D CNN methods. In addition, 2D-CNN FDD classifiers have been utilized as the pre-trained model for transfer learning among different chillers (Liu et al., 2021) and different AHUs (Fan et al., 2022).

Apart from these, some studies fully utilized the strengths of CNNs and RNNs, and combined them to detect and diagnose HVAC faults. Sun et al. (2019) proposed a convolution-sequence architecture consisting of a CNN, an encoder-decoder and a RNN to preprocess operational data for a heat pump system. Based on the convolution-sequence model, another CNN model was employed to diagnose the faults. The diagnosis accuracy of the proposed method was 1% higher than the pure-CNN method. Wang et al. (2020) combined 1D-CNN and GRU to diagnose chiller faults and the combined 1D-CNN-GRU model outperformed 1D-CNN-LSTM, 1D-CNN, GRU, LSTM, BPNN and combined PCA and BPNN model. Han et al. (2022b) developed a 1D-CNN-LSTM model with fused attention mechanism for chiller FDD. The fused attention mechanism facilitated the refinement of extracted features, and therefore the developed model achieved a diagnosis accuracy of 97%.

2.4.2.1.2 Supervised Regression Methods

Regression methods are typically used to establish a baseline HVAC system model by predicting system operational conditions. There are two main groups of supervised regression methods in the literature: support vector regression (SVR)-based methods and ANN-based regression methods.

SVR is an algorithm derived from the SVM, which is based on a similar mechanism to the SVM. The algorithm aims to search for the optimal hyperplane with the maximum margin to minimize the errors of data points outside the margin. Yang et al. (2013) developed a SVR-based prediction model to capture normal operation conditions of AHU sensors. Based on the SVR prediction model, statistical residuals were calculated between predictions and real sensor measurements to detect large sensor offsets, and fractal correlation dimension-based residuals were used to detect small sensor offsets. Zhao, Wang and Xiao combined SVR with exponentially-weighted moving average (EWMA) control charts to detect chiller faults (2013a), cooling tower and heat exchanger faults (2013b). The SVR was applied to establishing a reference chiller model and the EWMA control charts were used to address the limitation of t-statistic-based strategy from the previous work in detecting chiller faults. Compared with the method combining multiple linear regression (MLR) and t-statistic, the proposed method indicated substantial enhancement in detecting chiller faults at low intensity levels. Liu et al. (2018) applied the SVR to predicting VRF system power consumption and the EWMA control chart was used to diagnose refrigerant charge faults. The diagnosis accuracy of the combined SVR-EWMA control chart approach could reach at least 99.0% for refrigerant undercharge faults and at least 90.3% for refrigerant overcharge faults. Tran et al. (2016) improved the combined SVR-EWMA control chart approach by replacing the standard SVR with the nonlinear least squares support vector regression to diagnose chiller faults. The proposed method improved the diagnosis accuracy of the standard SVR-EWMA control chart method from 0.5% to 8.15% for different chiller faults.

As for ANN-based regression methods, Lee et al. (2004) applied the general regression neural-network (GRNN) to establishing AHU baseline models for detecting faults. Based on the residuals between predictions and actual measurements generated by GRNN models, the AHU faults could be identified by expert rules. Mavromatidis, Acha and Shah (2013) employed ANN to predict the baseline energy performance of a HVAC system in a supermarket and the fault detection of the HVAC system was carried out based on the predicted baseline system energy performance. Du et al. (2014a; 2014b) proposed a dual neural network, which combined a basic neural network with an auxiliary neural network, to detect AHU faults. The basic neural network was utilized to predict the SA temperature and the auxiliary neural network was employed to predict another crucial feature for the basic neural network, i.e., the return water temperature. The detected faults were then diagnosed to

their corresponding fault classes using the subtractive clustering analysis (2014a) and the PCA (2014b).

In addition to these two groups of supervised regression methods, van Every et al. (2017) applied Gaussian process regression to predicting the AHU measurements. The prediction error and variance based on the Gaussian process regression model were fed to a SVM model for fault detection. Wang and Gao (2017) employed kernel-based partial least square to model the reference performance for condenser and evaporator of a chiller for fault detection. Karami and Wang (2018) used Gaussian mixture model regression to construct a reference model for chillers and the proposed method was able to detect and identify multiple simultaneous faults in a multi-chiller plant. Tran, Chen and Jiang (2016) compared the FDD performances among the combinations of three regression methods, i.e., multiple linear regression, Kriging and radial basis function (RBF), and EWMA control charts. The comparison results indicated the highest diagnosis accuracy for the RBF-based FDD method.

2.4.2.2 Semi-supervised Learning-based FDD Methods

A sufficient labelled fault dataset is essential for supervised FDD methods. However, there is usually a limited amount of labelled fault data available for HVAC system FDD. Semi-supervised learning methods are appropriate alternatives to deal with labelled fault data insufficiency situations. Semi-supervised methods categorize unlabelled data into labelled classes based on fault and normal data patterns learnt from the limited amount of labelled data, and thus increase the amount of labelled fault data for the FDD task.

Elnour, Meskin and Al-Naemi (2020) proposed a semi-supervised VAV system sensor fault diagnosis method using Auto-Associative neural network (AANN). The AANN was capable of fixing sensor faults, filling missing sensor values using validated sensor data, removing sensor noise and correcting faulty sensor values. The fault diagnosis accuracy of the proposed method was 40% higher than the PCA-based sensor fault diagnosis method. Albayati et al. (2022) compared the FDD performance of two semi-supervised methods, i.e., the combined SVM and k-nearest neighbors (KNN) labelling method and the combined SVM, clustering and KNN labelling method with a supervised method using SVM. Based on the performance comparison, a hybrid method of the supervised and semi-supervised methods was proposed to determine which method of the two methods to be used for RTU FDD, which could reach an

average fault diagnosis accuracy of 95.7%. Yan et al. (2018) carried out semi-supervised learning using five machine learning algorithms, including SVM, RF, KNN, CART and ELM, for detecting and diagnosing AHU faults. The test results revealed that semi-supervised SVM could reach a fault diagnosis accuracy between 80% and 89% with only 30 real samples for each fault class. However, this proposed method could only enrich the fault dataset after fault occurrences. Fan et al. (2021b) also utilized self-training strategy to develop a semi-supervised FDD method to detect and diagnose AHU faults. The method was able to improve the fault detection rate for unseen faults by up to 27.9%. Fan et al. (2021c) further improved the semi-supervised FDD method to achieve a maximum recall of 10.06% under the minimum number of labelled fault samples of 280. In addition, Martinez-Viol et al. (2022) developed a semi-supervised transfer learning method based on neural networks and the developed method could achieve a F1 score of 92% (F1 score is a metric suitable for evaluating machine learning model performance under class-imbalanced data distributions). Deng et al. (2023) employed neighbour-optimized graph convolutional network for a semi-supervised FDD for chiller faults. The proposed method created an association graph between the limited number of labelled samples and a great amount of unlabelled samples based on spatial and temporal correlations among data samples. As a result, the average fault diagnosis accuracy of the proposed method could reach 97.6% when the amount of training data samples from ASHRAE research project 1043 is 240, and 94.1% when the amount of training data samples from a real chiller dataset is 720.

Moreover, GAN (Fig. 2.19) has been regarded as a promising algorithm candidate for semi-supervised FDD methods. Li et al. (2021a) proposed a semi-GAN chiller FDD framework where the discriminator of GAN was responsible for not only identifying whether each synthetic sample is real or fake, but also predicting the fault class each sample belongs to. Hence, both labelled and unlabelled fault data can be fed into the training process of the framework. The proposed framework could achieve a fault diagnosis accuracy of 84% with 80 labelled samples and 16000 unlabelled samples, compared with the supervised methods which could reach a diagnosis accuracy up to 65%. Beyond this work, Li et al. (2021b) developed a semi-supervised FDD method for VAV systems using a modified GAN which substituted the binary discriminator in the vanilla GAN with a multiclass classifier. Self-training was also used to reinforce the robustness of the method against the normal and fault data class imbalance in the training set. The proposed method was capable of enhancing the

fault diagnosis accuracy by approximately 10%, compared with semi-supervised SVM and semi-supervised neural network. Zhang et al. (2023a) proposed a end-to-end semi-supervised FDD method by combining variational auto-encoder (VAE), GAN, and bagging classifier for VRF heat pump systems. The combined model demonstrated a high average fault detection rate up to 98.9% for unknown faults.

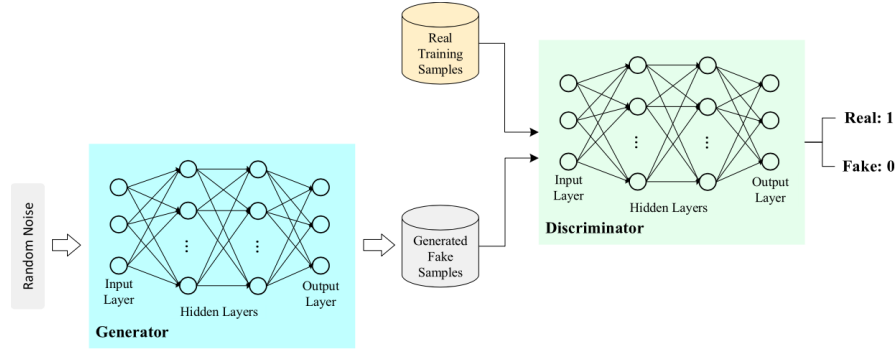


Figure 2.19 Diagram of a GAN architecture. (Li et al., 2021a)

2.4.2.3 Unsupervised Learning-based FDD Methods

When the data labels are not available from data collections, unsupervised learning FDD methods are appropriate for detecting and identifying faults in such a situation. Overall, there are two popular categories of unsupervised FDD methods in the literature: (1) PCA-based methods; (2) clustering-based methods.

PCA is one of the most common algorithms utilized in unsupervised FDD methods. It converts the high-dimensional data into a new coordinate system which can represent the important data information using less dimensions than the original data (Wold, Esbensen and Geladi, 1987). The extracted important data information in the low-dimension space can then be used for FDD. Zhao et al. (2019) explained the mechanism of PCA-based FDD approach as shown in Figure 2.20: the PCA transforms the original data space into two orthogonal subspaces, i.e., principal component subspace and residual subspace, which stands for the normal data space and the fault data space, respectively. According to PCA-transformed subspaces, fault detection can be implemented using the t-squared statistic or the Q-statistic, and fault diagnosis can be achieved based on the Q-statistic and Q-contribution plot (Wang and Xiao, 2004; Wang and Cui, 2005). Wang and Xiao (2006) further improved the previous PCA-based AHU FDD method (Wang and Xiao, 2004) by incorporating a condition-based adaptive scheme to contribute to detecting slowly evolving faults.

The integration of PCA with joint angle analysis (JAA) is also commonly used for HVAC FDD. Du and Jin applied the combined PCA and JAA approach to detecting and identifying individual AHU sensor offset (2007a), multiple AHU sensor faults (2007b) and actuator faults (2007c). In addition, the PCA and Fisher discriminant analysis were used to detect and identify individual AHU sensor fault (Du, Jin and Wu, 2007), and multiple AHU sensor and actuator faults (Du and Jin, 2008). Apart from these, wavelet analysis has also been integrated with the PCA method for chiller sensor FDD (Xu, Xiao and Wang, 2008) and AHU FDD (Li and Wen, 2014). Wavelet analysis contributed to removing noise in the original data in these work and the combined wavelet analysis PCA method indicated superior FDD performances compared with the traditional PCA method. Guo et al. (2017) combined the Satizky-Golay (SG) method with PCA for VRF system sensor FDD and the SG method was used to smooth the initial data before being fed into the PCA model. The proposed SG-PCA FDD method was able to identify large sensor offsets, but indicated low detection and diagnosis accuracies for small sensor offsets. Guo et al. (2017) further proposed a modularized PCA method with expert-based multivariate decoupling for VRF system FDD. The modularized PCA method was advanced in not only detecting faults, but also locating the fault sources. The expert-based multivariate decoupling was then used to identify the fault types. Mao et al. (2018) incorporated Empirical Mode Decomposition (EMD) threshold denoising into the PCA model for chiller sensor fault detection. The EMD threshold denoising was advantageous in removing noise but keeping the crucial information in the original data. The fault detection efficiency of the proposed method was demonstrated to be higher than the traditional PCA method. Li and Hu (2018) applied ensemble empirical mode decomposition (EEMD) denoising to removing the noise and extracting the crucial information in the original data. Based on the extracted information, the PCA model was used to detect chiller sensor faults. The validation results showed that the EEMD-PCA method improved the average fault detection rate by 22% compared with the traditional PCA method. Wang et al. (2018) integrated PCA with Bayesian network to detect and diagnose chiller faults. The PCA decomposed the normal data space into principle component subspace and residual subspace, and the residual subspace, rather than principle component subspace, was employed to establish the Bayesian network. The fault diagnosis accuracies for seven typical chiller faults by the proposed method was demonstrated to be higher than those by the PCA- Bayesian network developed based on the principle component subspace.

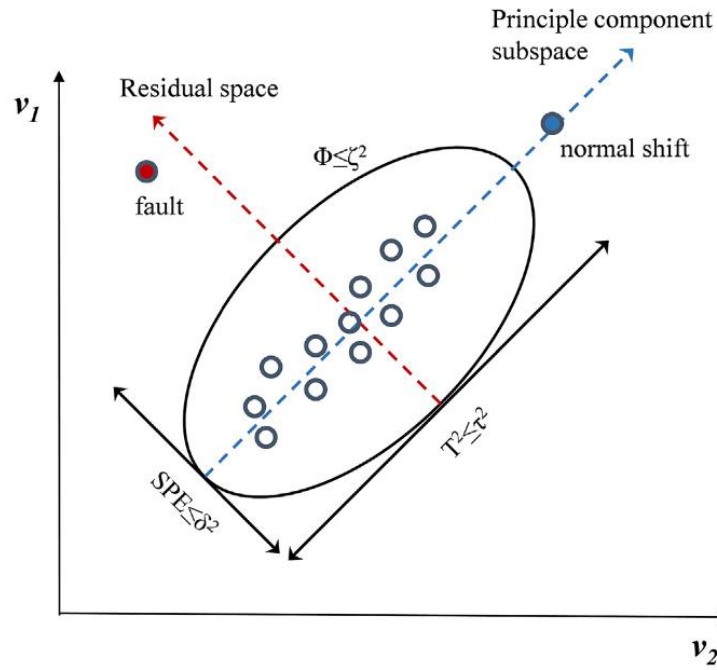


Figure 2.20 PCA-based fault detection demonstration. (Zhao et al., 2019)

Clustering-based methods divide HVAC operational data into different clusters, i.e., normal class and fault classes, based on the geometric distances among clusters (Fig. 2.21). Zogg, Shafai and Geering (2006) developed a vector clustering method for heat pump fault detection. Du et al. (2014) utilized subtractive clustering analysis to isolate AHU faults into different clusters. Chen and Wen (2017) applied Symbolic Aggregate approximation technique to producing data clusters under similar weather conditions. Base on the data clusters, a PCA model was used to detect VAV system faults. Luo et al. (2019) proposed a k-means clustering-based chiller sensor FDD approach. The threshold for sensor fault detection was determined using the featuring centroid score acquired from normal sensor measurement data and the sensor fault diagnosis dataset was produced based on centroid score patterns from various chiller sensor faults. Liu et al. (2019) extracted meaningful features from the original data using 1D-CNN and implemented WaveCluster clustering method to indicate different fault classes as clusters.

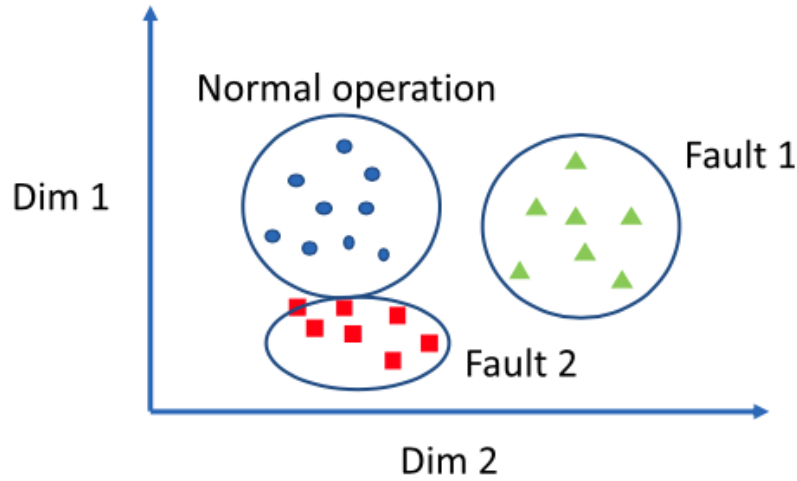


Figure 2.21 A clustering-based FDD method. (Chen et al., 2022)

Density-Based Spatial Clustering of Applications with Noise (DBSCAN) is a popular and effective clustering algorithm for separating fault data into different groups. Capozzoli, Lauro and Khan (2015) employed the DBSCAN approach to categorize fault data into clusters and a ANN was used to detect HVAC faults. Li and Hu (2018) adopted the DBSCAN technique to group chiller operational data into clusters of different working conditions and dual sub-PCA models, instead of individual PCA model, were used for detecting and diagnosing chiller faults. Compared with the traditional individual PCA approach, the developed method enhanced fault detection rate by 29.8% and fault diagnosis rate by 27.9%. Wen et al. (2023) also applied DBSCAN to separate the AHU data into clusters under a variety of operating conditions and used the SG method to smooth the data in each cluster. The preprocessed data was then fed into the PCA model for detecting and identifying AHU individual and multiple faults. Compared with the SG-PCA method, the proposed method substantially mitigated the required data amount for FDD by 13.7%.

2.4.3 Data-driven FDD Practical Challenges and Potential Solutions

Although data-driven FDD methods have been extensively proposed in existing studies, there are still some challenges remaining in the practical applications of these methods to real world HVAC fault scenarios. One of the most common challenges for applying data-driven methods to HVAC FDD in reality is the limited amount of fault data available for training data-driven FDD, which leads to undesired performances of data-driven FDD models. In addition, the class imbalance between the minority classes, i.e., fault data samples, and the majority class,

i.e., normal data samples, induces trained FDD models to be biased towards the normal class, which further results in FDD performance degradations.

2.4.3.1 Labeled Fault Data Insufficiency

Due to the poor HVAC operational data storage in old buildings and the low HVAC fault occurrences in newly-constructed buildings, the amount of fault data is usually lacking for supporting the development of data-driven FDD methods. Besides, the process for labelling HVAC operational data is considerably labourious and time-consuming, which causes the insufficiency of labelled fault data available for training the mainstream data-driven FDD methods, i.e., supervised learning-based FDD methods. To conquer this issue, semi-supervised methods have been alternatively applied to the insufficient labelled fault data situation. Semi-supervised FDD methods can capture the underlying patterns from unlabelled data and iteratively feed the training dataset with confidently labelled test fault samples. A variety of semi-supervised methods have been employed in existing studies and discussed in subsection 2.4.2.2. In particular, one of the advanced generative modeling techniques, GAN, has been widely-used to address the insufficient labelled fault data issue in VAV and AHU systems (Yan et al., 2018; Li et al., 2021b), chillers (Li et al., 2021a), VRF heat pump systems (Zhang et al., 2023a), etc.

Apart from semi-supervised GAN methods, other cutting-edge deep learning strategies have also been utilized to tackle the labelled fault data insufficiency. Zhang et al. (2022c) proposed a new feature transformation transfer learning strategy for dealing with an extreme scenario that only limited amount of normal data is available for the target HVAC system. The test results demonstrated the proposed transfer learning-based FDD method could achieve a fault diagnosis accuracy up to 0.90, which enhanced the diagnosis accuracy for conventional transfer-learning-based methods by up to 0.21. Fan et al. (2023a) implemented three self-supervised learning methods based on two self-prediction methods and a contrastive learning method (Siamese network), respectively, to extract useful information from unlabelled HVAC operational data. Based on the preprocess data, a transformer-based architecture was used to diagnose AHU faults. Compared with conventional supervised FDD methods, the proposed methods not only decreased the number of labelled fault data required for maintaining high FDD performances, but also enhanced the fault diagnosis rate by 8.44%. Chen, Xiao and Guo (2023) developed a similarity learning-based FDD method using Siamese networks for real

world HVAC fault scenarios, including labelled data insufficiency and, fault and normal class data imbalance. Compared with the baseline dual LSTM architecture, the developed method was proven to improve fault diagnosis accuracy by up to 45.7% when the amount of labelled fault samples was lacking, but presented lower fault diagnosis accuracy when the amount of inter-class imbalanced faults was over 12. Fan et al. (2023b) employed graph convolutional networks (GCN) to fully using both labelled and unlabelled HVAC data for AHU FDD. The 2D HVAC time series data was converted into association graphs using KNN and GCN was then applied to inferring discriminative information based on the association graphs for detecting and diagnosing faults. The proposed GCN-based method was tested under data scenarios containing different amounts of training fault samples, and was able to improve the average performance improvement ratio from 2.9% to 7.3% compared with fully connected neural networks and KNN.

2.4.3.2 Fault and Normal Data Class Imbalance

The fault data insufficiency issue leads to another major practical challenge for data-driven FDD methods to achieve robust performances in reality, which is the imbalance between the amount of normal data samples and fault data samples. Class imbalance is a common issue in many real-world anomaly detection applications and can be addressed from either the data side or the algorithm side. Class imbalance imposes a biases on the trained FDD models, and hence leads the FDD models to perform well on detecting the majority class, i.e., the normal data class, but misclassify the minority group, i.e., the fault classes. In terms of the algorithm-side solutions, the cost-sensitive learning and adding new loss functions are the two typical options for mitigating the influence of class imbalance on the trained model classification performance (Johnson and Khoshgoftaar, 2019). Cost-sensitive learning assigns a weight or a penalty score to each minority class, in order to reduce the model's bias towards the majority class (Krawczyk, 2016). Yan et al. (2018) embedded the costs of false positive and false negative into the cost-sensitive classification accuracy, different from the conventional classification accuracy, and employed this metric in the training of a baseline SVM model. Han et al. (2022b) applied dynamic cross-entropy to adjusting the weights of a 1D-CNN-LSTM classifier in real time, in order to improve the classifier's focus on hard-diagnosed fault classes.

Apart from the algorithm-side solutions, the efforts of most HVAC FDD studies focused on proposing data-side solutions. One of the most prevalent data-side solutions is data augmentation. FDD researchers used various data augmentation techniques, containing random data sampling and generative modeling, to increase the amount of fault samples in the training dataset. Typical random data sampling techniques consist of random under-sampling, random over-sampling, synthetic minority over-sampling technique (SMOTE) and its variants. Fan et al. (2019) performed a combined PCA and SMOTE method to increase the amount of fault samples in the training dataset at an imbalanced ratio of 5% between fault and normal data. Based on the oversampled data, SVM was employed to diagnose chiller faults. Compared with the original SVM method, the combined PCA-SMOTE-SVM method enhanced the fault diagnosis accuracy by 0.8%. Zhou et al. (2021) integrated the PCA-SMOTE technique with three classification models, including BPNN, CART and SVM, to diagnose VRF system faults. The integrated methods were tested under the scenarios where the amount of fault samples ranged from 50 to 10. The test results revealed that the PCA-SMOTE technique was able to enhance the geometric mean accuracy for three classification models by around 44% in average. Zhang et al. (2023b) combined random under-sampling and SMOTE to rebalance the training datasets for developing chiller FDD models. The random under-sampling was used to undersample the normal class and the SMOTE was employed to oversample the fault classes. Based on the rebalanced training datasets, the ensemble ELM models were utilized to diagnose chiller faults. The fault diagnosis accuracy of the proposed resampling-based ensemble ELM method was 9.4%-11.0% higher than other ensemble ELM methods. Shen et al. (2023) integrated a stable SMOTE technique, which could mitigate the overlap of the synthetic fault samples, with self-attention mechanism based temporal convolutional network (STCN) for chiller FDD. Compared with other traditional SMOTE techniques, the stable SMOTE technique was able to facilitate the STCN to achieve higher recall and precision scores.

Apart from traditional random data sampling techniques, generative modeling has played a critical role in addressing fault and normal class imbalance in HVAC operational data. VAE (Kingma and Welling, 2014), as an extension to the vanilla autoencoders, has been attempted in recent studies. Fan et al. (2021a) compared the AHU fault diagnosis improvements, driven by conventional random data sampling techniques and a cutting edge generative modeling network, i.e., VAE, under imbalanced data scenarios. According to the evaluation results, the

VAE-based methods resulted in most significant fault diagnosis enhancements when the amount of training fault samples was limited, such as, 20 for each fault class. The SMOTE-based methods led to highest fault diagnosis improvements when the imbalanced ratio between normal data and fault data was higher than 15 and the amount of training fault samples was higher than 60.

Beyond the above-mentioned methods, GAN, as a robust generative modeling algorithm, has extensively proved its ability in generating synthetic fault samples with high realness compared with real fault samples. Zhong et al. (2019) integrated a generative modeling technique, Wasserstein GAN (WGAN) with conventional machine learning-based FDD classifiers for AHUs. The WGAN is responsible for producing synthetic fault samples for the training set. The fault diagnosis accuracy for the combination of WGAN and ensemble SVM approach could reach 90.4% with only 40 initial real samples for each fault class. Yan et al. (2020) further enhanced Zhong et al.'s work (2019) by modifying the WGAN to a new extension of GAN, conditional Wasserstein GAN (CWGAN). Compared with WGAN, CWGAN produces synthetic AHU fault samples by taking into account fault class labels. As a result, an nearly uniform distribution of fault samples from different classes can be obtained by CWGAN. The combination of CWGAN with a single RF model could achieve a fault detection rate of approximately 1 and a fault diagnosis rate of 87.3% with 40 initial real samples for each fault class. Yan, Chong and Mo (2020) implemented CWGAN to support chiller FDD under imbalanced data scenarios. The fault diagnosis accuracies for CWGAN-aided FDD classifiers were at least 10% higher than those FDD classifiers trained without incorporating synthetic fault samples generated by CWGAN.

Moreover, the hybrid use of GAN and VAE has been explored in several recent work. Yan et al. (2022) updated the CWGAN with gradient penalty (GP) to enhance the realness of synthetic fault samples, and applied VAE to selecting high-quality synthetic samples. Compared with the ensemble-learning-based fault sample selection strategy in Zhong et al. (2019)'s work, VAE demonstrated superior performance in improving fault diagnosis accuracy. Yan (2021) further proposed a GAN-based high-quality fault sample selection scheme for CWGAN-GP, and the proposed method indicated a more robust high-quality fault sample selection ability than VAE. Besides, Zhang et al. (2022d) combined WGAN-GP with

VAE for producing synthetic fault samples for improving the FDD performances of ensemble classifiers under imbalanced data scenarios.

Recently, one of the state-of-the-art deep learning architectures, Transformer (Vaswani et al., 2017), has been integrated with GAN to enhance the quality of synthetic fault samples. Yan et al. (2023a) substituted the generator in the WGAN architecture with the Transformer and adopted transfer learning strategy to diagnose AHU faults. Compared with WGAN-based ensemble classifiers, including SVM, KNN, decision tree, RF and multi layer perceptron, the Transformer-based WGAN with transfer learning (TWGAN-TL) FDD approach maximized the fault diagnosis accuracy by 15.0% when the number of available real fault samples is 20. Yan et al. (2023b) further proposed a Transformer-based CWGAN architecture for fault data generation and used deep reinforcement learning (DRL) for high-quality synthetic fault sample selection. The proposed Transformer-based CWGAN with DRL approach improved the fault diagnosis accuracy of the previous TWGAN-TL approach (Yan et al., 2023a) by 3.66%.

2.4.3.3 Poor Cross-Operation-Condition Generalization

In addition to the above-mentioned two challenges, the fault data intra-class imbalance among different HVAC system operation conditions induces the poor generalization ability of data-driven FDD methods for detecting and diagnosing faults which occur under different system operation conditions from the operation conditions that the fault samples in the training dataset belong to. The HVAC system operation conditions depend on climatic conditions of the target building location, and hence this challenge can also be regarded as the intra-class imbalance within each fault class among different climatic conditions. Based on the exhaustive search of the literature, there are only a few studies exploring the solutions for dealing with this practical issue and most work focus on proposing algorithm-side solutions, i.e., deep transfer learning-based FDD strategies. Transfer learning captures the common knowledge between a historical system operating condition (also known as the source domain) and a new system operating condition (also known as the target domain) for HVAC system FDD. Liang et al. (2022) proposed a domain-adaption-based chiller FDD framework by learning a common domain-invariant feature between the source domain and the target domain so that the data distributions of the two domains are similar. As a result, the FDD classifiers learning the domain-invariant fault feature from the source domain can detect and

diagnose the faults in the target domain. Two domain-adaption methods, domain adversarial neural network (DANN) and partial unsupervised domain adaption (PUDA), were applied to the proposed framework, and an ANN was used as the FDD classifier, respectively.

Compared with baseline FDD classifiers, including SVM, RF and ANN, the PUDA-based FDD method was able to enhance the average fault diagnosis accuracy among all data scenarios by 14.5%. Li et al. (2023b) compared FDD performances under different chiller working conditions based on three different deep transfer learning strategies, containing network-based fine-tuning (FT), mapping-based domain-adaptive neural network (DaNN) and DANN. The FT-based method was found to outperform the other two methods and achieved an average fault diagnosis accuracy of 93% under cross-working-condition scenarios. Jiang et al. (2023) proposed a vision transformer generative adversarial domain generalization (VIT-GADG) method that captured domain-invariant information in source and latent domains and utilized the extracted information to diagnose faults in the target domain under unseen chiller operation conditions.

On the other hand, Li et al. (2022b) proposed an Adversarial Cross domain Data Generation method that utilized the encoder-decoder structures as fault data generators. The feature patterns and common knowledge were extracted from the source domain and used for fault data generation for the target domain. The evaluation experiments tested the cross-domain fault sample generation ability for exhausted air and OA damper stuck from one season to another season. Although the proposed method outperformed an instance-based transfer learning method, the gap in diagnosis accuracy, between the proposed method and the supervised classification method that accessed the label information from the target domain, still ranged from 2.3% to 30.5%.

2.4.4 Summary and Research Gaps

Existing data-driven FDD methods can be categorized into supervised learning, semi-supervised learning and unsupervised learning methods. Although the majority of them claim robust FDD performances for a variety of HVAC systems and components, including VAV systems, VRF systems, AHUs, chillers, heat pumps, etc, there are still three main practical challenges causing FDD performance degradations when these methods are applied to online HVAC FDD. The labelled fault data insufficiency, and fault and normal class imbalance issues have been paid significant attention to. Semi-supervised learning methods (Zhang et al., 2023a), transfer learning strategies (Zhang et al., 2022c), self-supervised learning method,

such as, contrastive learning based on Siamese network (Fan et al., 2023a), and graph-based methods, such GCN (Fan et al., 2023b), have been demonstrated to be effective in dealing with limited labelled fault data situations. As for the fault and normal class imbalance issue, cost-sensitive learning (Yan et al., 2018) and adding new loss functions (Han et al., 2022b) are two typical solutions from the algorithm perspective. In terms of data-side solutions, data augmentation techniques, such SMOTE technique (Zhou et al., 2021), and generative modeling algorithms, including GAN (Yan, Chong and Mo, 2020), VAE (Fan et al., 2021a) and the hybrid models of them (Yan et al., 2022) , have been extensively explored.

Apart from these two practical challenges, the generalization weakness under unseen system working conditions for data-driven FDD methods has not yet been well addressed. Based on the exhaustive search and sophisticated review of the relevant studies, only three studies investigated the solutions to this practical issue from the algorithm perspective, and only one study considered tackling the issue from the data perspective. Consequently, there is still a significant gap in enhancing FDD performances under different HVAC system working conditions. Due to the dependence of HVAC system working conditions on outdoor climatic conditions, the required task can also be regarded as addressing FDD performance degradations under the fault data intra-class imbalance scenarios among different climatic conditions. Moreover, the impact of this fault data intra-class imbalance on FDD performance has not yet been quantified.

2.5 Research Gaps Summary

The review of fault modeling field is mainly used to pave the way for fault impact analysis and data-driven FDD, and thus identifying research gaps in the fault modeling field were not the focus of this chapter. To the author's best knowledge, the research gaps from the fault impact analysis field and data-driven FDD field are concluded as below:

1. Apart from the work investigating HVAC fault impacts under climatic conditions of different climate zones, HVAC fault impacts under changing climatic conditions (changing climate periods and seasonal weather condition variations) have not yet been evaluated.
2. The interactions among a wider array of multiple HVAC faults, such as the interactions among sensor and actuator faults have not yet been investigated.

3. The impact of the fault data intra-class imbalance among different climatic conditions (system working conditions) on FDD performance has not yet been quantified and the resulting FDD performance degradation has not yet been well addressed.

Chapter 3 Fault Modeling and Simulation

3.1 General Workflow

Fault modeling and simulation is an effective alternative method to resolve the HVAC fault data scarcity issue, especially the data scarcity under some HVAC system operating conditions. In addition, the evaluation of HVAC fault impacts under different climatic conditions can also be carried out through modeling and simulating the faults in BPS tools. This chapter will go through the entire workflow of fault modeling and simulation, as presented in Figure 3.1a. In this study, the fault models were developed based on the modeling methods described in the work from Li and O'Neill (2018), Cheung and Braun (2015), and Kim et al. (2019). By incorporating the fault models into a widely-used BPS software, EnergyPlus (U.S. DOE, 2020), fault simulations were conducted to generate simulated fault data for further fault impact analysis and FDD framework development. Section 3.2 will explain the fault characteristics and modeling approaches for each fault. Section 3.3 will introduce the case study building model for this study and the verification of the model. Section 3.4 will describe the contemporary and future weather files and explain the methods for generating these weather files. Section 3.5 will show the processes for carrying out fault simulations. It should be noted that this section has been published in the paper 'Assessment of HVAC system operational fault impacts and multiple faults interactions under climate change' (Zhong, Calautit and Wu, 2022) and hence some contents from the paper is reused in this section.

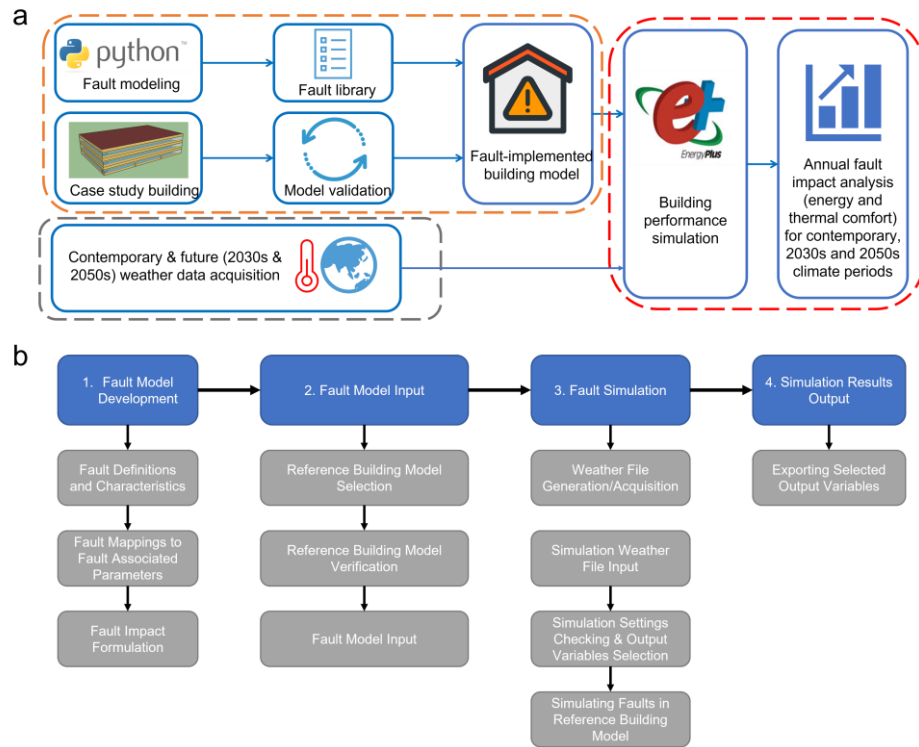


Figure 3.1 (a) The workflow for evaluating HVAC fault impacts; (b) Detailed fault modeling and simulation steps.

3.2 Fault Model Development

The steps for developing fault models are illustrated in Figure 3.1b. Firstly, the faults and their corresponding intensities were defined. Next, the associated parameter(s) in EnergyPlus, as shown in Table 3.1, were linked to each fault, and the fault impact on the associated parameter(s) was formulated using Python (van Rossum and Drake, 2019). After the fault impact formulation, the fault models were then introduced into the verified EnergyPlus building model, which will be described in section 3.3.

Five fault categories, including control, sensor, DX packaged air conditioner, VAV terminal unit, and other faults, were taken into account in this study. The definitions of intensities, their ranges, and the associated parameters in EnergyPlus for these faults are described in Table 3.1. The definition and the modeling approach for each fault are explained in subsections 3.2.1-3.2.5.

Table 3.1 Descriptions of fault characteristics (Cheung and Braun, 2015; Kim et al., 2019).

Fault name	Fault intensity	Fault intensity range	Associated parameters
Control faults			
No overnight setback	Overnight setback is set up or not	0 or 1	Heating/cooling schedules setpoints
Extended morning setback	Hours that the setpoint for the occupied period is extended forward in the morning	0 to 3 hours	Time for ending the unoccupied period in heating/cooling schedules
Delayed evening setback	Hours that the setpoint for the occupied period is delayed in the evening	0 to 3 hours	Time for starting the unoccupied period in heating/cooling schedules
Occupant manual changes of setpoint	The temperature threshold that occupants will alter the thermostat setpoints and the degree(s) of setpoints altered	User-defined	Time when the setpoint is altered and the altered setpoint is added to heating/cooling schedules
Sensor faults			
Cooling/heating coil SA temperature sensor bias	Sensor offset in coil SA temperature	±4 °C	Fault availability and severity schedules, coil/controller/thermostat object name and type, as well as reference sensor offset (U. S. DOE, 2020a)
Economizer OA temperature sensor bias	Sensor offset in OA temperature	±4 °C	
Economizer RA temperature sensor bias	Sensor offset in RA temperature	±4 °C	
Thermostat offset	Thermostat offset in zone air temperature	±3 °C	
Packaged air conditioner faults			
Supply fan motor efficiency degradation	The percentage of fan motor efficiency reduction	0 to 30%	Fan total efficiency in fan objects
Supply fan stuck at a fixed speed	Fan airflow rate stuck percentage	0 to 100%	EMS actuator: fan air mass flow rate
OA damper stuck	OA damper stuck position percentage	0 to 100%	Minimum and maximum OA fraction schedules
OA damper leakage	The ratio of the leaked OA flow rate to the OA flow rate at the full operation condition	0 to 40%	Minimum OA fraction schedules
VAV terminal unit faults			
VAV reheat coil complete failure	VAV reheat coil completely broken down or not	0 or 1	Reheat coil availability schedule(s)
Other faults			
SA duct leakage	The ratio of the leaked airflow rate to the SAflow rate	0 to 30%	Constant downstream leakage fraction

3.2.1 Control Faults

Most HVAC control faults are related to the heating/cooling thermostat setpoints of HVAC systems. The following faults were modelled in this study based on previous studies from Cheung and Braun (2015) and Kim et al. (2019):

(1) HVAC setpoint setback fault: no overnight setback

Thermostat setpoint setback is usually set to increase setpoints in the cooling season and decrease setpoints in the heating season during the building unoccupied period to reduce HVAC energy consumption. This fault occurs when the setpoint setback is not set for the unoccupied period accidentally, and thus a constant setpoint for the occupied period is used for all times.

Modeling approach: Altering the thermostat setpoint for the unoccupied period to a constant setpoint immediately before switching from the occupied period to the unoccupied period.

(2) HVAC setpoint setback fault: extended morning setback

The setpoint for the precooling/preheating/building opening time is mistakenly extended forward to the unoccupied period.

Modeling approach: Extending forward the time for the initial change of setpoint in the morning.

(3) HVAC setpoint setback fault: delayed evening setback

Similar to the extended morning setback fault, this fault delays the last shift of the thermostat setpoint from the occupied period to the unoccupied period.

Modeling approach: Delaying the time when the setpoint for the occupied period is shifted to that for the unoccupied period.

(4) Occupant manual changes of setpoint

Due to extreme outdoor temperatures or difference in thermal comfort preference of occupants, occupants may feel considerably hot or cold at the moment when they enter the conditioned space and require prompt cooler or warmer thermal sensation. Consequently, they

may adjust the thermostat setpoint manually to a lower or higher setpoint and not set back the setpoint to the original value after feeling thermally comfortable.

Modeling approach: In the heating season, this fault is modeled by defining the upper limit of the OA temperature at which occupants will manually alter the heating setpoint and the resultant increase in the setpoint. In the cooling season, similarly, this fault is modeled by setting the lower limit of the OA temperature at which occupants will alter the cooling setpoint and the resultant decrease in the setpoint.

3.2.2 Sensor Faults

A wide range of sensors are installed in HVAC systems in order to monitor system performances. These sensors contain temperature sensors, humidity sensors, enthalpy sensors, pressure sensors, flow rate sensors, etc. Sensor faults are typically caused by the aging of sensors, which leads to the drift of sensor measurements from the calibrated readings (Cheung and Braun, 2015). If the faulty sensor is not calibrated promptly, the fault may substantially affect system energy performance and occupant thermal comfort. A variety of fault objects are embedded in EnergyPlus and contribute to a simplified approach for modeling faults by entering the user-defined fault information into the required fields of fault objects. The following sensor faults were modelled in this study:

(1) Cooling/heating coil SA temperature sensor bias

The reading of the coil SA temperature sensor deviates from the actual temperature of SA leaving the coil.

Modeling approach: Establishing the availability and severity schedules for selected faulty coil(s) with user-defined reference sensor offset. The details for modeling this fault can be found in EnergyPlus Input Output Reference (U. S. DOE, 2020a).

(2) Economizer OA temperature sensor bias

The reading of OA temperature sensor deviates from the actual OA temperature.

Modeling approach: Establishing the availability and severity schedules for selected faulty economizer controller(s) with user-defined reference sensor offset. The details for modeling this fault can be found in EnergyPlus Input Output Reference (U. S. DOE, 2020a).

(3) *Economizer RA temperature sensor bias*

The reading of RA temperature sensor deviates from the actual RA temperature.

Modeling approach: Establishing the availability and severity schedules for selected faulty economizer controller(s) with user-defined reference sensor offset. The details for modeling this fault can be found in EnergyPlus Input Output Reference (U. S. DOE, 2020a).

(4) *Thermostat offset*

The thermostat reading deviates from the actual zone air temperature.

Modeling approach: Establishing the availability and severity schedules for selected faulty zone thermostat(s) with user-defined reference sensor offset. The details for modeling this fault can be found in EnergyPlus Input Output Reference (U. S. DOE, 2020a).

3.2.3 Packaged Air Conditioner Faults

Packaged air conditioner faults or RTUs are typically used in small scale buildings, which serves the function of ventilation, cooling and (or) heating for the conditioned space. The following faults were modelled according to the work from Cheung and Braun (2015) and Kim et al. (2019):

(1) *Supply fan motor efficiency degradation*

The supply fan motor efficiency degradation is a typical fault symptom for packaged air conditioners due to bearing or (and) stator winding faults.

Modeling approach: According to EnergyPlus Engineering Reference (U. S. DOE, 2020b), the fan total efficiency η_{total} is calculated by:

$$\eta_{total} = \eta_{motor} \times \eta_{impeller} \quad (3.1)$$

where η_{motor} is the motor efficiency and $\eta_{impeller}$ is the impeller efficiency.

To model the fault, it is essential to calculate the impeller efficiency using the original fan total efficiency $\eta_{total,ori}$ and motor efficiency $\eta_{motor,ori}$:

$$\eta_{impeller} = \frac{\eta_{total,ori}}{\eta_{motor,ori}} \quad (3.2)$$

Then, the faulty fan total efficiency $\eta_{total,f}$ can be determined by:

$$\eta_{total,f} = \eta_{motor,f} \times \eta_{impeller} \quad (3.3)$$

(2) Supply fan stuck at a fixed speed

The supply fan is stuck at a fixed speed due to a non-functioning fan controller.

Modeling approach: Developing EMS program codes to set the fan actuator to the faulty fixed fan air mass flow rate $\dot{m}_f$ which is calculated by:

$$\dot{m}_f = \dot{m}_{max} \times \theta \quad (3.4)$$

where $\dot{m}_{max}$ is the fan air mass flow rate at the maximum speed of fan (kg/s); θ is the fan speed stuck percentage.

(3) OA damper stuck

The OA damper is stuck at a fixed position, which can be caused by various factors, such as economizer control faults or seized actuators.

Modeling approach: Establishing the same minimum and maximum OA fraction schedules to fix the OA damper position.

(4) OA damper leakage

The OA leaks from the OA damper into the packaged air conditioners.

Modeling approach: This fault can be modelled by modifying the minimum OA fraction to the leaking ratio.

3.2.4 VAV Terminal Unit Faults

VAV terminal unit is a critical component in a VAV system used to adjust the airflow rate based on each zone's cooling/heating demand. Due to the limited available input parameters

of VAV terminal units in EnergyPlus, only the fault model of VAV reheat coil complete failure was developed in this study.

VAV reheat coil complete failure

The reheat coil in the VAV terminal unit completely breaks down.

Modeling approach: Modifying the availability schedule of reheat coils as the status of always off.

3.2.5 Other Faults

SA duct leakage

Apart from the above categories of faults, SA duct leakage was also modelled in this study. The fault occurs when SA leaks from the SA duct to an unconditioned space in a building or to the ambient environment due to various factors, such as duct disconnection and poor installation of duct fittings.

Modeling approach: In this study, this fault assumes to occur at downstream of VAV terminal unit. The modeling of this fault at upstream of terminal unit cannot be achieved in EnergyPlus since the only relevant parameter ‘Nominal Upstream Leakage Fraction’ is a ratio of leaked airflow rate to the design SAflow rate and can only lead to a constant leaked airflow rate when the actual SAflow rate varies. The assumed fault is modelled by setting the constant downstream leakage fraction in ‘Zone HVAC: Air Distribution Unit’ to the leaking ratio.

3.3 Case Study Reference Building Model

A verified EnergyPlus case study building model is essential for fault simulations based on the developed fault models. The DOE prototype commercial building model collection (Deru et al., 2011) is an ideal collection for implementing the faults models. This collection contains 16 types of commercial buildings, including offices, schools, restaurants, retail store, hospital, etc, under 19 climate locations, and the building models complies with ASHRAE 90.1-2004, 62.1-2004, and 62-1999 standards. In this research work, the prototype medium office building model, commonly used for modeling the building energy consumption (Zhang et al., 2022a), assessing the HVAC fault impacts (Li and O’Neill, 2019) and predicting the power usage (Huang et al., 2021) of commercial buildings under various climatic conditions, was used as the case study for generating simulated fault data and performing fault impact

evaluations under different seasons/climate periods. Detailed descriptions of the prototype building model development can be found in Deru et al.'s work (2011). This established prototype building model and its inputs, which embodies an extensive collection of building domain research and knowledge, were used to avoid repeating the foundational work of building energy modeling in this thesis, which is irrelevant to the main interest of the thesis. The detailed descriptions for the case study building and its HVAC system are provided in subsection 3.3.1. The verification results for the model are presented in subsection 3.3.2.

3.3.1 Building and HVAC Systems Descriptions

The medium office building is a three-storey building (Fig. 3.2a) with one core and four perimeter zones on each storey (Fig. 3.2b). The total floor area of the building is 4979.6 m². With regard to the HVAC system, a VAV system coupled with three packaged air conditioners was employed in the building, and each air conditioner has a DX cooling coil and a gas-fueled heating coil. As for the zone air distribution, each VAV terminal unit is equipped with a damper to adjust the actual SA flow rate for each zone and an electric reheating coil to meet the required zone SA temperature. The design day conditions for HVAC system sizing were adjusted to those of London (ASHRAE, 2017).

In terms of HVAC default control settings, the VAV system operates between 6 am to 10 pm on weekdays, and 6 am to 5 pm on Saturday. The zone thermostat setpoint for the occupied period is set to 24 °C for cooling and 21 °C for heating, while that for the unoccupied period is set to 26.7 °C for cooling and 15.6 °C for heating. To meet the required zone air temperature, the setpoint for the SA temperature leaving the air conditioning units is set to 12.8 °C for cooling and 15.6 °C for heating. Apart from these, an economizer is used to control OA intake for each packaged air conditioning unit depending on the difference between the OA temperature and RA temperature. The minimum OA flow rate per area is 0.43 L/s×m², and the maximum intake of OA is 4.81 m³/s for the bottom floor, 4.29 m³/s for the middle floor and 4.75 m³/s for the top floor. The details of building geometry and HVAC system specifications are shown in Table 3.2.

Table 3.2 Building geometry and HVAC system specifications.

Geometry	
Storeys	Three
Thermal zone layout	One core and four perimeter zones on each storey
Total floor area	4979.6 m ²
HVAC system configuration	
Air system type	VAV system
Cooling system type	DX packaged air conditioning unit
Heating system type	Gas furnace in the packaged air conditioning unit
Zone air distribution unit	VAV terminal unit with damper and electric reheating coil
Design day conditions for system and zone sizings	Winter: dry-bulb temperature of -2.5 °C and corresponding dewpoint of -6.1 °C over 99.6% time; Summer: dry-bulb temperature of 28.3 °C and corresponding wet-bulb temperature of 18.6 °C below 0.4% time.
HVAC control settings	
System operation schedule	Weekdays: 6 am to 10 pm; Saturday: 6 am to 5 pm
Thermostat setpoint (occupied period)	Cooling: 24 °C; Heating: 21 °C
Thermostat setpoint (unoccupied period)	Cooling: 26.7 °C; Heating: 15.6 °C
Central system SA temperature setpoint	Cooling: 12.8 °C; Heating: 15.6 °C
Economizer control type	Set the OA flow rate to its minimum when the dry-bulb temperature of OA is higher than that of RA.
Minimum OA flow rate per area	0.43 L/s×m ²
Maximum OA flow rate for each floor	Bottom: 4.81 m ³ /s; Middle: 4.29 m ³ /s; Top: 4.75 m ³ /s

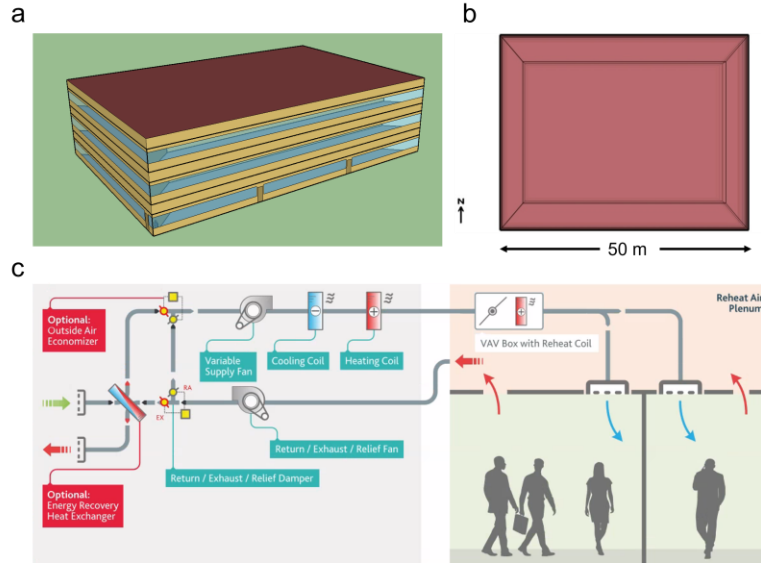


Figure 3.2 (a) Case study medium office building model; (b) One core and four perimeter zones on the floor plan; (c) A typical VAV system layout with reheat coils and outdoor air economizer (IES, 2016).

3.3.2 Model Verification

Due to the limitation of unavailable real energy measurements for this prototype building model, the simulation results from both the original model and the modified model with design day conditions of London were compared with the DOE source simulation results (Deru et al., 2011), in order to verify the validity of both models. The two metrics, i.e., coefficient of variation of root mean square error (CV(RMSE)) and normalised mean bias error (NMBE), were calculated using Eqn. 3.5 and 3.6 and used for verifying the simulation results of this study.

As presented in Figure 3.3, the two metrics for all the energy usages, except the heating coil gas energy, are lower than the absolute value of 1% for the present original model. The only large deviations for the heating coil gas energy may be attributed to the version difference between the version of EnergyPlus tool used in this study, version 9.3, and the model version used for producing the source results by DOE, version 8.0. The use of EnergyPlus 9.3 in this study is mainly due to the new fault modeling capability in this version.

$$CV(RMSE) = \frac{1}{\bar{y}} \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \times 100 \quad (3.5)$$

$$NMBE = \frac{1}{\bar{y}} \frac{\sum_{i=1}^n (y_i - \hat{y}_i)}{n} \times 100 \quad (3.6)$$

In addition, the adjustment of design day conditions of the model led to higher errors, in terms of the two metrics, for all the energy usages. However, the CV(RMSE) and NMBE for these energy usages of the present original model and the modified model are both within the accepted ranges (15% and 5%) indicated in ASHRAE guideline 14-2014 (ASHRAE, 2014), and thus the case study building model is verified to be used for further fault simulations under different climate conditions. It should be noted that further model developments and validations should be conducted in the future to ensure that the short-term and dynamic behaviours of the HVAC system and its components are captured.

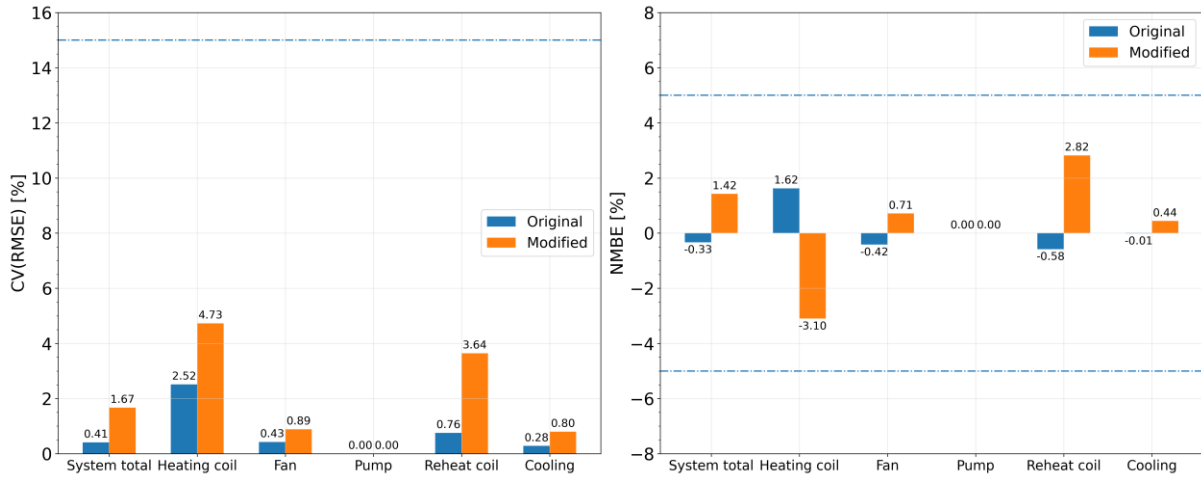


Figure 3.3. CV(RMSE) and NMBE for system total and end energy usages of the present original model and the modified model, compared with the DOE source results.

3.4 Contemporary and Future Weather Data Generation

The fault simulations were implemented using a contemporary weather file and future weather files (the 2030s and 2050s) of London. The contemporary weather file was produced from Meteonorm 8 (Remund et al., 2020), a comprehensive meteorological program for generating contemporary and future weather data at most locations worldwide in various formats of weather files. The contemporary weather data refers to the recent weather data based on radiation data from 1996 to 2015 and temperatures, wind, precipitation, etc., from 2000 to 2019 (Remund et al., 2020).

As for the future weather files, two different Intergovernmental Panel on Climate Change (IPCC) climate scenarios, i.e., Special Report on Emission Scenarios (SRES) medium emissions scenario (A1B) (IPCC, 2007) and Representative Concentration Pathways (RCP) 8.5 scenario (IPCC, 2013), were used to generate them. The SRES A1B scenario weather files was provided by Eames, Kershaw and Coley (2011) using the UKCP09 weather generator (Jones et al., 2010), and the weather files generated using the average monthly temperature for each same month between 1961 and 1990 at the 50th percentile were selected for this study. The RCP 8.5 scenario weather files were generated by Meteonorm 8 (Remund et al., 2020) in the present study and the steps for generating the weather files will be explained in detail in subsection 3.4.4.

3.4.1 IPCC Climate Scenarios

The SRES scenarios are greenhouse gas emissions scenarios used for future climate change projections and firstly published in the IPCC third assessment report (IPCC, 2000). There are several storylines predicting the future developments in the SRES scenarios, as presented in Figure 3.4. These storylines produce four sets of scenarios, including A1, A2, B1 and B2 families. The A1 storyline depicts a future world where there will be a significantly fast economic growth and innovative technology developments. The global population will reach its peak in the mid-century and the difference in per capita income decreases regionally. Based on the A1 storyline, there are three scenario groups classified by their fossil-fuel focuses: fossil-fuel intensive scenario (A1FI), non-fossil fuel sources scenario (A1T), and a balanced scenario across all energy sources (A1B). The A2 storyline represents an increasing global population and a regionally-oriented economic growth in the future. The per capita income and technology development are slower than those in other storylines. The B1 storyline describes a similar future world to the A1 storyline, but the economic structures will be transitioned to an information and service economy. The focus of this storyline will be proposing global economic, social, and environmental sustainability solutions. The B2 storyline indicates a future world with an increasing global population, but at a slower rate than the A2 storyline. The economic growth will be at an intermediate rate and the technology development will be less faster than that in the B1 and A1 storylines. The focus of this storyline will be initiating local economic, social, and environmental sustainability solutions.

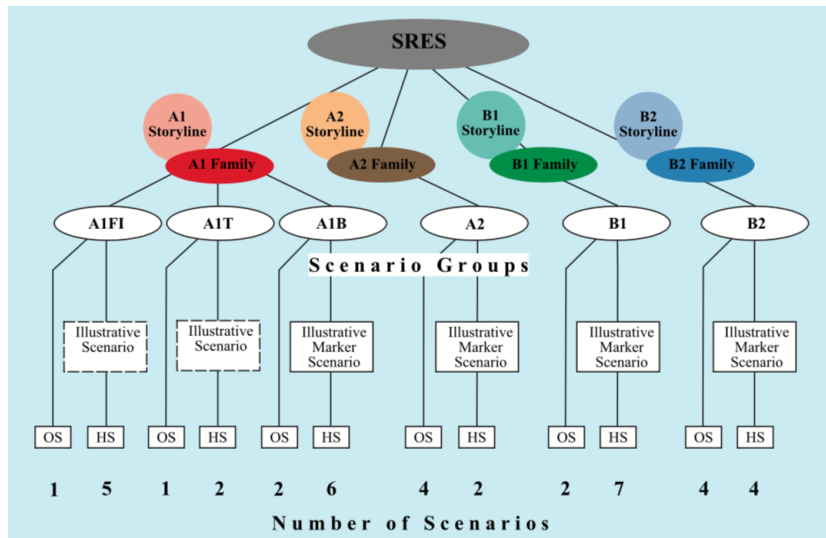


Figure 3.4 Schematic diagram of SRES scenarios (IPCC, 2000).

The RCP scenarios are new climate change scenarios proposed in the IPCC fifth assessment report (IPCC, 2013), including four scenarios representing different total radiative forcing in the end of this century relative to 1750: a greenhouse gas emission mitigation scenario (RCP2.6), two stable scenarios (RCP4.5, RCP6.0) and a considerably high greenhouse gas emission scenario (RCP8.5). Compared with the SRES used in the third and fourth IPCC assessment reports (IPCC, 2000; IPCC, 2007), RCP scenarios incorporate a series of climate policies in this century while SRESs don't consider the climate policies.

3.4.2 UKCP09 Weather Generator Introduction

The UKCP09 weather generator incorporates stochastic models to produce synthetic time-series weather data. The synthetic time-series weather data closely captures the features of real weather but is not actually observed. The stochastic process of the weather generator means that the weather conditions at the current time step are not entirely dependent on the conditions at the previous time step. The general mechanism of the weather generator is based on a stochastic rainfall model. The observations of daily rainfall sum and other weather conditions from a benchmark climate (1961–1990 for rainfall and 1961–1995 for the other conditions) are used for the calibration of the weather generator. The range of potential climate change futures are defined by monthly change factors for each grid according to the UKCP09 probabilistic projections. The stochastic rainfall model is refitted based on future daily rainfall perturbations. The other weather condition parameters are produced based on the rainfall sequences. The detailed theory about the UKCP 09 weather generator can be found in Jones et al.'s report (2010).

3.4.3 Meteonorm Introduction

Meteonorm is a comprehensive meteorological program for providing meteorological data sources for engineering simulation tools. The program is mainly based on the solar radiation calculation for surfaces of arbitrary orientations at any specified location. The uncertainty of the program is clarified as follows: the hourly model slightly overestimates the solar radiation on angled surfaces by 0–3%. The deviations from measured values is $\pm 10\%$ for monthly results and $\pm 6\%$ for yearly results. The program contains climatological databases including both ground weather stations' data and satellite data. Thus, the program is capable of generating future weather data based on RCP2.6, RCP4.5 and RCP 8.5 scenarios for future decades from 2030 to 2100. The detailed software guidance and theory can be found in its manuals (Remund et al., 2020a; Remund et al., 2020b).

3.4.4 Meteonorm Future Climate Change Data Generation

The Meteonorm future climate change data generation consists of five steps. The first step is to select the weather station location where the climate change weather file is to be based on. The London Heathrow weather station was selected in this case, as presented in Figure 3.5. The second step is to select the baseline database for generating the future climate change data. There two options provided in Meteonorm, which are employing the embedded climate data in Meteonorm or importing the user's own data, as shown in Figure 3.6. In this work, the climate data from Meteonorm was selected.

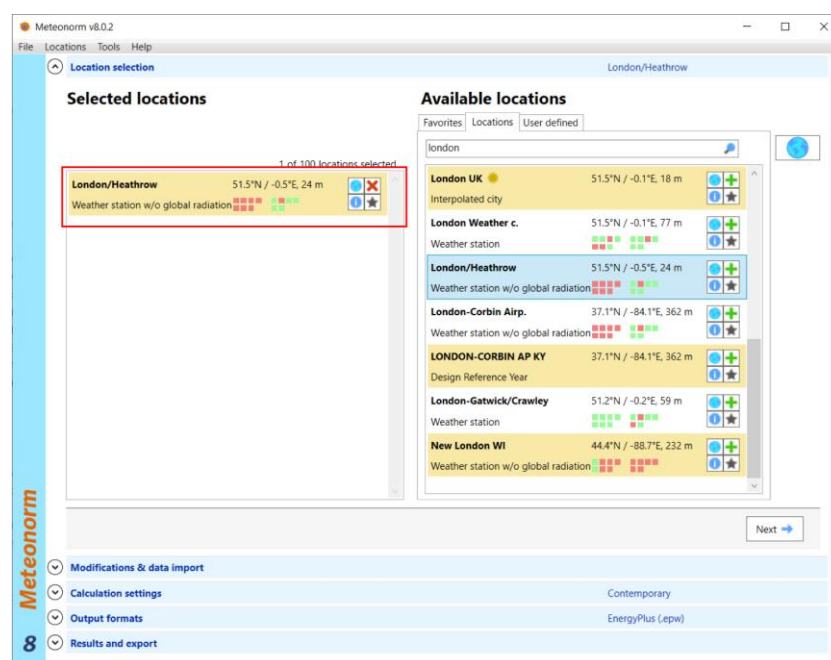


Figure 3.5 Weather station selection for future climate change data generation.

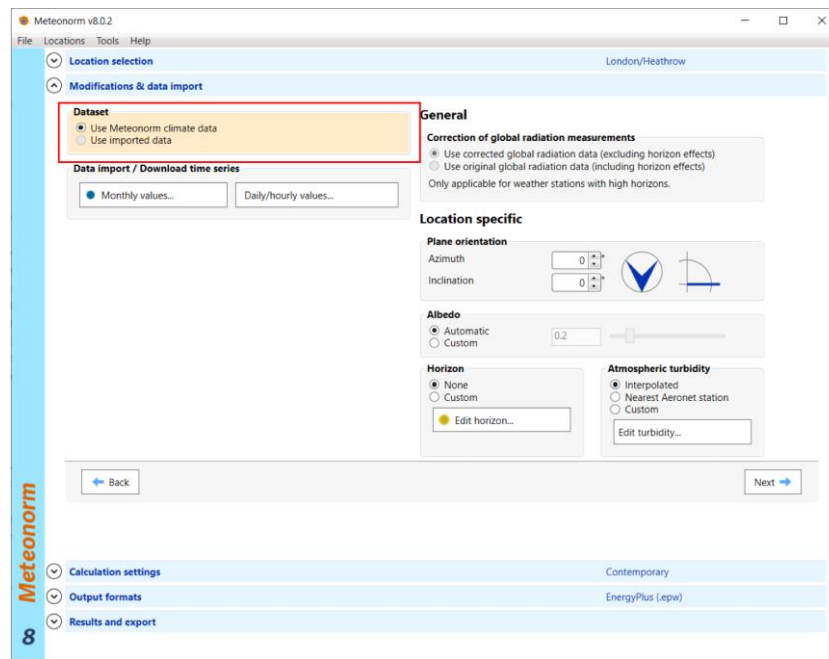


Figure 3.6 Baseline climate database selection for future climate change data generation.

Then, the calculation methods for generating the future climate change data need to be defined. As displayed in Figure 3.7, the program can produce weather data for three periods including historic, contemporary and future periods. The historic-period calculation is based on the radiation data from 1981 – 1990, and temperature, wind and precipitation data from 1961 – 1990. The contemporary-period calculation relies on the radiation data from 1996 – 2015, and temperature, wind and precipitation data from 2000 – 2019. Besides, the program is capable of generating future weather data for every decade from 2030 to 2100. As for the climate change scenarios for future climate data generation, RCP 2.6, RCP 4.5, RCP 8.5 and two RCP scenarios taking into account the urban effects according to the horizon 2020 climate-fit.city project are available. In addition, the radiation models remained default settings and the standard (hour) temperature model was used. The average 10-year monthly extreme values were employed for both temperature and radiation models. Next, the generated weather data can be exported as various building simulation software formats, including the ‘epw’ format used for EnergyPlus in this work (Fig. 3.8). After completing all these settings, the generated weather data can be exported in different time intervals, as presented in Figure 3.9. The hourly data was exported in this case.

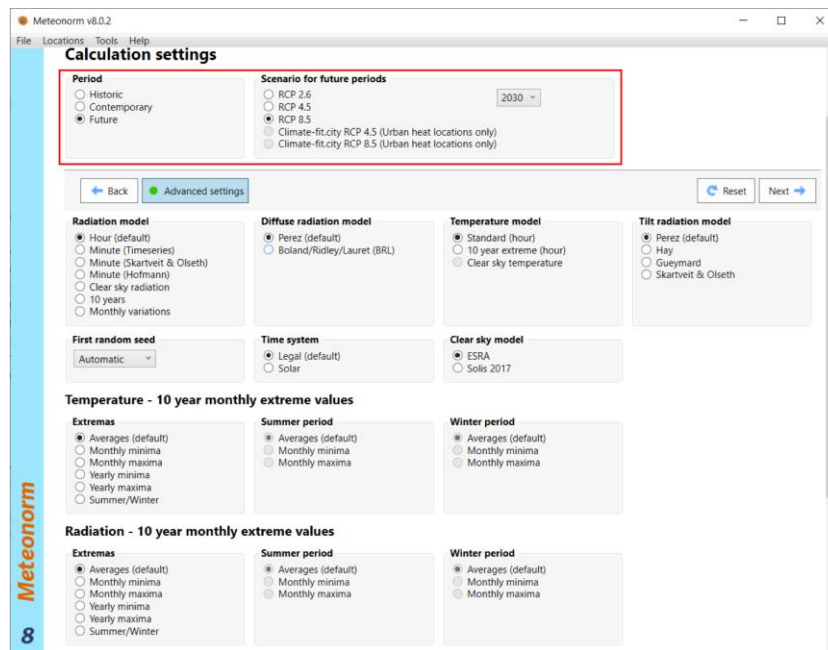


Figure 3.7 Calculation settings for future climate change data generation.

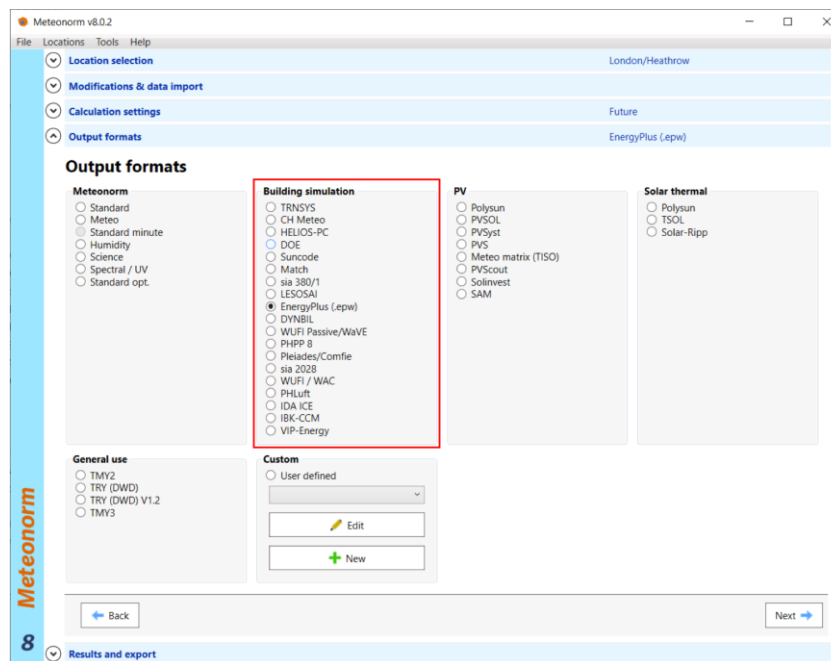


Figure 3.8 Weather data output format for future climate change data generation.

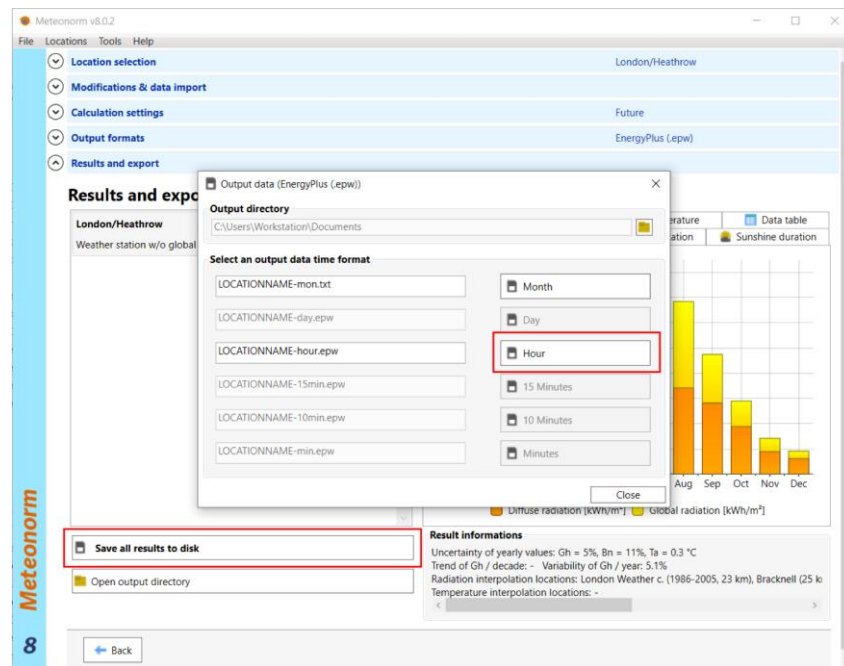


Figure 3.9 Results export for future climate change data generation.

3.5 Fault Simulation

The developed fault models were introduced into the prototype medium-office building, and then the fault impacts on the HVAC system of the prototype building were simulated based on EnergyPlus. The inputs and settings for fault simulations were explained in subsection 3.5.1. The defined outputs from fault simulations were described in subsection 3.5.2.

3.5.1 Fault Simulation Inputs and Settings

The fault intensity for each fault introduced into the building model is presented in Table 3.3. Apart from the control faults, the other faults are assumed to occur in the core zone on the bottom floor or in the component serving the bottom floor to simplify the analysis of fault impacts. In terms of simulation settings, as listed in Table 3.4, six timesteps per hour were used to calculate the zone heat transfer load. The minimum and maximum number of warmup days for convergence were set to 6 and 25, respectively. As for the heat load calculation algorithms, polygon clipping (DOE, 2020b) was used for the shadow calculation. Thermal Analysis Research Program (TARP) (Walton, 1983) was used to calculate the external and internal surface convections. Moreover, the conduction transfer function (DOE, 2020b) was selected as the heat balance algorithm. The details of these governing equations can be found in (DOE, 2020b). Regarding evaluating the thermal comfort of occupants, the widely used Fanger's comfort model (Fanger, 1970) was selected.

Table 3.3. Fault intensity input for each fault model.

Fault name	Fault code	Fault intensity
No overnight setback	NS	1 (binary)
Extended morning setback	ES	2 hours
Delayed evening setback	DS	2 hours
Occupant manual changes of setpoint (heating season)	OMC-H	+2 °C when OA temperature is below -2.5 °C
Occupant manual changes of setpoint (cooling season)	OMC-C	-2 °C when OA temperature is above 26.2 °C
Cooling coil SA temperature sensor bias	CT-P/N	± 2 °C
Heating coil SA temperature sensor bias	HT-P/N	± 2 °C
Economizer OA temperature sensor bias	OAT-P/N	± 2 °C
Economizer RA temperature sensor bias	RAT-P/N	± 2 °C
Thermostat offset	TH-P/N	± 2 °C
Supply fan motor efficiency degradation	SFD	15%
Supply fan stuck at fixed speed	SFmin/50/max	0.25, 0.5, 1
OA damper stuck	OD0/50/100	0, 50%, 100%
OA damper leakage	ODL	40%
VAV reheat coil complete failure	VRCF	1 (binary)
VAV damper leakage	VDL	40%
SA duct leakage	SDL	30%

Table 3.4 Fault simulation EnergyPlus settings.

Number of timesteps per hour	6
Maximum number of warmup days	25
Minimum number of warmup days	6
Shadow calculation method	Polygon clipping
Internal surface convection algorithm	TARP
External surface convection algorithm	TARP
Heat balance algorithm	Conduction transfer function
Thermal comfort model	Fanger

3.5.2 Fault Simulation Outputs

The output reports of EnergyPlus simulations need to be defined before running the simulation, as EnergyPlus only shows the selected outputs after simulations. To represent fault impacts on system energy consumptions and occupant thermal comfort, the following variables were selected as simulation outputs: air system electric energy, air system gas energy (heating coil gas energy in this case as it is the only gas end-use in the air system), air

system fan electric energy, cooling coil electric energy and air system heating coil electric energy (reheat coil electric energy) for each packaged air conditioner as well as predicted percentage dissatisfied (PPD) for each thermal zone on the bottom floor.

Chapter 4 Fault Impact Analysis under Changing Climatic Conditions

4.1 Fault Impact Analysis Overview

To support the development of data-driven FDD methods for HVAC systems, a fault impact analysis is typically needed for the following reasons: Firstly and most importantly, there are usually tens of faults occurring during the lifecycle of an HVAC system, and each of them has a different impact on system energy usages, occupant thermal comfort, operating cost, etc. A fault impact analysis on these faults can help FDD researchers identify the prioritisation of faults for developing an FDD framework since an FDD method may not be able to detect and diagnose all kinds of faults with high detection and diagnosis accuracies. For instance, the predicted accuracy for all the 19 faults covered in the study of Wang, Feng and Liu (2021) varies from 80% to nearly 100%. Thus, maximising the detection and diagnosis accuracies for the faults with substantial impacts and compromising on the accuracies for the faults with considerably less impacts should be an objective for optimising the FDD framework. Secondly, multiple faults existing in an HVAC system is a common case in reality. To investigate the interactions among multiple faults and the combined fault effects, it is critical to perform a fault impact analysis on both single fault and multiple faults cases (Li and O'Neill, 2018). In the present thesis, the purpose of the fault impact analysis is to discover the correlations between fault impacts and climate periods/seasons, and the resulting findings from this chapter stimulate the need for evaluating the developed FDD classifier's generalization ability under different seasons in Chapter 6. This chapter will focus on two main aspects of the fault impact analysis under changing climatic conditions. Section 4.2 will present the fault impact analysis results across different climate periods and discuss the key findings. Section 4.3 will describe fault impact seasonal patterns, and their variations driven by climate change.

4.2 Fault Impact Analysis across Different Climate Periods

The influence of climate change on HVAC system performances has become prominent, and HVAC operational fault impacts may also vary with climate change. This section presents a fault impact assessment for HVAC operational single fault and multiple faults of a VAV system under the contemporary, 2030s and 2050s climate periods based on a validated prototype medium-office building model (Deru et al., 2011). The fault types considered in this assessment and their corresponding fault intensities have been listed in Table 3.3 in Chapter 3.

There are three main contributions to the corresponding domain: (1) The system energy and thermal comfort impact indicators, based on system energy and thermal comfort deviations from the baseline case for each single fault or multiple faults combination, have been proposed to rank each fault case, respectively, under each climate period. Besides, the fault cause and effect analyses were provided for the faults with substantial impacts; (2) The interactions among multiple faults were investigated by comparing actual combined impacts of multiple faults with algebraic summations of the fault impacts of each single fault included in the multiple faults case. The comparison led to two interaction patterns: synergetic effect and antagonistic effect; (3) The single fault and multiple faults impacts variations from the contemporary to 2030s period, and from the contemporary to 2050s period were assessed. It should be noted that this section has been published in the paper ‘Assessment of HVAC system operational fault impacts and multiple faults interactions under climate change’ (Zhong, Calautit and Wu, 2022) and hence the content from the paper is reused in this section.

The structure of this section is as follows: The system energy, and thermal comfort impact indicators used for ranking each single fault or each multiple faults combination can be found in subsection 4.2.1. The fault impacts and rankings for single faults will be discussed in subsection 4.2.2. The rankings for multiple faults combinations and the interactions among them will be described in subsection 4.2.3. The fault impact variation of each single fault or multiple faults combination across different climate periods will be presented in subsection 4.2.4. The comparison of the findings from this section with previous work will be discussed in subsection 4.2.5. The key fault impact analysis findings across different climate periods will be summarised in subsection 4.2.6 and the potential future directions based on the key findings will also be suggested in this subsection.

4.2.1 System Energy and Thermal Comfort Impact Ranking Indicators

Each system total or end energy usage for each fault case E_{faulty} was compared with that from the baseline case E_{normal} to calculate the energy variation percentage δ between each fault case and the normal case:

$$\delta = \frac{E_{faulty} - E_{normal}}{E_{normal}} \quad (4.1)$$

Each fault case was ranked based on the maximum value between the summation of all the variations in system energy usages, and the maximum value among these variations (Eqn. 4.2), and the deviation φ in the mean of the PPDs for all the zones on the bottom floor from

the baseline normal case (Eqn. 4.3), respectively. The reason for using Eqn. 4.2 to calculate the energy impact ranking indicator is not only to highlight fault cases which have insignificant effects on all the energy usages but lead to a significant energy increase in total, but also not to neglect fault cases which lead to decreases in most of the energy usages but result in a substantial increase in one particular energy usage.

$$\delta_{rank} = \max \left\{ \sum_i^n \delta_i, \delta_{max} \right\} \quad (4.2)$$

where δ_{rank} denotes the percentage indicator used for ranking the fault cases; δ_i denotes the percentage variation of each type of system energy usage; δ_{max} denotes the maximum percentage variation among all types of energy usages.

$$\varphi = \frac{1}{n} \left(\sum_{i=1}^n PPD_{i,f} - \sum_{i=1}^n PPD_{i,n} \right) \quad (4.3)$$

where $PPD_{i,f}$ is the PPD for each zone on the bottom floor under each fault case; $PPD_{i,n}$ is the PPD for the same zone under the normal case; n is the number of zones on the bottom floor.

4.2.2 Single Fault Impact and Rankings

A general overview of the rankings for single faults, in terms of energy and thermal comfort impacts, is presented in Table 4.1. Each single fault is ranked based on the proposed energy and thermal comfort indicators from Eqn. (4.2) and (4.3) under the current, 2030s and 2050s periods. According to Table 4.1, the energy and thermal comfort impact ranks for most of the single faults remain the same among the three climate periods. The types of the top 6 faults in the energy impact ranking under three climate periods are the same, but the ranks for some of the top 6 faults vary across the three climate periods. Supply fan stuck at maximum speed, OA damper stuck fully open and thermostat positive offset kept the first, second and sixth in the ranking. However, no overnight setback dropped from third under the current period to fifth under the 2030s and 2050s periods. Cooling coil SA temperature sensor positive bias and supply fan stuck at minimum speed ascended from fifth and fourth under the current period to fourth and third under the two future periods, respectively. Apart from these faults, a substantial variation in the rank was observed from OA damper stuck fully closed which descended from seventh under the current and 2030s periods to tenth under the 2050s period.

Table 4.1 System energy and thermal comfort impact rankings for each single fault under the current, 2030s and 2050s periods.

<i>Fault code</i>	<i>Fault type</i>	<i>Energy ranking</i>			<i>Thermal comfort ranking</i>		
		<i>Current</i>	<i>2030s</i>	<i>2050s</i>	<i>Current</i>	<i>2030s</i>	<i>2050s</i>
NS	No overnight setback	3	5	5	11	6	5
ES	Extended morning setback	22	23	23	18	16	16
DS	Delayed evening setback	13	13	13	9	9	9
OMC-H	Occupant manual heating setpoint change	16	16	16	20	20	20
OMC-C	Occupant manual cooling setpoint change	9	10	11	5	5	5
CT-P	Cooling coil SA temperature sensor positive bias	5	4	4	4	4	4
CT-N	Cooling coil SA temperature sensor negative bias	18	18	17	24	24	24
HT-P	Heating coil SA temperature sensor positive bias	14	14	14	6	8	8
HT-N	Heating coil SA temperature sensor negative bias	10	9	8	21	21	21
OAT-P	Economizer OA temperature sensor positive bias	19	19	19	12	12	12
OAT-N	Economizer OA temperature sensor negative bias	21	21	21	15	17	18
RAT-P	Economizer RA temperature sensor positive bias	23	22	22	19	19	19
RAT-N	Economizer RA temperature sensor negative bias	25	25	25	13	13	13
TH-P	Thermostat positive offset	6	6	6	2	2	2
TH-N	Thermostat negative offset	15	15	15	23	23	23
SFD	Supply fan motor efficiency degradation	17	17	18	14	14	14
SFmin	Supply fan stuck at min speed	4	3	3	25	25	25
SF50	Supply fan stuck at 50% speed	12	12	12	3	3	3
SFmax	Supply fan stuck at max speed	1	1	1	1	1	1
OD0	OA damper stuck fully closed	7	7	10	7	7	7
OD50	OA damper stuck 50% position	8	8	7	10	10	10
OD100	OA damper stuck fully open	2	2	2	8	11	11
ODL	OA damper leakage	20	20	20	17	17	17
VRCF	VAV reheat coils complete failure	24	24	24	15	15	15
SDL	SA duct leakage	11	11	9	22	22	22

In terms of the thermal comfort impact ranking, eight out of the top 10 faults remain the same ranks while heating coil SA temperature sensor positive bias and OA damper stuck fully open

descended from sixth and eighth under the current period to eighth and eleventh under the two future periods, respectively. Besides, no overnight setback showed the largest increase in the rank from eleventh under the current period to sixth under the 2030s period.

4.2.2.1 Cause and Impact Analysis for Prioritized Single Faults

The energy and thermal comfort impacts of the top single faults in Table 4.1 under the current, 2030s and 2050s period are shown in Figure 4.1–4.3, respectively. The deviations of system energy end usages from the baseline normal case for the top 7 faults are illustrated in Figure 4.1a, 4.2a and 4.3a. The deviations of the PPD from the baseline normal case for the top 10 faults are presented in Figures 4.1b, 4.2b and 4.3b. Under the current period, most single faults demonstrate large discrepancies in energy impacts but show slight deviations in terms of the influence on thermal comfort, except for the supply fan stuck at maximum speed. This fault led to the most significant impact on system energy end usages and thermal comfort. Due to the largest requirement of SA flow rate, which was achieved by increasing the amount of both return and outdoor air in the mixed air (Fig. 4.4a), it caused substantial increases in electricity usage of the air system total, reheating coils, cooling coil and fan by 109.72%, 111.41%, 40.49% and 226.95% (Fig. 4.1a), respectively. The only decrease for this fault occurred for the heating coil gas energy by 22.53% (Fig. 4.1a). In addition, the PPD was 9.70% higher than that of the baseline normal case (Fig. 4.1b). On the contrary, supply fan stuck at minimum speed led to decreases of 12-16% in electricity usages but resulted in a considerable increase of 104.10% in heating coil gas energy (Fig. 4.1a). This effect was attributed to the use of full OA as the SA at the minimum speed of the supply fan (Fig. 4.4b).

Apart from the supply fan stuck faults, no overnight thermostat setback indicated significant influences on the air system total and reheating coils electricity usages by +30.68% and +64.06%, respectively, but only caused a slight increase of 1.82% in cooling coil electricity usage (Fig. 4.1a). The difference in the electricity usage between the cooling coil and reheating coils can be explained by the setpoint difference of 2.7 °C and - 5.4 °C between the unoccupied and occupied periods for cooling and heating, respectively (Table 3.2). In addition, another factor for this difference is that the summer OA temperature in the evening is typically significantly lower than the thermostat setpoints in both occupied and unoccupied periods (Fig. 4.4c). Moreover, this fault increased the room air temperature during the

unoccupied period, and thus the heating coil gas energy decreased slightly by 2.44% (Fig. 4.1a).

As for the OA damper stuck faults, OA damper stuck fully open resulted in an enormous increase of 230.22% in heating coil gas energy (Fig. 4.1a) due to the low winter OA temperature. In contrast, the OA damper stuck fully closed led to a large rise of 51.37% in cooling coil electric energy (Fig. 4.1a) since the lower summer OA temperature than the zone air temperature in the morning of summer days or during the whole day of some summer days can contribute to the reduction in the mixed air temperature for the packaged air conditioners (Fig.4.4c).

With respect to sensor faults, the positive bias in the cooling coil SA temperature sensor resulted in a lower actual temperature of SA leaving the cooling coil, and thus caused a rise in cooling coil electric energy of 15.55% (Fig. 4.1a). Meanwhile, this lower SA temperature led to a significant increase in heating coil gas energy of 70.72% (Fig. 4.1a) since cooling is turned on in the heating season to mitigate the overheating effect on zones (Fig. 4.4d). In addition, this fault slightly affected the occupant thermal comfort in the cooling season as the PPD increased by 1.06% (Fig. 4.1b).

As for thermostat positive offset, it showed a substantial effect on occupant thermal comfort (Fig. 4.1b). The main reason for the effect is that the thermostat reading reached the zone air temperature setpoint when the actual zone air temperature was lower than the setpoint. As a result, the temperature of SA leaving the heating coil during the heating season (Fig. 4.4e) and that leaving the cooling coil during the cooling season (Fig. 4.4f) were both lower than the actual required SA temperature, which led to the decrease in heating coil gas energy of 6.84% and the increase in cooling coil electric energy of 9.57% (Fig. 4.1a), respectively. To compensate for the less heating effect of SA from the packaged air conditioner, the SA flow rate was increased in the heating season (Fig. 4.4e). Moreover, the SA flow rate in the cooling season was also increased (Fig. 4.4f) to reduce the actual air temperature of the faulty zone to ensure that the positive-deviated reading of the thermostat meets the zone air temperature setpoint.

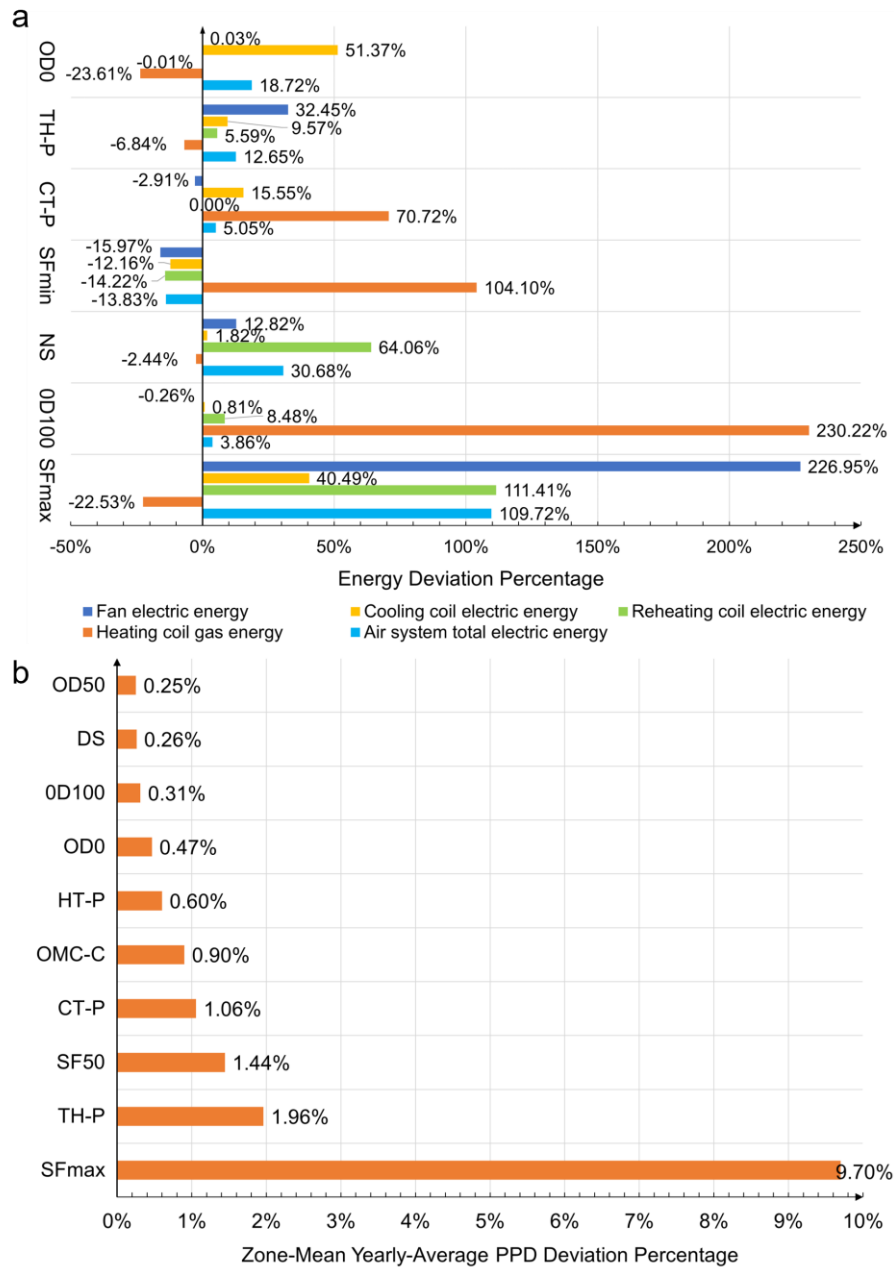


Figure 4.1. Single fault impacts under the current climate period: (a) System energy usage deviations; (b) PPD deviations for the top-ranked faults.

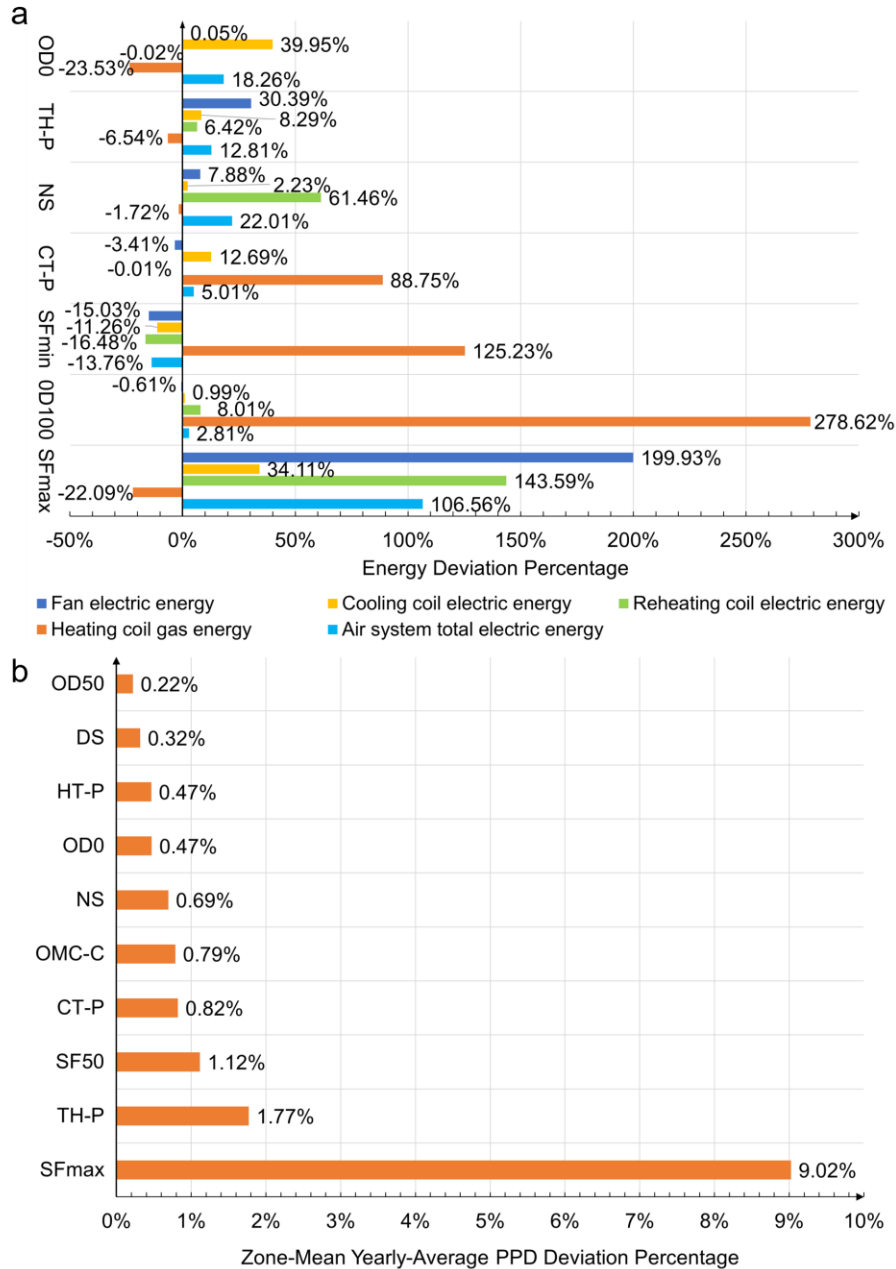


Figure 4.2. Single fault impacts under the 2030s climate conditions: (a) System energy usage deviations; (b) PPD deviations for the top-ranked faults.

Under the 2030s and 2050s period, similar trends in energy end usages for the top 7 faults can still be observed in Figure 4.2a and 4.3a. Supply fan stuck at maximum speed still resulted in the most significant energy impact under both periods (Fig. 4.2a and 4.3a). Compared with the ranks under the current period, the rank ascending under both periods for supply fan stuck at minimum speed and cooling coil SA temperature sensor positive bias, as presented in Table 4.1, were mainly due to their intensified impacts on heating coil gas energy (Fig. 4.2a and 4.3a). Apart from these, the rank descending for no overnight setback, as presented in Table

4.1, was caused by the reduced effect on the fan and reheat coil electric energy (Fig. 4.2a and 4.3a). As for the only difference among the top 7 faults in the energy impact ranking between the current and 2050s periods, the OA damper stuck at 50% position resulted in large deviations in heating coil gas energy and cooling coil electric energy from the baseline normal case (Fig. 4.3a). Regarding the PPD for the top 7 faults, the most significant impact of 9.02% under both periods was also led by supply fan stuck at maximum speed (Fig. 4.2b and 4.3b).

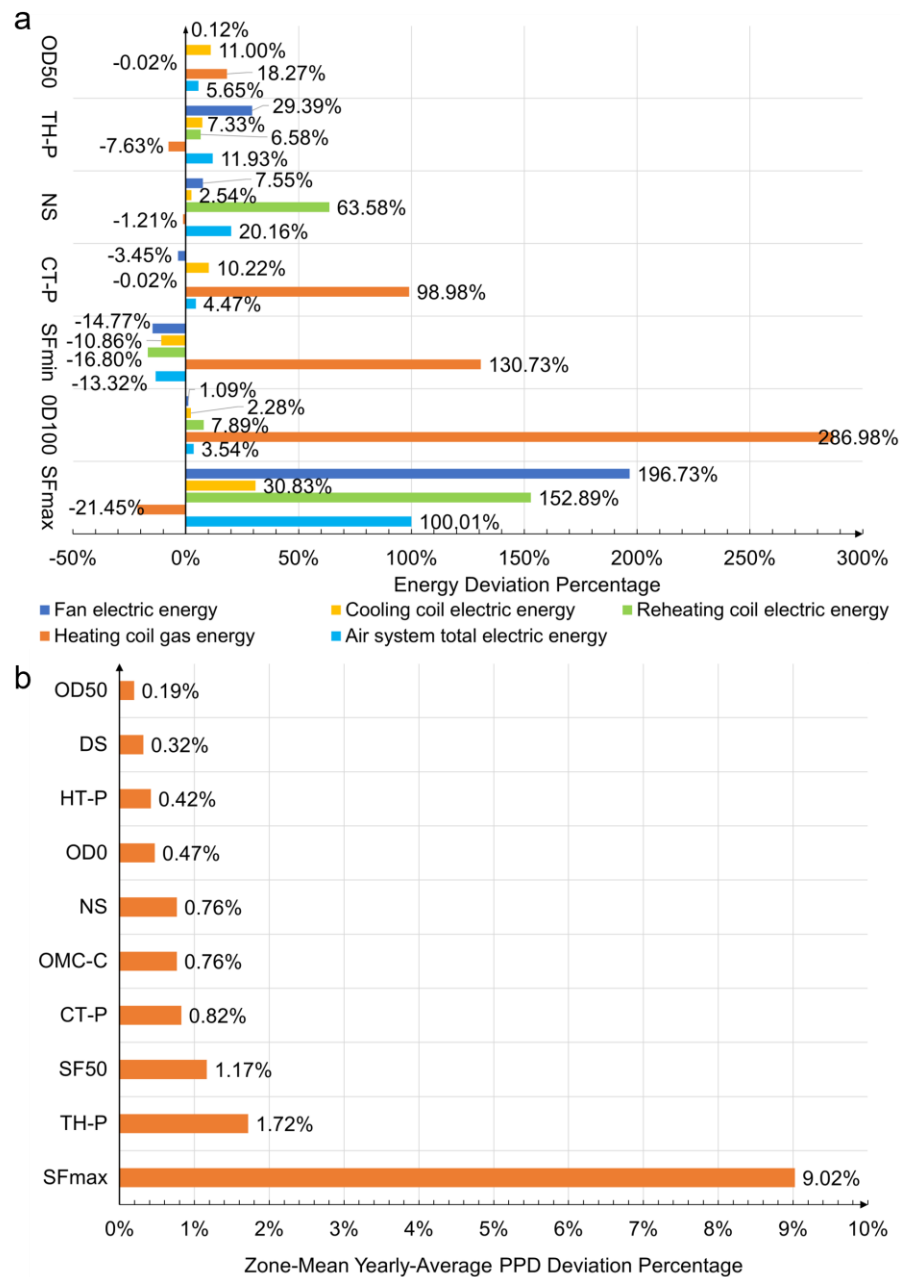


Figure 4.3. Single fault impacts under the 2050s climate period: (a) System energy usage deviations; (b) PPD deviations for the top-ranked faults.

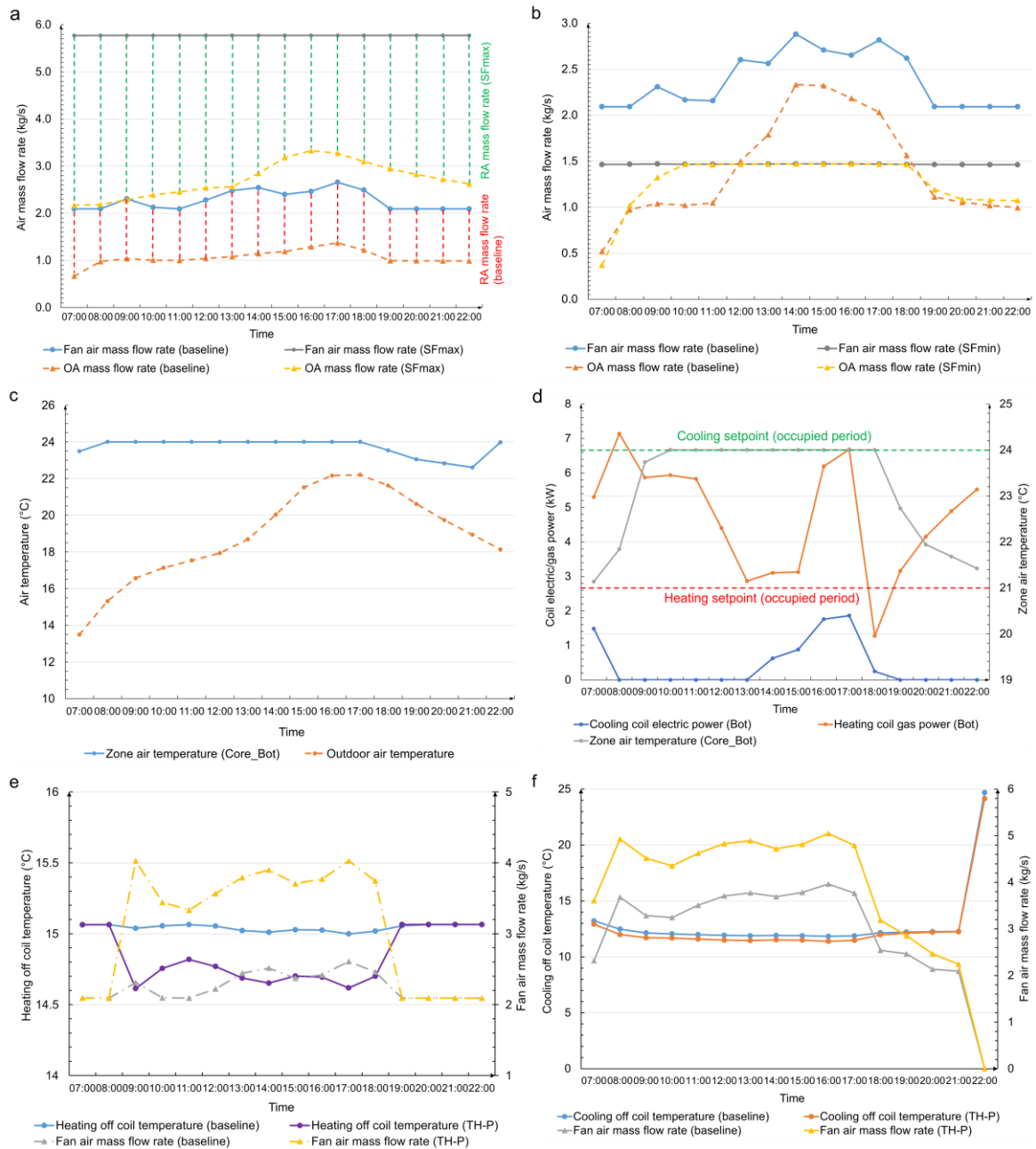


Figure 4.4. Causes for the impacts of single faults: (a) Supply fan stuck at maximum speed; (b) Supply fan stuck at minimum speed; (c) No overnight setback and OA damper stuck fully closed; (d) Cooling coil SA temperature sensor positive bias; (e-f) Thermostat positive offset.

4.2.3 Multiple Faults Impact and Rankings

A general overview of the rankings for each multiple faults combination, in terms of energy and thermal comfort impacts, are listed in Table 4.2. Each fault combination is ranked based on the proposed energy and thermal comfort indicators from Eqn. (4.2) and (4.3) under the current, 2030s and 2050s periods. The differences between the actual combined effects and the algebraic sums were found to be within two categories: synergetic effect and antagonistic

effect, which are discussed in subsections 4.2.3.1 and 4.2.3.2. The synergetic effect represents the phenomenon that the actual combined impact of a fault combination is greater than the algebraic sum of the impact of each single fault within this combination. On the contrary, the antagonistic effect means that the actual combined impact of a fault combination is smaller than the algebraic sum of the impact of each single fault within this combination. Based on Table 4.2, most multiple faults combinations remain in the same ranks across climate periods. The types of the top 5 combinations in both energy and thermal comfort rankings keep the same, but the ranks of the top 5 combinations in terms of thermal comfort vary significantly from the current and 2030s periods to 2050s period. The combination of supply fan stuck at maximum speed and OA damper stuck fully open, ranked second under the current and 2030s periods, dropped to fifth under the 2050s period. Subsequently, under the 2050s period, the combinations of supply fan stuck at maximum speed and OA damper stuck fully closed, thermostat positive offset and supply fan stuck at maximum speed, and no overnight setback and supply fan stuck at maximum speed ascended by one rank from their third, fourth and fifth place under the current and 2030s periods.

Table 4.2. System energy and thermal comfort impact rankings for each multiple faults combination under the current, 2030s and 2050s periods.

<i>Fault type</i>	<i>Energy ranking</i>			<i>Thermal comfort ranking</i>		
	<i>Current</i>	<i>2030s</i>	<i>2050s</i>	<i>Current</i>	<i>2030s</i>	<i>2050s</i>
No overnight setback+Cooling coil SA temperature sensor positive bias	9	9	9	12	10	10
No overnight setback+Supply fan stuck at min speed	15	16	15	15	15	15
No overnight setback+Supply fan stuck at max speed	3	3	3	5	5	4
No overnight setback+OA damper stuck fully closed	14	17	17	13	12	11
No overnight setback+OA damper stuck fully open	5	5	5	14	14	13
Cooling coil SA temperature sensor positive bias+Supply fan stuck at min speed	17	15	13	16	17	17
Cooling coil SA temperature sensor positive bias+Supply fan stuck at max speed	2	2	2	1	1	1
Cooling coil SA temperature sensor positive bias+OA damper stuck fully closed	13	12	12	10	11	12
Cooling coil SA temperature sensor positive bias+OA damper stuck fully open	8	8	8	11	13	14
Supply fan stuck at min speed+OA damper stuck fully closed	19	19	19	17	18	18
Supply fan stuck at min speed+OA damper stuck fully open	12	10	10	18	19	19
Supply fan stuck at max speed+OA damper stuck fully closed	4	4	4	3	3	2
Supply fan stuck at max speed+OA damper stuck fully open	1	1	1	2	2	5
No overnight setback+Thermostat positive offset	11	14	16	6	6	6
Cooling coil SA temperature sensor positive bias+Thermostat positive offset	10	11	11	7	7	7
Thermostat positive offset+Supply fan stuck at min speed	16	13	14	19	16	16
Thermostat positive offset+Supply fan stuck at max speed	6	6	6	4	4	3
Thermostat positive offset+OA damper stuck fully closed	18	18	18	8	8	8
Thermostat positive offset+OA damper stuck fully open	7	7	7	9	9	9

4.2.3.1 Multiple Faults Synergetic Effect

The differences between actual combined effects and algebraic sums, in terms of system energy usages and occupant thermal comfort, for the top 5 combinations under the current period are illustrated in Figures 4.5a and 4.5b, respectively. The synergetic effect, which means that the actual combined effect is larger than the algebraic sum, is indicated in system energy usages of the top 5 combinations. Under the current period, some cases demonstrate that the actual combined effects on energy usages are significantly higher than the algebraic sums. The combination of cooling coil SA temperature sensor positive bias and supply fan stuck at maximum speed showed an enormous increase of 135.17% from the algebraic sum in heating coil gas energy (Fig. 4.5a). It should be noted that the two single faults led to an increase of 70.72% and a decrease of 22.53% in heating coil gas energy (Fig. 4.1a), respectively. One of the main reasons for this phenomenon is that cooling coil SA temperature sensor positive bias caused SA temperature difference of 1 °C to 2 °C across the heating coil (Fig. 4.6a), which was zero when only supply fan stuck at maximum speed occurred since the large portion of RA in the mixed air offsets the need for heating the SA. Another reason is that the SA flow rate was increased considerably (Fig. 4.6a) due to the maximum operating speed of the supply fan.

Similarly, substantial increases of 90.47% and 68.51% from the algebraic sum in heating coil gas energy were led by the combinations of no overnight setback and OA damper stuck fully open, and supply fan stuck at maximum speed and OA damper stuck fully open (Fig. 4.5a). Although no overnight setback and supply fan stuck at maximum speed individually resulted in decreases in heating coil gas energy (Fig. 4.1a), the SA temperature difference across the heating coil was increased substantially (Fig. 4.6b-4.6c) when OA damper stuck fully open occurred. As a result, the actual combined effect of these combinations on heating coil gas energy was considerable.

Apart from the large increases from the algebraic sum in heating coil gas energy, the combination of supply fan stuck at maximum speed and OA damper stuck fully open also resulted in a significant increase of 82.26% from the algebraic sum in reheating coil electric energy (Fig. 4.5a). The large deviation can be attributed to the significantly increased zone SA temperature difference across reheating coils caused by both faults and the considerably increased zone SA flow rate due to the increased SA flow rate from the packaged air

conditioner (Fig. 4.6d). Besides, the combination of no overnight setback and supply fan stuck at maximum speed led to a higher variation of 29.34% than the algebraic sum for air system total electric energy, as a result of the increases from the algebraic sum of 49.38% and 40.52% in reheating coil and fan electric energy, respectively.

With respect to electricity consumed by cooling coils, a large increase of 22.82% from the algebraic sum in cooling coil electricity usage was caused by the combination of cooling coil SA temperature sensor positive bias and supply fan stuck at maximum speed (Fig. 4.5a). This increase was due to the increased SA flow rate caused by the supply fan stuck at maximum speed and the increased SA temperature difference across the cooling coil led by the cooling coil SA temperature sensor positive bias (Fig. 4.6e).

Based on variations in the synergetic effects for system energy usages across the three climate periods, as presented in Figure 4.7a-4.7e, the synergetic effect on heating coil gas energy and reheating coil electric energy for several fault combinations is intensified with the transition from the current to the 2050s period. The most significant increase from the algebraic sum in heating coil gas energy was caused by the combination of cooling coil SA temperature sensor positive bias and supply fan stuck at maximum speed, which increased from 135.17% under the current period to 151.56% under the 2030s period, and to 163.67% under the 2050s period (Fig. 4.7b). Similarly, the increase from the algebraic sum in heating coil gas energy for the combination of supply fan stuck at maximum speed and outdoor air damper stuck fully open grew from 68.51% under the current period to 72.73% under the 2030s period, and further to 76.38% under the 2050s period (Fig. 4.7b). However, a special case was the combination of no overnight setback and outdoor air damper stuck fully open, which led to a decrease in the deviation from the algebraic sum in heating coil gas energy from +90.47% under the current period to +83.23% under the 2030s period, but resulted in a decrease to +85.20% under the 2050s period (Fig. 4.7b).

In addition, the increase from the algebraic sum in reheating coil electric energy for the combination of supply fan stuck at maximum speed, and OA damper stuck fully open was intensified across the three climate periods from 82.26% under the current period to 103.90% under the 2030s period and 105.97% under the 2050s period (Fig. 4.7c). Moreover, the combination of no overnight setback and supply fan stuck at maximum speed also led to an

increasing trend in the deviation from the algebraic sum in reheating coil electric energy across climate periods from +49.38% to +53.23% and +56.60% (Fig. 4.7c).

On the other hand, the synergetic effect on the air system total, cooling coil and fan electric energy is reduced across the three climate periods (Fig. 4.7a, 4.7d, 4.7e). The deviations from the algebraic sums in the air system total electric energy for all the top 5 combinations decreased across the three climate periods (Fig. 4.7a). The most drastic decrease in the deviation from the algebraic sum of air system total electricity was caused by the combination of supply fan stuck at maximum speed and OA damper stuck fully open from +36.02% to +33.76%, and to +29.68% (Fig. 4.7a).

Besides, the decrease in the deviation from the algebraic sum was also observed from the cooling coil electric energy for the combination of supply fan stuck at maximum speed and OA damper stuck fully closed, and cooling coil SA temperature sensor positive bias and supply fan stuck at maximum speed. The deviation from the algebraic sum for the combination of supply fan stuck at maximum speed and OA damper stuck fully closed dramatically reduced from +46.02% to +33.71%, and +25.94% across the three climate periods (Fig. 4.7d). Compared with this combination, the reductions across the three climate periods for the combination of cooling coil SA temperature sensor positive bias and supply fan stuck at maximum speed were less intense, which were from +22.82% to +17.53%, and to +13.29% (Fig. 4.7d).

In addition, the decrease in the deviation from the algebraic sum was indicated in the fan electric energy for the combination of no overnight setback and supply fan stuck at maximum speed. The deviation of this combination drastically reduced from +40.52% under the current period to +29.21% under the 2030s period, and slightly decreased to +28.15% under the 2050s period (Fig. 4.7e). As for the synergetic effect on thermal comfort, the variation from the algebraic sum in the PPD for the combination of cooling coil SA temperature sensor positive bias and supply fan stuck at maximum speed increased from +1.84% under the current period to +2.07% under 2030s period, and then decreased to +1.86% under 2050s period (Fig. 4.7f).

4.2.3.2 Multiple Faults Antagonistic Effect

The differences between actual combined effects and algebraic sums, in terms of system energy usages, for the fault combinations with severe antagonistic effects under the current period are shown in Figure 4.8. In contrast to the synergetic effect, the antagonistic effect, which represents a less actual combined effect than the algebraic sum, is also presented in energy end usages of some multiple faults combinations. Under the current period, the combinations of supply fan stuck at minimum speed with OA damper stuck fully closed, and with OA damper stuck fully open indicated substantial decreases of 104.45% and 166.97% from the algebraic sums in heating coil gas energy, respectively (Fig. 4.8). The supply fan stuck at minimum speed and OA damper stuck fully open individually caused significant increases in heating coil gas energy of 104.10% and 230.22% (Fig. 4.1a), while OA damper stuck fully closed individually led to a decrease in heating coil gas energy of 23.61% (Fig. 4.1a). A common reason for them was that the minimum operation speed of the supply fan limited OA's intake, thus reducing the need for heating the mixed air (Fig. 4.6f-4.6g). In addition, the fully closed OA damper led to the full RA ratio for the mixed air, therefore minimising the SA temperature difference across the heating coil to zero (Fig. 4.6f).

As for the antagonistic effects on other system energy end usages, the largest decrease of 38.02% from the algebraic sum in reheating coil electric energy was caused by the combination of thermostat positive offset and supply fan stuck at maximum speed (Fig. 4.8). As shown in Fig. 4.6h, the supply fan stuck at maximum speed induced large SA temperature difference across the reheat coil, but thermostat positive offset minimised the need for reheating the SA to zero and thus led to this large decrease. Apart from this, the combination of thermostat positive offset and supply fan stuck at minimum speed also contributed to a similar trend that the actual air system total, reheating coil, cooling coil and fan electricity usages decreased from the algebraic sums of those for each single fault by 12.86%, 7.11%, 8.67% and 31.90% (Fig. 4.8). In terms of thermal comfort, the antagonistic effects were presented in the three out of the top 5 combinations. The most significant antagonistic effect occurred at the combination of thermostat positive offset and supply fan stuck at maximum speed with a decrease of 1.81% from the algebraic sum (Fig. 4.5b).

The variations in the antagonistic effects on system energy usages across the three climate periods for these combinations are plotted in Figure 4.7g-4.7j. The most remarkable effects are

presented on heating coil gas energy for several fault combinations. The largest decrease of 212.55% from the algebraic sum (Fig. 4.7g) was triggered by the combination of supply fan stuck at minimum speed and OA damper stuck fully open under the 2050s period, which was intensified from 166.97% under the current period (Fig. 4.7g). Similarly, the decreases from the algebraic sums for the combinations of supply fan stuck at minimum speed and OA damper stuck fully closed, cooling coil SA temperature sensor positive bias and OA damper stuck fully open, as well as cooling coil SA temperature sensor positive bias and supply fan stuck at minimum speed were also boosted from 104.45% to 125.71% and further to 131.24%, from 46.68% to 54.96% and further to 56.37%, and from 50.40% to 61.79% and further to 66.31%, respectively across the three climate periods (Fig. 4.7g).

In contrast, the deviations from the algebraic sums in cooling coil electric energy for all the fault combinations, except the combination of supply fan stuck at minimum speed and OA damper stuck fully open, were mitigated across the three climate periods (Fig. 4.7i). Among these combinations, the combination of supply fan stuck at minimum speed and OA damper stuck fully closed resulted in the most drastic reduction in the decreases from the algebraic sum from 10.18% to 7.70% and 5.26% across the three climate periods.

Moreover, considerable enhancement in the antagonistic effect on reheating coil electric energy was indicated by the combination of thermostat positive offset and supply fan stuck at maximum speed with decreases from 38.02% to 50.56% and further to 54.09% across the three climate periods (Fig. 4.7h). Regarding the air system's total electric energy, it can be obviously observed that the variations in the antagonistic effects for all the fault combinations are insignificant (Fig. 4.7j). Apart from these, an obvious difference among the climate periods is that the synergetic effect on PPD for the combination of supply fan stuck at max speed and OA damper stuck fully open decreased from + 0.84% under current period to +0.31% under 2030s period, and further being converted to the antagonistic effect of -0.44% period (Fig. 4.7f).

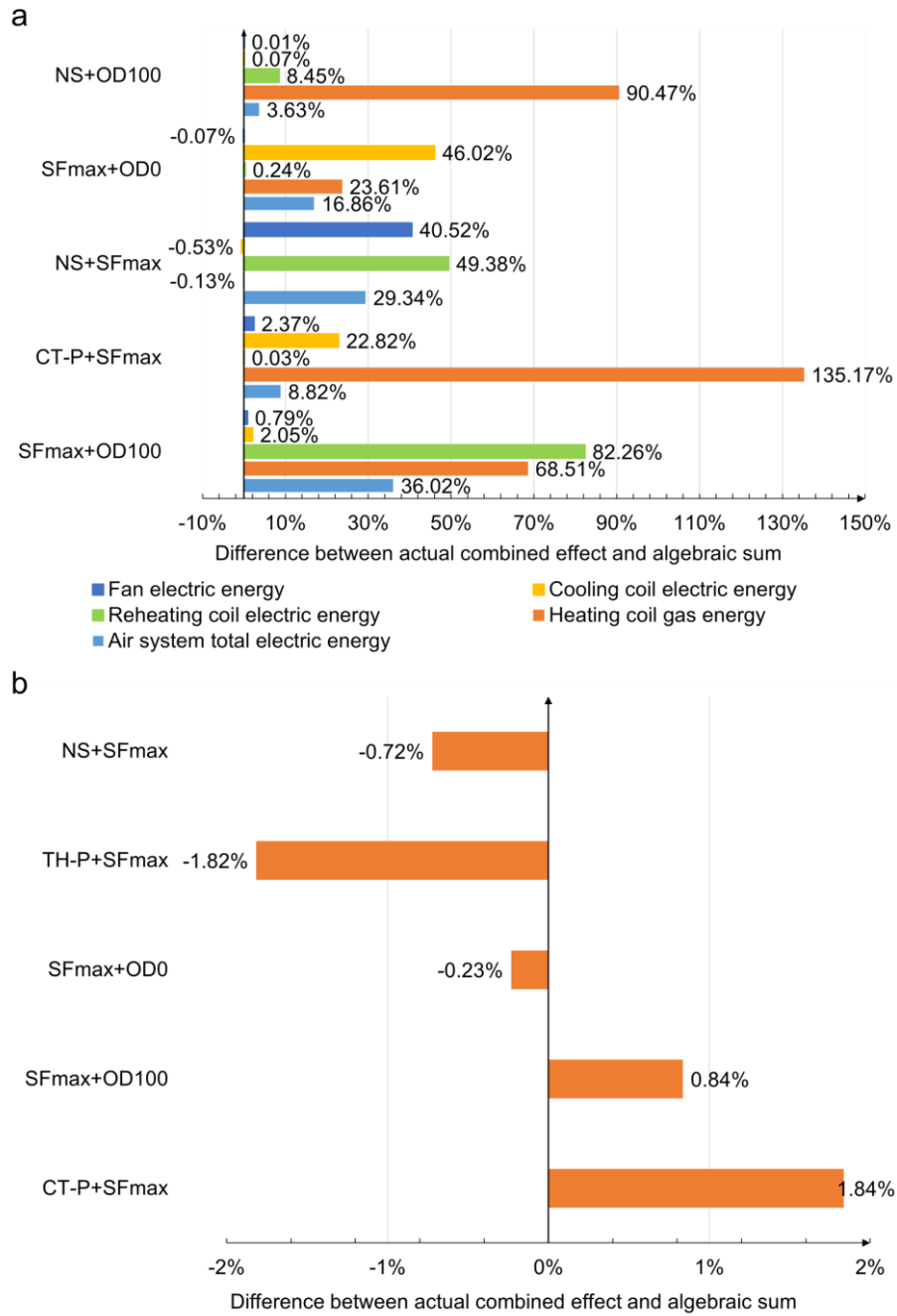


Figure 4.5. Multiple faults interactions for the top 5 combinations under the current climate period: (a) Difference between actual combined effect on (a) system energy consumptions; (b) PPD, and the algebraic sums of each single-fault impact.

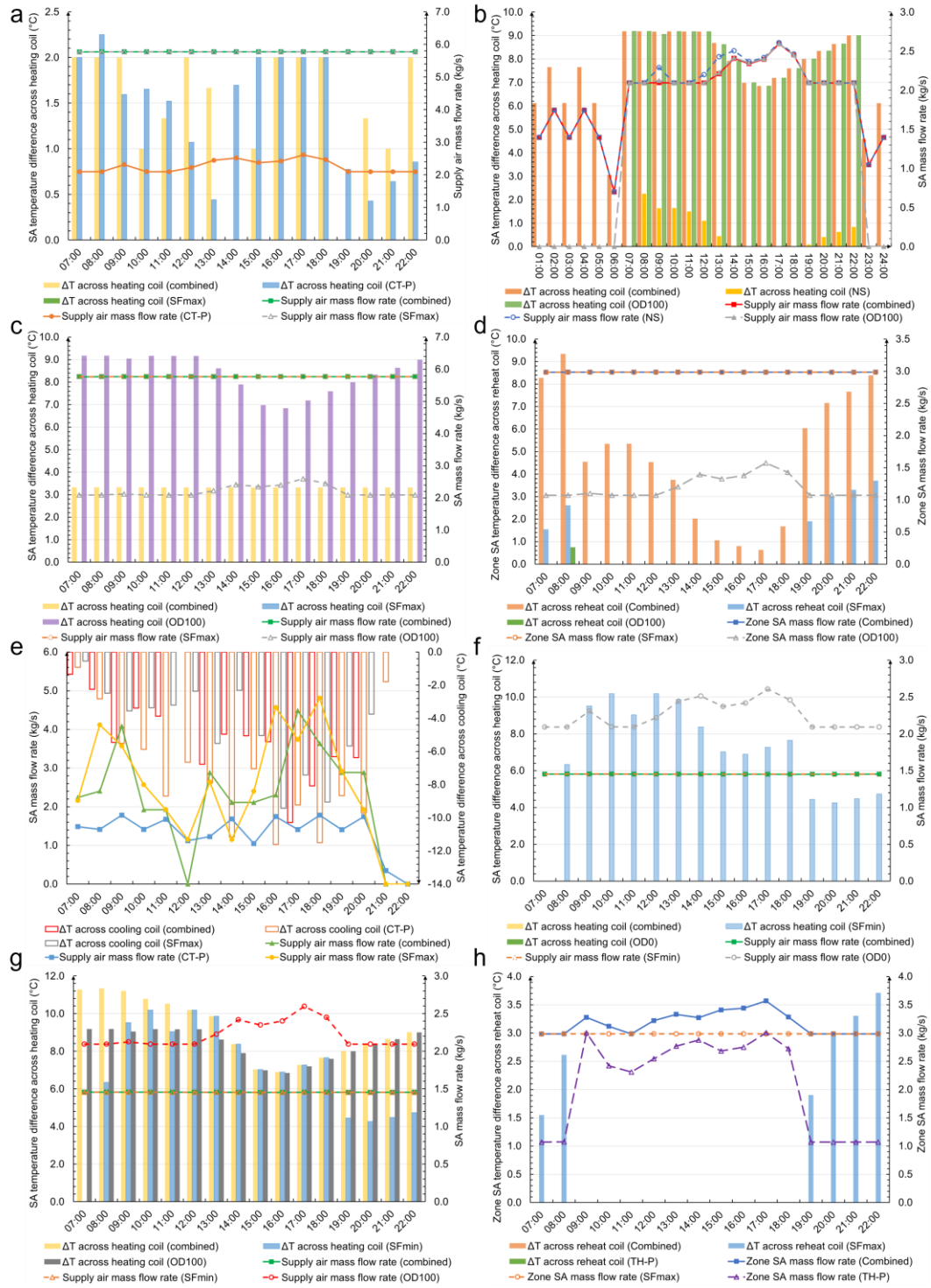


Figure 4.6. Causes for multiple faults (a-e) synergetic effects and (f-h) antagonistic effects.

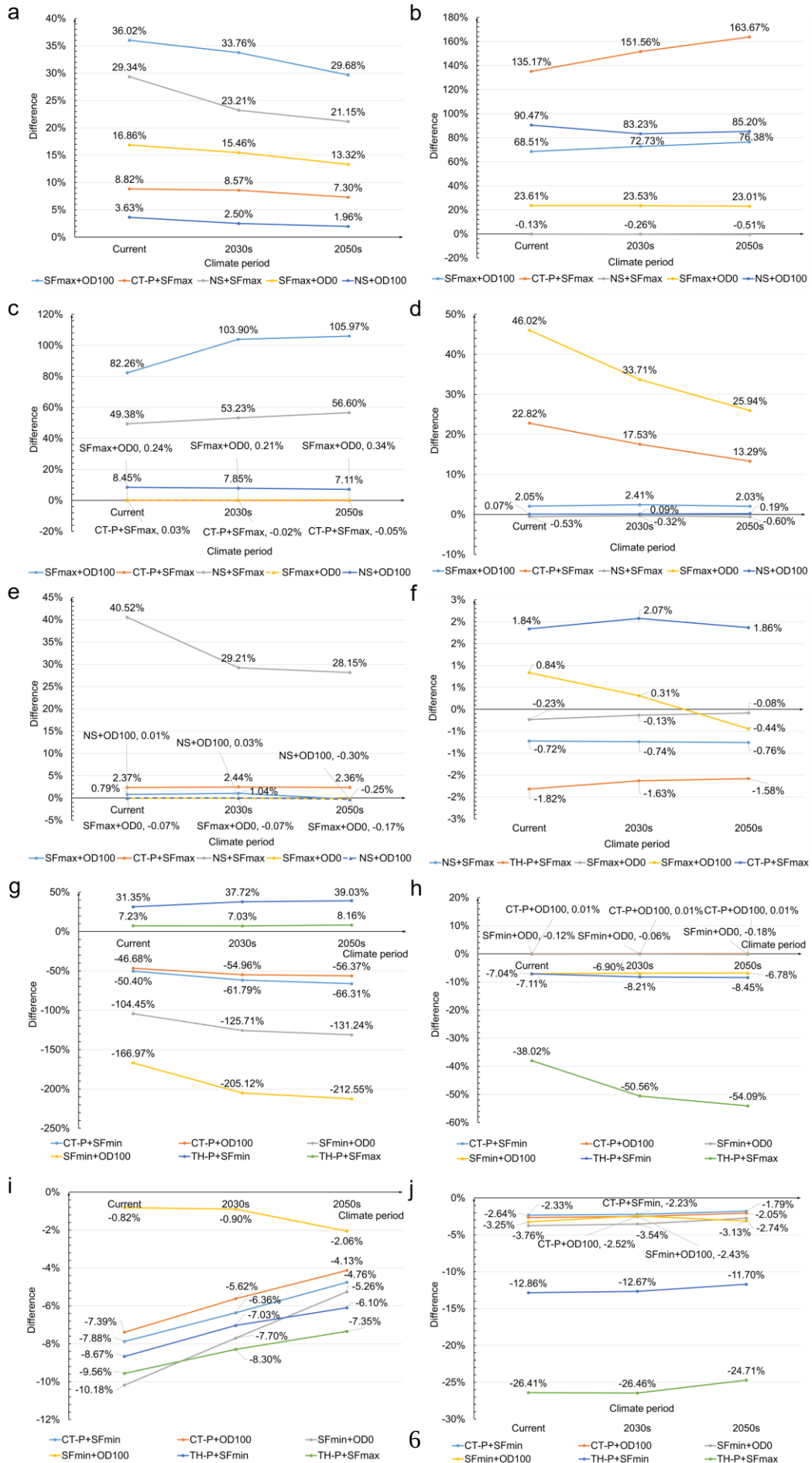


Figure 4.7. Variations across climate periods in the difference between actual combined effect and algebraic sum (‘Difference’ in the figures) of (a) air system total electric energy; (b) heating coil gas energy; (c) reheating coil electric energy; (d) cooling coil electric energy; (e) fan electric energy; (f) PPD for the top 5 fault combinations; (g) heating coil gas energy (h) reheating coil electric energy; (i) cooling coil electric energy; (j) air system total electric energy for the fault combinations with severe antagonistic effects.

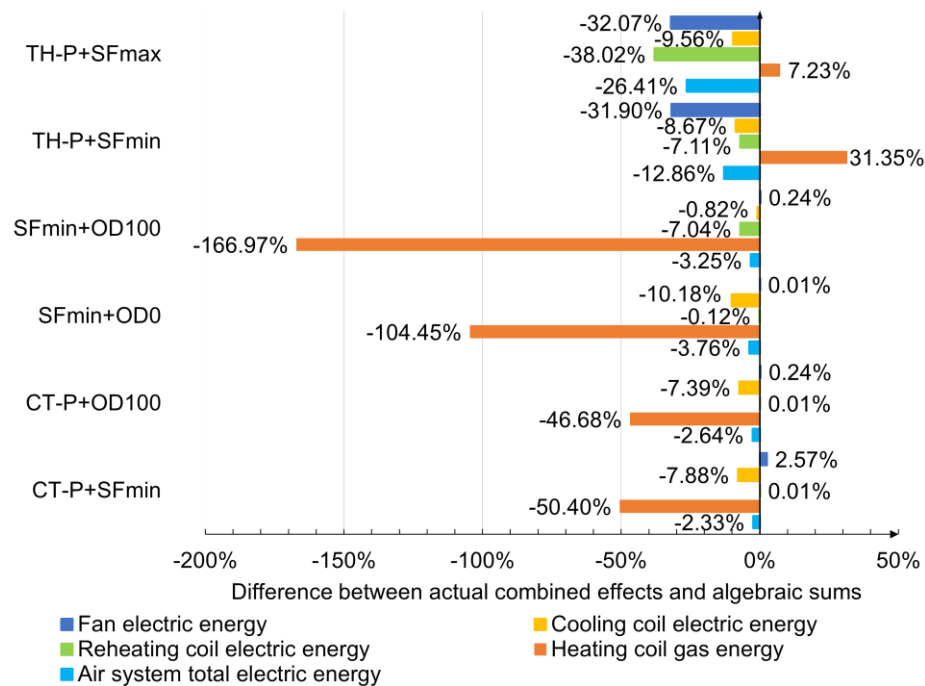


Figure 4.8. The antagonistic effects of multiple faults combinations on system energy consumptions under the current climate period.

4.2.4 Fault Impact Variations across Different Climate Periods

This subsection provides the variations of single and multiple faults impacts on system energy end usages and occupant thermal comfort from the current to the 2030s period, and from the current to the 2050s period, respectively. The impact variations of each single fault across these periods are presented in 4.2.4.1 and the impact variations of each multiple faults combination across these periods are shown in 4.2.4.2.

4.2.4.1 Single Fault Impact Variations

The variations of energy end usages and the PPD for each single fault between the 2030s and current periods, and between the 2050s and current periods are presented in Figure 4.9. The fault codes presented in each bar chart refer to those in Table 4.1. Overall, the cooling coil

and fan electric energy for all single faults increased by the ranges of 29.5 to 35.5 GJ (Fig. 4.9b) and 5.4 to 9.6 GJ (Fig. 4.9e), respectively, from the current to 2030s period and grown by the range of 56.3 to 72.0 GJ (Fig. 4.9b) and 6.9 to 11.5 GJ (Fig. 4.9e), respectively, from the current to 2050s period. In contrast, the heating coil gas energy and reheating coils electric energy for all single faults reduced by the range of 8.5 to 13.7 GJ (Fig. 4.9c) and 33.0 to 61.5 GJ (Fig. 4.9d), respectively, from the current to 2030s period and decreased by the range of 11.1 to 24.0 GJ (Fig. 4.9c) and 38.6 to 69.6 GJ (Fig. 4.9d), respectively, from the current to 2050s period.

As for the most notable variations in the above-mentioned energy end usages from the current to 2030s period, supply fan stuck at maximum, and minimum speeds led to the largest and smallest increases in cooling coil electric energy of 35.5 and 29.5 GJ (Fig. 4.9b). These two faults also resulted in the least decrease in heating coil gas energy of 8.5 GJ (Fig. 4.9c) and reheating coil electric energy of 33.1 GJ (Fig. 4.9d), respectively. The most reduction in heating coil gas energy of 13.7 GJ (Fig. 4.9c) and in reheating coil electric energy of 61.5 GJ (Fig. 4.9d) were caused by OA damper stuck fully open and no overnight setback, respectively. Moreover, no overnight setback also induced the least increase in fan electric energy by 5.4 GJ and cooling coil SA temperature sensor negative bias triggered off the largest increase in fan electric energy by 9.6 GJ (Fig. 4.9e). With respect to the variations from the current to 2050s period, the most and least variations in heating coil gas energy, reheating coil electric energy and fan electric energy were led by the same faults as those from the current to 2030s period (Fig. 4.9c-4.9e). In addition, the smallest increase in cooling coil electric energy of 56.3 GJ was induced by OA damper stuck fully closed, while the largest increase in cooling coil electric energy of 72.0 GJ was still caused by supply fan stuck at maximum speed (Fig. 4.9b).

Apart from these, most single faults increased by around 2.0 to 7.0 GJ in air system total electric energy from the current to 2030s period (Fig. 4.9a). Regarding the variations from the current to 2050s period, the air system total electric energy increased dramatically by 20.2 to 32.5 GJ compared with the 2030s period. The largest increase in air system total electric energy of 34.3 GJ was caused by heating coil SA temperature sensor negative bias from the current to 2050s period (Fig. 4.9a). Besides, no overnight setback resulted in the most significant decrease of 22.8 GJ in air system total electric energy from current to 2030s period

(Fig. 4.9a) and the least increase of 3.56 GJ from current to 2050s period, mainly due to the most significant reduction in reheating coil electric energy. Besides, delayed evening setback and supply stuck at maximum speed also led to slight decreases of 1.43 GJ and 1.20 GJ in air system total electric energy from current to 2030s period.

In terms of thermal comfort, the PPDs for all single faults decreased by 2.2% to 3.6% from the current to 2030s period and reduced by 2.3% to 4.0% from the current to 2050s period (Fig. 4.9f). The increasing OA temperature in winter can explain these trends as the occupant thermal discomfort mainly originates from the heating season (Fig. 4.9g-4.9h). In detail, the most and least reductions from the two climate period transitions were both caused by supply fan stuck at maximum and minimum speeds, respectively (Fig. 4.9f).

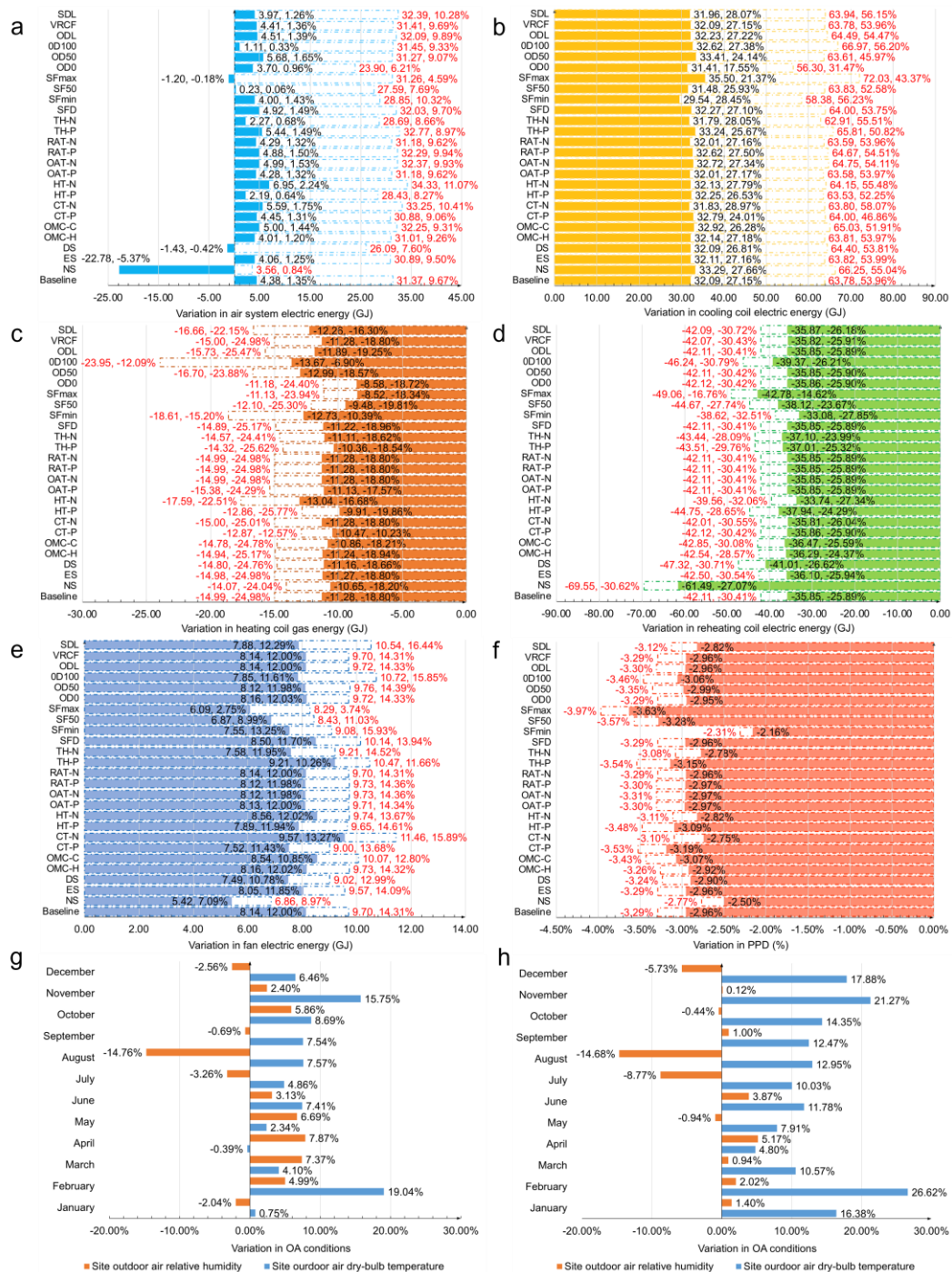


Figure 4.9. (a-f) Variations in air system total electric energy, cooling coil electric energy, heating coil gas energy, reheating coil electric energy, fan electric energy and PPD for single faults from the current to the 2030s period (in solid fill & black labels) and from the current to the 2050s period (in dash-dot line & red labels). (First label for each bar: the variation amount, second label: the variation percentage); Variations in OA dry-bulb temperature and relative humidity between (g) the current and 2030s periods; (h) the current and 2050s periods.

4.2.4.2 Multiple Fault Combinations Impact Variations

The variations in energy end usages and the PPD for each fault combination from the current to 2030s, and from the current to 2050s are illustrated in Figure 4.10. The fault codes presented in each bar chart refer to those in Table 4.2. Around half of the multiple faults combinations, mostly involved with no overnight setback and supply fan stuck at maximum speed, resulted in decreases in air system total electric energy from 1.4 GJ to 47.2 GJ between the current and 2030s periods, and three of them with no overnight setback induced decreases from 1.2 GJ to 16.5 GJ between the current and 2050s periods (Fig. 4.10a). In contrast to the small increases and even significant decreases in air system total electric energy between the current and 2030s periods, the variations in air system total electric energy for most fault combinations increased substantially under the 2050s period (Fig. 4.10a). The most increase in air system total electric energy from the current period was 5.1 GJ by the combination of thermostat positive offset and supply fan stuck at minimum speed under the 2030s period and 35.3 GJ by the combination of thermostat positive offset and OA damper stuck fully open under the 2050s period (Fig. 4.10a). Besides, the largest decrease in air system total electric energy from the current period was 47.2 GJ under the 2030s period and 16.5 GJ under the 2050s period by the combination of no overnight setback and supply fan stuck at maximum speed (Fig. 4.10a).

Additionally, the approximately double increases of 53.4 GJ to 76.5 GJ in cooling coil electric energy between the current and 2050s periods, compared with those of 29.3 GJ to 37.2 GJ between the current and 2030s periods (Fig. 4.10b), are responsible for the significant increases in air system total electric energy under the 2050s period. The most significant increase under both climate period transitions was led by the combination of supply fan stuck at maximum speed and OA damper stuck fully open, while the least increase under both climate period transitions was induced by the combination of supply fan stuck at minimum speed and OA damper stuck fully closed (Fig. 4.10b).

Apart from these, the heating coil gas energy and reheating coil electric energy decreased by 7.5 GJ to 26.8 GJ and 32.8 GJ to 82.2 GJ, respectively, from the current to 2030s period, and reduced by 9.8 GJ to 39.0 GJ and 38.3 GJ to 90.3 GJ, respectively, from current to 2050s period (Fig. 4.10c-4.10d). The most substantial decrease in heating coil gas energy under both climate period transitions was caused by the combination of no overnight setback and OA damper stuck

fully open, and the least decrease was led by the combination of cooling coil SA temperature sensor positive bias and OA damper stuck fully closed (Fig. 4.10c). As for reheating coil electric energy, it was obviously observed that the fault combinations associated with no overnight setback resulted in more significant decreases than the remaining combinations (Fig. 4.10d). As a result, this can explain the phenomena for the combinations with no overnight setback that indicated large decreases in air system total electric energy under the 2030s period and even slight decreases under the 2050s period.

Moreover, the fault combinations, except no overnight setback and supply fan stuck at maximum speed, resulted in increases in fan electric energy by the range of 4.9 to 9.4 GJ from the current to 2030s period, and 6.3 to 12.2 GJ from the current to 2050s period (Fig. 4.10e). The largest increases within both climate period transitions were triggered by the combination of thermostat positive offset and OA damper stuck fully open (Fig. 4.10e). In addition, the fault combinations related to no overnight setback led to less increases than the other faults under both climate period transitions (Fig. 4.10e).

With respect to thermal comfort, the PPDs for fault combinations reduced by the range of 1.9% to 4.3% from the current to 2030s period, and 2.1% to 5.4% from the current to 2050s period (Fig. 4.10f). Remarkably, the combinations associated with no overnight setback and supply fan stuck at minimum speed indicated less decreases than the remaining combinations under both climate period transitions (Fig. 4.10f). Among these combinations, the smallest decreases in the PPD under both climate period transitions was induced by thermostat positive offset and supply fan stuck at minimum speed. On the other hand, the largest decreases in the PPD under both climate period transitions were led by the combination of supply fan stuck at maximum speed and OA damper stuck fully open.

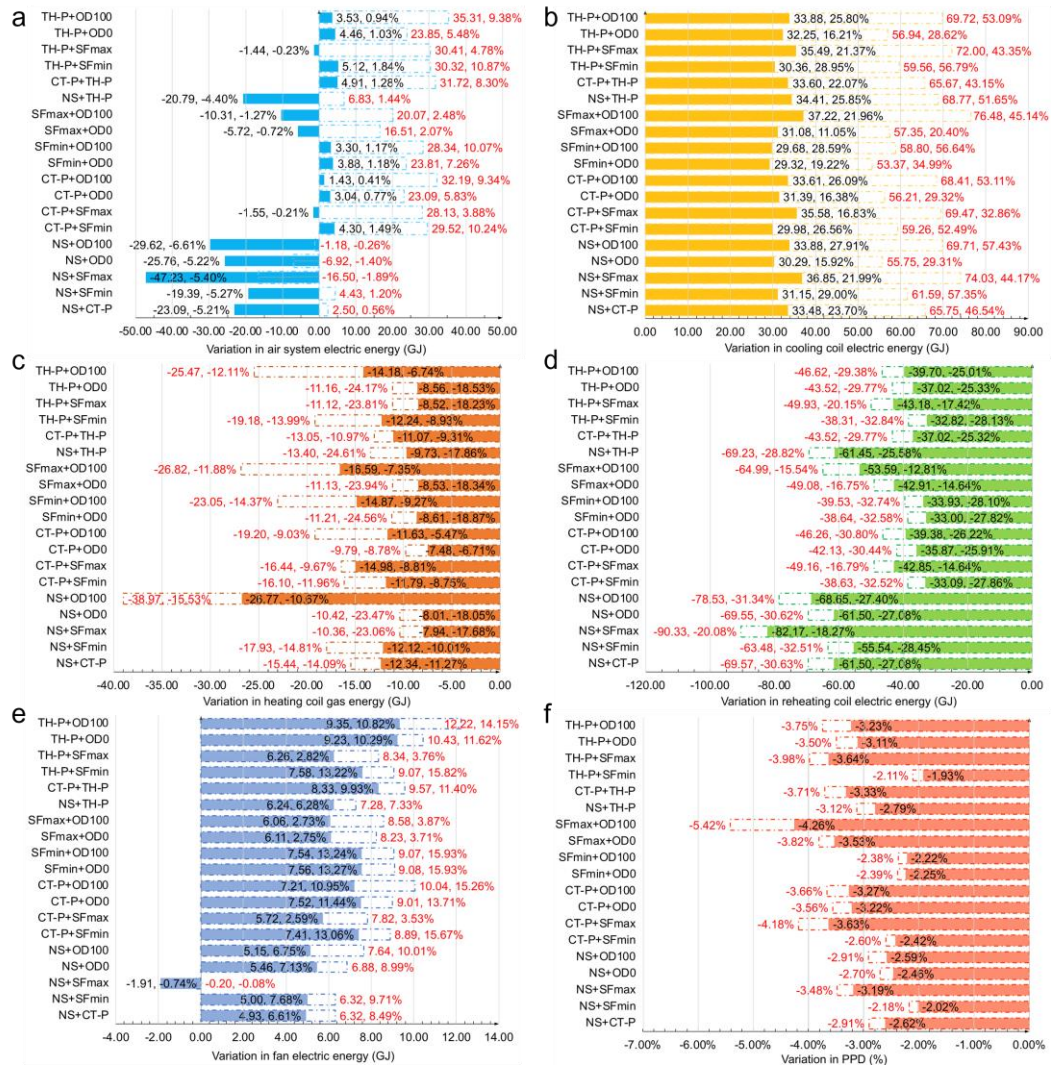


Figure 4.10. (a-f) Variations in air system total electric energy, cooling coil electric energy, heating coil gas energy, reheating coil electric energy, fan electric energy and PPD for multiple faults combinations from the current period to the 2030s period (in solid fill & black labels) and from the current period to the 2050s period (in dash-dot line & red labels). (First label for each bar: the variation amount, second label: the variation percentage).

4.2.5 Comparisons with Findings from Previous Studies

The findings from the results of this study are compared with those from previous studies and discussed in terms of three aspects: (1) single fault impacts; (2) multiple faults interactions; (3) HVAC system performance variations driven by climate change.

From the perspective of single fault impacts, Lu et al. (2021) also implemented a comprehensive evaluation of fault impacts using the Modelica simulations and highlighted that no overnight setback is one of the faults among the top faults in the energy impact

ranking under the current period, which conforms with the finding of this study that this fault is ranked third in terms of energy impact. In addition, OA damper stuck, and thermostat positive bias were found to be within the top faults in the thermal comfort impact ranking, which are also verified in this study.

In terms of multiple faults interactions, Zhou et al. (2022) and Hu et al. (2021b) investigated the interactions among multiple common faults from a different HVAC system to that in this study, VRF system and heat pump system, respectively. Both studies indicated that the impact of multiple faults on the system performances, in terms of cooling capacity and COP, is not an algebraic summation of the impact of each single fault. The impact of each fault can be intensified or cancelled when these faults occur simultaneously, which corresponds to the proposed synergetic and antagonistic effects for multiple faults in the present paper. In this study, the summation rule is also found to be not applicable for the fault impacts of a VAV system and the synergetic and antagonistic effects among multiple faults are verified. However, there is a unique finding proposed in this study, i.e., the synergetic effect of multiple faults can lead to a significantly increased combined fault impact in the case that one fault causes a decrease in an energy end usage and another fault induces an increase in the same energy end usage, and similarly, the antagonistic effect can result in a substantially decreased combined fault impact in the same case.

As for the fault impact variations driven by climate change, due to the lack of studies exploring this, the general trends of fault impact variations across different climate periods are compared with the trends in HVAC system performance variations under normal conditions. Waddicor et al. (2016) indicated that the decrease in system heating energy was higher than the increase in system cooling energy from 2020 to 2050 under the climate of Turin. In addition, Kharseh et al. (2014) pointed out the same phenomenon from the current to 2050s climate period under the cold climate of Stockholm. The results of this study show a similar trend in fault impact variations from the current to 2030s climate period under the marine climate of London, but demonstrated a reverse trend for most single faults and almost half of multiple faults combinations from the current to 2050s climate period.

4.2.6 Key Findings and Potential Future Work

The results of this section are useful for FDD researchers to not only prioritise the FDD for faults with significant impacts, but also pay attention to faults with substantial variations across climate periods, in order to ensure that the FDD framework can adapt to the impact changes in prioritized faults caused by climate change. In addition, according to the variations in fault impact patterns across climate periods, it is worth highlighting that data-driven FDD methods are required to be updated timely using the latest fault measurements from real buildings or fault data simulated by well-calibrated building models. The main findings of this section are summarized as follows:

- (1) Supply fan stuck at maximum speed led to the most significant impacts on both system energy usages and occupant thermal comfort. In addition, the combinations of this fault with OA damper stuck fully open, and this fault with cooling coil SA temperature sensor positive bias were ranked first in the energy and thermal comfort impact ranking, respectively.
- (2) The synergetic effect of multiple faults can lead to a significantly increased combined fault impact in the case that one fault causes a decrease in an energy end usage and another fault induces an increase in the same energy end usage. Similarly, the antagonistic effect can result in a substantially decreased combined fault impact in the same case.
- (3) Changing climate conditions can have a significant influence on the synergetic and antagonistic effects. The increase from the algebraic sum in heating coil gas energy for the combination of cooling coil SA temperature sensor positive bias and supply fan stuck at maximum speed was intensified from 135.17% under the current period to 163.67% under the 2050s period. On the other hand, the largest decrease from the algebraic sum in the same energy usage for the combination of supply fan stuck at minimum speed, and OA damper stuck fully open was also intensified from 166.97% under the current period to 212.55% under the 2050s period.
- (4) Under the marine climate, the decrease in system total heating energy is higher than the increase in system cooling energy from the current to 2030s period for all single and multiple faults. However, the increase in system cooling energy exceeds the decrease in system total heating energy from the current to 2050s period for most single faults and almost half of multiple faults combinations.
- (5) Heating coil SA temperature sensor negative bias, and the combination of thermostat positive offset and OA damper stuck fully open led to the most increase in air system total electric energy by 34.3 GJ and 35.3 GJ, respectively, from the current to 2050s period.

The future work can explore the fault impacts under different climate zones, such as hot climates, to compare the differences in fault impacts under various climate zones. In addition, the fault impacts from other types of HVAC systems, such as fan coil systems, are worth exploring. Besides, the impacts of faults related to the refrigerant cycle of HVAC systems can be further evaluated, and other system performance metrics, such as COP, can be introduced. Moreover, the occurrence frequency of each fault can be taken into account to quantify the fault impacts in a more realistic way if general occurrence frequency data is available. Regarding the integration of fault impact analyses results into the development of data-driven FDD methods, the findings related to fault impact trends across different climate periods and the interactions among multiple faults are worth to be incorporated, in order to enhance the performances of data-driven FDD models and reduce the data amount required for training the models.

4.3 Climate Change-driven Fault Impact Seasonal Pattern Variations

The fault impacts vary significantly with changing climates as HVAC systems operate at different load conditions under different weather conditions, as shown in section 4.2. Thus, it is also important for investigating the seasonal patterns of HVAC fault impacts and exploring the effect of shifting seasons driven by climate change on these seasonal patterns. The main contributions of this section are described as follows: (1) The variations in fault impacts including system energy, thermal comfort levels and energy costs along with seasons were evaluated and seasonal patterns of fault impacts were described. (2) The season onset and end dates for future climate periods (2030s and 2050s) were predicted and the variations in seasonal fault impact patterns among different periods were assessed. Subsection 4.3.1 describes the enhanced fault modeling and simulation workflow by incorporating current and future season onset and end dates into seasonal fault impact analysis. Subsection 4.3.2 explains the method for forecasting future season onset and end dates. Subsection 4.3.3 introduces fault impact indicators used for seasonal fault impact analysis. Subsection 4.3.4 analyses season patterns of fault impacts. Subsection 4.3.5 discusses the effect of shifting seasons driven by climate change on fault impact seasonal patterns. Subsection 4.3.6 shows the potential applications of seasonal fault impact analysis. Subsection 4.3.7 summarises the key findings based on the seasonal fault impact analysis and recommends potential future studies based on the work of this section.

4.3.1 Enhanced Fault Modeling and Simulation Workflow

To identify the seasonal fault impact patterns and investigate the influence of climate-change-driven season shifts on the patterns, the fault modeling and simulation workflow (Fig. 3.1a) in section 3.1 was modified, as shown in Figure 4.11. The addition of current and future season onsets and ends predictions (green block) define the boundary among seasons, which divides the current and future weather data into corresponding seasonal segments. The rest of the parts (orange, red and grey blocks) follow the methods described in Chapter 3.

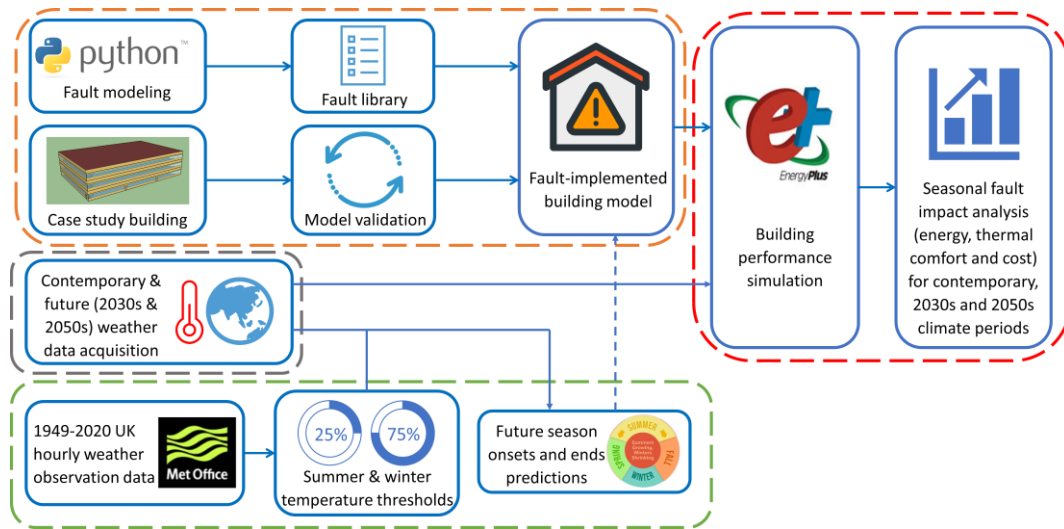


Figure 4.11. The workflow for the proposed method including four main parts: (1) fault implementation to the case study building (orange block), (2) contemporary and future weather data acquisition (grey block), (3) future season onsets and ends predictions (green block) and (4) fault simulations (red block).

4.3.2 Season Onset and End Dates Predictions

The four seasons were divided based on summer and winter temperature thresholds following the works of Christidis et al. (2007) and Park et al. (2018). A third-degree polynomial (Fig. 4.12) was used to fit the raw daily temperature observations of London from 1949 to 2020 (UK Met Office, 2021), in order to ensure that there were only two intersections between the daily temperature observation and each temperature threshold. The summer and winter temperature thresholds were defined as the 75th and 25th percentiles of daily average temperature over the period of 1949–2020, respectively (Wang et al., 2021) and were calculated as 16.3 °C and 7.3 °C, respectively. The intersections between the polynomial fit and each threshold are the onset and end dates of summer and winter, respectively.

Consequently, spring was defined as the transition period from winter to summer with rising

daily air temperatures and autumn was the transition period from summer to winter with reducing daily air temperatures. The onset and end dates as well as the length of each season for the current, 2030s and 2050s periods were therefore evaluated and presented in Table 4.3. It should be noted that this section didn't aim to accurately predict the season onset and end dates, but rather to explore the seasonal fault impacts and to evaluate the variations in seasonal fault impact patterns under the case of shifting seasons driven by climate change.

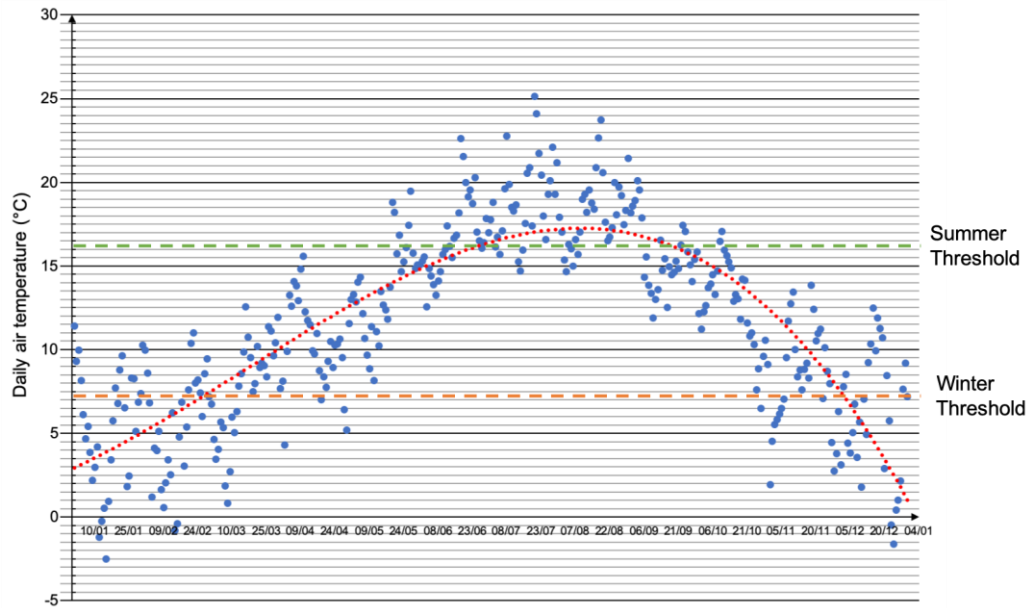


Figure 4.12 An example of the third-degree polynomial fit (red dot line) of the raw daily temperature observations and intersections with summer (green dash line) and winter (orange dash line) thresholds.

Table 4.3. Predicted season onset and end dates and season lengths (date format: day/month).

Season	Spring			Summer			Autumn			Winter		
	Onset	End	Length	Onset	End	Length	Onset	End	Length	Onset	End	Length
Current	14/3	19/5	68	20/5	8/9	111	9/9	27/11	81	28/11	13/3	105
2030s	8/3	31/5	85	1/6	12/10	133	13/10	23/12	73	24/12	7/3	74
2050s	16/2	5/5	79	6/5	6/10	153	7/10	22/12	78	23/12	15/2	55

4.3.3 Seasonal Fault Impact Indicators

The faults and their intensities included in this section are same as those provided in Table 3.3 in subsection 3.5.1. The system total electric and gas energy as well as PPD were extracted from simulation results and used for comparing the impacts of each fault under seasons of the current, 2030s and 2050s periods. The system electricity and gas costs were calculated as the

seasonal fault impact indicators, based on the 2020 UK annual electricity and natural gas price of 0.154 £/kWh and 0.049 £/kWh are used (UK Department for Business, Energy and Industrial Strategy, 2021).

4.3.4 Fault Impact Seasonal Patterns Analysis

The variations from the baseline case in system total electric and gas energy usages as well as PPD for 26 fault cases under seasons of three climate periods are illustrated in Figure 4.13. Within each climate period, some faults presented remarkable seasonal patterns in terms of fault impacts. Supply fan stuck at maximum speed (SFmax) indicated substantial increases in system electric energy usage for all the seasons and the increase was most significant in current winter, 2030s summer and 2050s summer. In addition, this fault led to most decreases in system gas energy usage under each winter and considerable increases in PPD under each summer. The patterns similar to SFmax for the variations in system gas energy usage were also presented by supply fan stuck at 50% speed (SF50) and OA damper stuck fully closed (OD0).

Apart from these, occupant manual changes of heating setpoints (OMC-H) showed remarkable increases in system electric energy usage in spring and autumn under each period, and current winter. Similar patterns were exhibited by no overnight setback (NS), but this fault induced most increases in each winter. However, these two faults presented completely different seasonal patterns in terms of PPD. OMC-H indicated obvious improvements in thermal comfort levels in each spring, each autumn under 2030s and 2050s periods, and current winter. NS presented a different pattern that the increase in PPD was exhibited in each summer while the decrease in PPD was shown in each winter.

In contrast to SFmax, supply fan stuck at minimum speed (SFmin) indicated remarkable decreases in system electric energy usage for all the seasons and the largest decrease occurred in each summer. Besides, substantial increases in system gas energy usage were observed in spring, autumn and winter under each period and the most significant decrease in PPD was presented in each summer. Similar to SFmin, the increasing patterns for system gas energy usage in spring, autumn and winter under each period were also observed from cooling coil SA temperature sensor positive bias (CT-P) and OA damper stuck fully open (OD100). Moreover, the increasing patterns of OD100 were most profound compared with the increasing patterns of other faults.

4.3.5 Influence of Shifting Seasons on Fault Impact Seasonal Patterns

With the changes in season lengths and OA temperature amplitudes of seasons, the impacts of some faults on system electric energy usage were phenomenally intensified, such as SF max, and the impacts of some faults on system gas energy usage were mitigated, such as OD100. In addition, most faults with seasonal patterns under the current period still kept similar patterns in future climate periods. Besides, PPD was the impact criterion that was not heavily affected by the shifting seasons. The only two notable fault impact pattern variations occurred at SFmax and SFmin between 2030s autumn and current autumn, and between 2050s spring and 2030s spring.

As for cost distributions of all the faults (Fig. 4.14), the seasonal pattern of gas cost distribution remains similar among each climate period. The only remarkable difference among three periods was that the gas cost distribution decreased with the shortening of winter length from the current to 2050s period. However, the electricity cost distribution pattern under each period was significantly different from each other. Under the current period, the electricity cost distributions for summer and winter were in a similar range between £4500 to £6000, and also considerably higher than those for spring and autumn. Under the 2030s period, the distribution for summer increased by around £1000 to £1500. The distribution for winter decreased by around £2000, but was still higher than that for the other seasons. As for the 2050s period, the distribution for summer further increased by around £2000, while the distribution for winter decreased to a similar range between £2000 to £3000 to that for autumn.

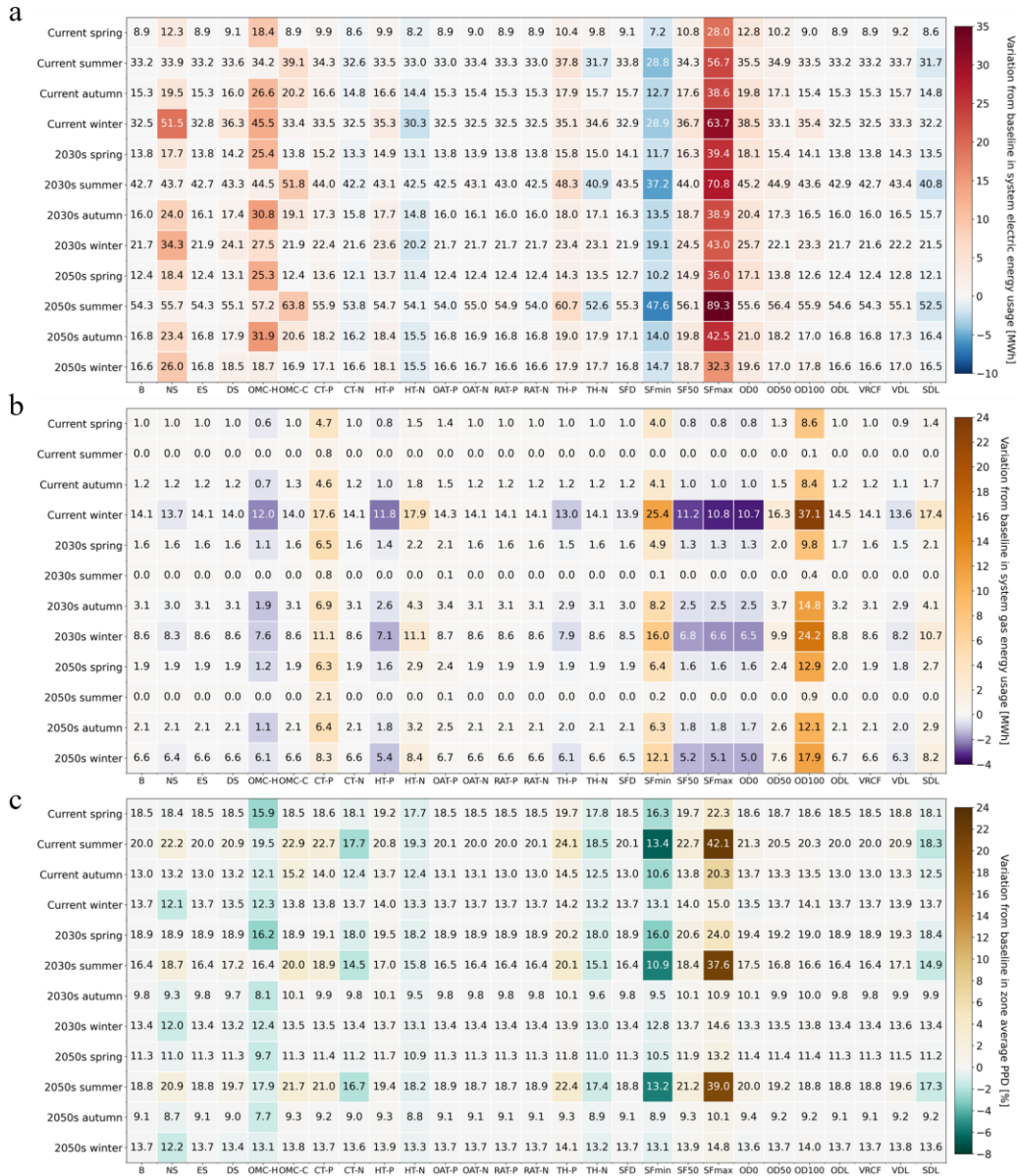


Figure 4.13 (a) System total electric energy usage; (b) System total gas energy usage; (c) Zone average PPD for all the faults under seasons of current, 2030s and 2050s climate periods. (the label on each grid represents the exact amount of the parameter rather than the variation).

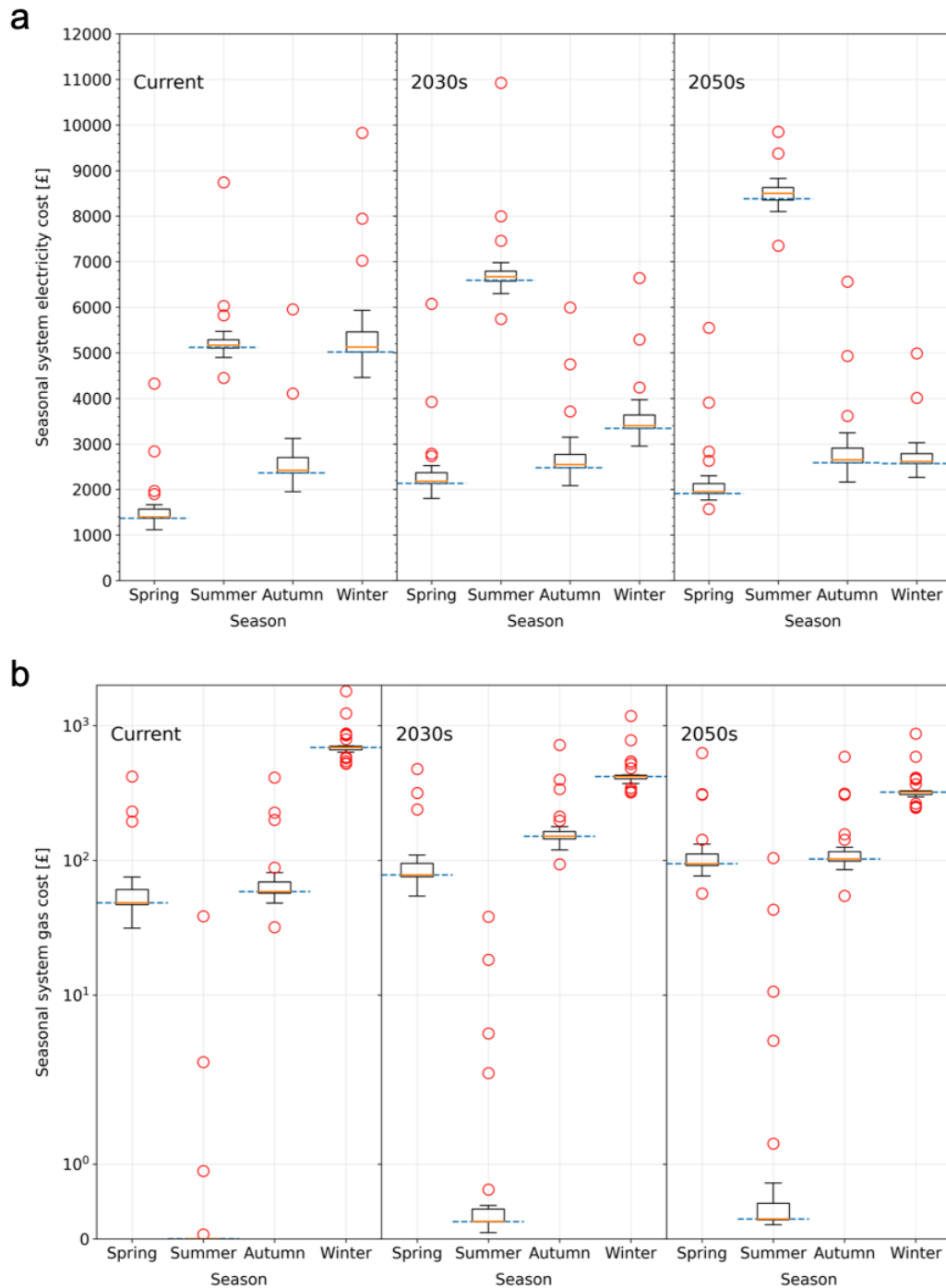


Figure 4.14 (a) System electricity and (b) system gas cost distributions for all the faults under seasons of current, 2030s and 2050s climate periods (blue dash lines indicate the baseline cost).

4.3.6 Potential Applications of Seasonal Fault Impact Analysis

The seasonal fault impact analysis in this section proves the fault behaviours vary with different seasons, and climate change-driven shifting seasons have phenomenal influences on some faults, such as SFmax and OD100. These findings provide a potential enhancement for

identifying the focus on faults with substantial seasonal impacts for developing and evaluating FDD methods. In addition, the investigations of future fault impact seasonal patterns can provide a proactive perspective for fault behaviours with climate change, in order to either update detection and diagnosis rules for rule-based FDD or recalibrate system physics-based models for model-based FDD, or improve machine learning model architectures for data-driven FDD.

4.3.7 Key Findings and Potential Future Work

In this section, the seasonal patterns of fault impacts were assessed under the contemporary, 2030s and 2050s climate periods, and the effect of climate change-driven shifting seasons on these seasonal patterns was explored. The main findings are summarised as follows:

- (1) Fault impact seasonal patterns were exhibited by a large number of faults. SFmax and SFmin presented most profound patterns which were also opposite to each other.
- (2) Most faults kept seasonal patterns along with changing climates. The impacts of some faults were phenomenally intensified on system electric energy usage and the impacts of some faults were mitigated on system gas energy usage. PPD was the impact factor that was not heavily affected by shifting seasons.
- (3) The gas cost distribution exhibited similar seasonal patterns in different climate periods while the electricity cost distribution presented changing seasonal patterns.

The future work will explore fault impact seasonal patterns under other climate zones, such as hot climate, to summarise the patterns from various climate zones. Besides, multiple faults cases will be considered to identify fault impact seasonal patterns for different fault combinations.

4.4 Fault Impact Analysis Summary

There are two main contributions achieved in this chapter: (1) the impacts of single and multiple faults and the interactions among multiple faults under different climate periods were explored; (2) the fault impact seasonal patterns were assessed under different climate periods and the effect of climate change-driven shifting seasons on these seasonal patterns were investigated. These contributions led to the key findings as below:

- (1) Under the marine climate, the decrease in system total heating energy is higher than the increase in system cooling energy from the current to 2030s period for all single and multiple faults. However, the increase in system cooling energy exceeds the decrease in system total

heating energy from the current to 2050s period for most single faults and almost half of multiple faults combinations.

(2) The synergetic effect of multiple faults can lead to a significantly increased combined fault impact in the case that one fault causes a decrease in an energy end usage and another fault induces an increase in the same energy end usage. Similarly, the antagonistic effect can result in a substantially decreased combined fault impact in the same case.

(3) Changing climate conditions can have a significant influence on single fault impacts and multiple faults' interactions including synergetic and antagonistic effects.

(4) Fault impact seasonal patterns were exhibited by a large number of faults and most faults kept seasonal patterns along with changing climates.

The future work can explore the fault impacts under different climate zones, such as hot climates, to compare the differences in fault impacts under various climate zones. In addition, the fault impacts from other types of HVAC systems, such as fan coil systems, are worth exploring. Moreover, the occurrence frequency of each fault can be taken into account to quantify the fault impacts in a more realistic way if general occurrence frequency data is available. Regarding the integration of fault impact analyses results into the development of data-driven FDD methods, the findings related to fault impact trends across different climate periods and the interactions among multiple faults are worth to be incorporated, in order to enhance the performances of data-driven FDD models and reduce the data amount required for training the models.

Chapter 5 Data-driven Fault Detection and Diagnosis Model Development

5.1 General Workflow

Data-driven FDD methods can automatically capture the underlying behaviours of faults through relevant features from HVAC system operational data, but is usually sensitive to the data quality and availability. Therefore, some essential procedures need to be implemented during the development of data-driven FDD methods. This chapter describes the processes for developing a data-driven FDD model of the present study. The model development workflow is illustrated in Figure 5.1. Section 5.2 describes the two main sources of fault datasets: (1) fault simulations based on a verified prototype medium office building model, as mentioned in section 3.5; (2) public research fault datasets (Wen and Li, 2011; Kim and Leach, 2019). Section 5.3 shows data pre-processing techniques implemented to improve the quality of input data for training the FDD model, including preliminary feature selection from raw HVAC time series data, outlier detection and data normalization. Section 5.4 presents the way to transform the 1D time series data into two-dimensional (2D) data using the sliding window method. Section 5.5 explains the method that splits the input dataset into training, validation and testing sets. Section 5.6 describes the proposed FDD architecture structures and the theory of deep learning algorithms used in the architecture. This section also lists the candidates of hyperparameters and the hyperparameters tuning method. Section 5.7 provides model evaluation metrics for assessing FDD performances under the training, validation and testing stages. It should be noted that most contents of this chapter have been published in the paper ‘Fault data seasonal imbalance and insufficiency impacts on data-driven heating, ventilation and air-conditioning fault detection and diagnosis performances for energy-efficient building operations’ (Zhong, Calautit and Wu, 2023) and hence those parts are reused in this chapter.

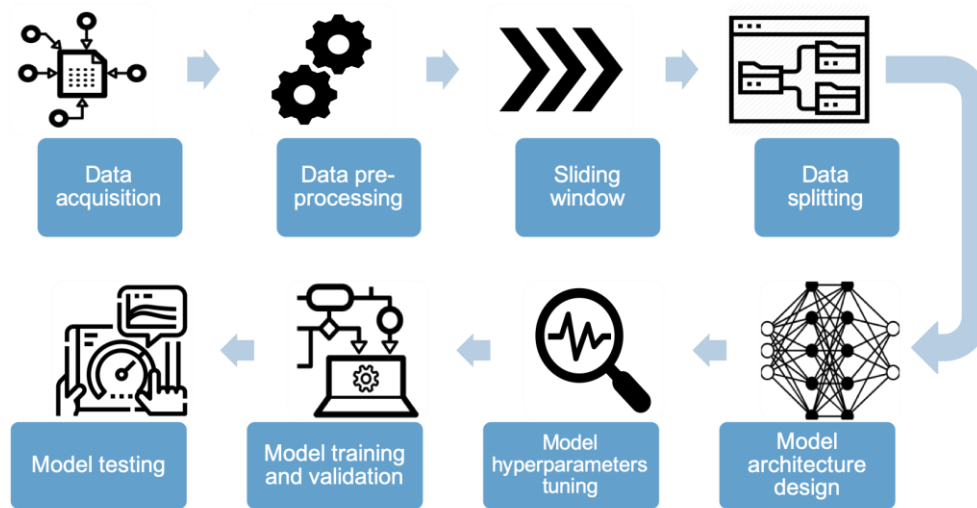


Figure 5.1 Data-driven fault detection and diagnosis model development workflow.

5.2 Fault Data Acquisition

The collected fault data in this thesis originated from three specific sources: (1) fault simulations based on a verified prototype medium office building model; (2) experimental fault measurements from ASHRAE research project 1312 (RP-1312) (Wen and Li, 2011); (3) experimental fault measurements from National Renewable Energy Laboratory (NREL)’s fault tests (Kim and Leach, 2019).

5.2.1 Description of Simulated Fault Dataset

The simulated fault dataset was generated using EnergyPlus fault simulations based on the prototype medium office building model, as mentioned in section 3.3, under the marine climate of London. 14 faults including control, sensor, packaged air conditioner and VAV terminal faults, as listed in Table 5.1, were simulated at a 5-minute interval for 48 hours under each season. The normal operating conditions of the VAV system were simulated at the same interval for adjacent 272 hours in total before and after fault occurrences. Thus, each fault case accompanied by the adjacent normal system operation produced 3840 data samples in total for 320 hours per season, and each one-season dataset has 53760 samples for 4480 hours. Overall, the ratio of the total amount of fault data to normal data is 0.18. The fault intensities, class labels and sample sizes for fault and normal classes are presented in Table 5.1.

Table 5.1 Simulated HVAC fault and normal data descriptions.

Class label	Fault type	Fault intensity	Sample size
0	Normal		45696 per season
1	No overnight setback	Setpoint setback is off	576 per season
2	Cooling coil SA temperature sensor positive bias	+2 °C	576 per season
3	Cooling coil SA temperature sensor negative bias	−2 °C	576 per season
4	Heating coil SA temperature positive sensor bias	+2 °C	576 per season
5	Heating coil SA temperature negative sensor bias	−2 °C	576 per season
6	OA temperature sensor positive bias	+2 °C	576 per season
7	OA temperature sensor negative bias	−2 °C	576 per season
8	Thermostat positive offset	+2 °C	576 per season
9	Thermostat negative offset	−2 °C	576 per season
10	Supply fan motor efficiency degradation	15%	576 per season
11	Supply fan stuck	Stuck at the 50% speed	576 per season
12	OA damper stuck	Stuck at the 25% position	576 per season
13	OA damper leakage	Leaking percentage 40%	576 per season
14	VAV damper leakage	Leaking percentage 40%	576 per season

5.2.2 Description of Real Fault Datasets

The two real fault datasets were gathered from ASHRAE RP-1312 (Wen and Li, 2011) and NREL’s fault tests (Kim and Leach, 2019). These datasets were used to verify the findings from the fault data evaluation experiments based on the simulated dataset. The ASHRAE RP-1312 carried out experimental measurements for an AHU in the Iowa Energy Resource Station under both fault and normal operating conditions. The Energy Resource Station is comprised of AHU–1, responsible for the common areas of the test facility; AHU–A and AHU–B, which are identical, and each serves four mirror zones, as shown in Figure 5.2a. Thus, a set of comparison experiments were carried out in these two AHUs. The fault tests were implemented using the AHU-A, and the comparative normal operating conditions were

monitored using the AHU-B. The measurements for 6 faults, implemented under multiple seasons, i.e., spring, summer and winter, were extracted from the original raw dataset. The faults were record at a one-minute interval from 6:00 to 18:00 for both AHUs. The fault types and their intensities, and sample sizes for fault and normal classes are presented in Table 5.2. The total amounts of data samples under spring, summer and winter are 17280, 17280 and 11520, respectively. Besides, the ratio of the total amount of fault data to normal data is 0.41.

Regarding the NREL's fault dataset, the fault tests were performed in a flexible research platform (FRP) (Fig. 5.2b). The FRP is equipped with a direct-expansion RTU and a reheat VAV system. 32 fault tests were conducted in the FRP, and five of them, including HVAC setpoint setback faults and thermostat biases were taken into account in the thesis. The reason for selecting these five faults is that only these faults were tested under multiple seasons, i.e., summer and winter. The fault data was recorded at a 15-minute interval, and each fault test lasted for 24 hours. The fault intensities and sample sizes for fault and normal classes are also shown in Table 5.2. Due to the limited amount of samples under each fault class, the fault amount under each season was extracted equally from the original raw dataset, in order to avoid the original large imbalance ratio between fault data under different seasons. The total amount of data samples is 1056 for each season, and the ratio of the total amount of fault data to normal data is 0.17.

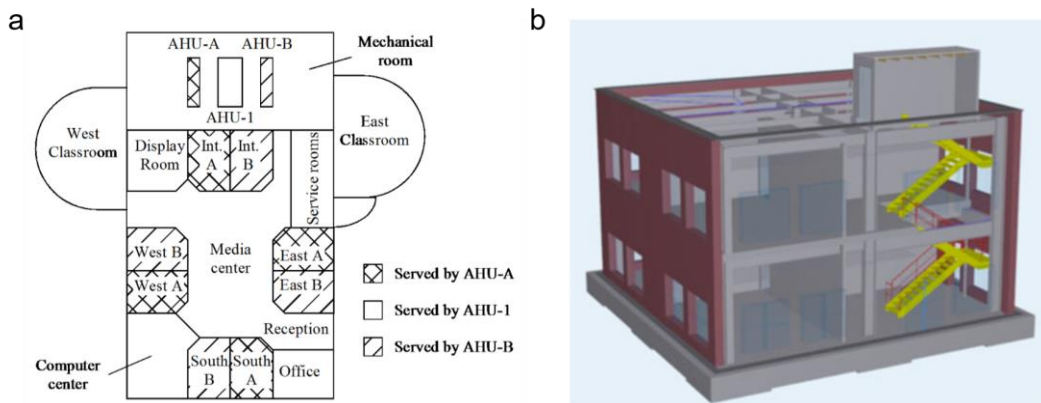


Figure 5.2 (a) The layout of Iowa Energy Resource Station; (b) a 3D model of the FRP (Buckberry and Bhandari, 2012).

Table 5.2. Real HVAC fault and normal data descriptions.

ASHRAE RP-1312 Dataset			
<i>Class label</i>	<i>Fault type</i>	<i>Fault intensity</i>	<i>Sample size (spring/summer/winter)</i>
0	Normal		11883/12664/8076
1	OA damper stuck fully closed	Stuck at the 0% position	722/722/722
2	EA damper stuck fully open	Stuck at the 100% position	1440/1126/1319
3	EA damper stuck fully closed	Stuck at the 0% position	722/721/721
4	Cooling coil valve stuck fully closed	Stuck at the 0% position	715/714/N/A
5	Cooling coil valve stuck fully open	Stuck at the 100% position	1077/612/682
6	Return fan complete failure	Off	721/721/N/A
NREL Dataset			
<i>Class label</i>	<i>Fault type</i>	<i>Fault intensity</i>	<i>Sample size (summer/winter)</i>
0	Normal		576/576
1	HVAC setback delayed onset	3 hours	96/96
2	HVAC setback early termination	3 hours	96/96
3	HVAC no setpoint setback	Off	96/96
4	Thermostat positive bias	2.2 °C	96/96
5	Thermostat negative bias	2.2 °C	96/96

5.3 Data Pre-processing Techniques

Data pre-processing is an essential process for achieving desirable data-driven FDD model training performance due to the intrinsic complexity of HVAC system operations and the data quality issues. This section introduces three data pre-processing techniques carried out to enhance the model input data quality. The preliminary feature selection, based on HVAC domain knowledge and the summary of HVAC system selected features from literature review in Chapter 2, from raw HVAC time series data are presented in subsection 5.3.1. The outlier detection technique used in this work is explained in subsection 5.3.2. The data normalization method for scaling the input features into similar scales is described in subsection 5.3.3.

5.3.1 Preliminary Feature Selection from Raw Data

The input features for training the data-driven FDD model were manually selected from each dataset referring to HVAC domain knowledge and the summary of selected features from literature. The feature selection follows the principles of filtering out the redundant features that measure similar parameters to other features and keeping the features relevant to the mechanism of each fault. According to the simulated fault dataset, four main categories of features were extracted, including weather features, building thermal zone features, packaged air conditioner features and energy-based features, as presented in Table 5.3. Regarding the ASHRAE RP-1312 dataset, the selected features could be grouped into coil features, damper features, fan features and air features, as listed in Table 5.4. As for the NREL's dataset, five categories, containing zone features, duct features, weather features, RTU features and energy-based features, were extracted, as shown in Table 5.5.

Table 5.3 Descriptions of selected features from the simulated fault dataset.

Category	Features	Units
<i>Weather Features</i>	OA dry-bulb temperature	°C
	OA wet-bulb temperature	°C
	OA relative humidity (RH)	%
	OA barometric pressure	Pa
	Wind speed	m/s
	Wind direction	°
	Rain Status	On/off
<i>Zone Features</i>	Zone air temperature	°C
	Zone RH	%
	VAV damper position	%
	VAV outlet temperature	°C
	VAV outlet airflow rate	m/s
<i>Packaged Air Conditioner Features</i>	Supply fan flow rate	m/s
	Supply fan differential pressure	Pa
	OA flow rate	m/s
	RA temperature	°C
	MA temperature	°C
	Cooling coil outlet temperature	°C
	Heating coil outlet temperature	°C
	SA temperature	°C
<i>Energy Features</i>	HVAC total electricity consumption	J
	HVAC gas consumption	J
	Heating electricity consumption	J
	Cooling electricity consumption	J
	Fan electricity consumption	J

Table 5.4 Descriptions of selected features from the ASHRAE RP-1312 dataset.

Category	Features	Units
<i>Coil Features</i>	Heating coil valve position	%
	Cooling coil valve position	%
	Heating coil discharge air temperature	°F
	Cooling coil discharge air temperature	°F
<i>Damper Features</i>	EA damper position	%
	RA damper position	%
	OA damper position	%
<i>Fan Features</i>	Supply fan power	W
	Return fan power	W
	Supply fan differential pressure	Inch of water
	Return fan differential pressure	Inch of water
	Supply fan speed	%
	Return fan speed	%
<i>Air Features</i>	SA flow rate	Cubic feet per minute (CFM)
	RA flow rate	CFM
	OA flow rate	CFM
	SA temperature	°F
	MA temperature	°F
	RA temperature	°F
	SA duct static pressure	Inch of water

Table 5.5 Descriptions of selected features from the NREL’s dataset.

Category	Features	Units
<i>Zone Features</i>	Room staircase temperature	°F
	Room temperature	°F
	Room staircase RH	%
	Room RH	%
<i>Duct Features</i>	Supply duct temperature	°F
	Return duct temperature	°F
	Exhaust duct temperature	°F
	Supply duct RH	%
	Return duct RH	%
	Exhaust duct RH	%
<i>Weather Features</i>	Barometric pressure	Pound per square inch
	OA temperature	°F
	OA RH	%
<i>RTU Features</i>	RTU unit fresh air intake temperature	°F
	RTU SA temperature	°F
	RTU RA temperature	°F
	RTU unit fresh air intake RH	%
	RTU SA RH	%
	RTU RA RH	%
	RTU actual volumetric air flow rate	CFM
	RTU standard volumetric air flow rate	CFM
	Total cooling capacity across the evaporator	W
	Gross coefficient of performance of RTU	
<i>Energy Features</i>	Exhaust fan electricity consumption	Wh
	RTU controller electricity consumption	Wh
	RTU total electricity consumption	Wh
	RTU compressor electricity consumption	Wh
	RTU condenser fan electricity consumption	Wh
	RTU evaporator fan electricity consumption	Wh
	VAV box reheat energy consumption	Wh
	Natural gas standard volumetric flow rate	Wh
	Total electric energy consumption	Wh
	Total heating electricity consumption	Wh
	Total cooling electricity consumption	Wh

5.3.2 Outlier Detection

Outliers are samples which are exceptionally distinct from the majority of the data samples (Kuhn and Johnson, 2018). There are two main groups of methods for outlier detection:

statistical methods and clustering methods. Generalised extreme studentised deviate (GESD) (Seem et al., 2007) method is a statistical method that determines the outlier boundary and amount based on typical statistics including standard deviations and means. Another useful technique is order-based methods, such as interquartile range (IQR) method (Yu et al., 2012), which detect outliers based on data percentage quantiles. Clustering methods are able to detect outliers in two approaches: (1) the statistical methods are used for outlier detection based on the data clusters identified by the clustering methods; (2) the clustering methods are applied directly for outlier detection.

In this thesis, the IQR method was used to detect the outliers within HVAC time series data. The IQR was given by the difference between the 75th and the 25th percentiles (Q_{75th} and Q_{25th}) of the data and can be used to detect outliers by setting upper and lower thresholds. The data sample values falling within the range between $Q_{75th} - 1.5(Q_{75th} - Q_{25th})$ and $Q_{25th} + 1.5(Q_{75th} - Q_{25th})$ were detected as outliers and the corresponding data samples were removed.

5.3.3 Data Normalization

Data normalization is a crucial procedure to ensure that the scales of feature values are similar. This procedure can minimize the effect of less weights of small-value features compared with large-value features on the trained FDD model. There are two popular options for normalizing HVAC system operational data, i.e., the max-min normalization method and z-score standardization method.

The max-min normalization method (Fig. 5.3) was employed to scale the features of each dataset. The maximum and minimum values of each feature were used to convert each data value into the range between zero and one:

$$X'_i = \frac{X_i - \min(X_i)}{\max(X_i) - \min(X_i)} \quad (5.1)$$

where X_i is the original data sequence of each feature, X'_i is the normalized data sequence of each feature, $\max(X_i)$ and $\min(X_i)$ are the maximum and minimum values of the data sequence of each feature.

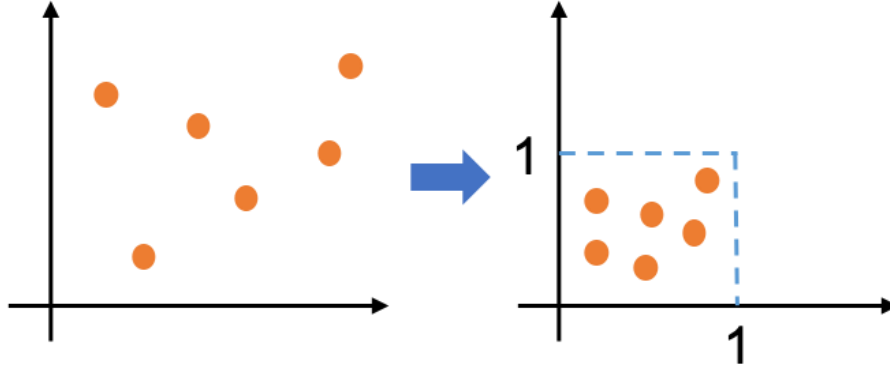


Figure 5.3 The max-min normalization method.

5.4 Sliding Window

After the data pre-processing, the 1D HVAC time series data was transformed into 2D data using a sliding window (Fig. 5.4) as a single data sample was not informative enough to indicate a fault occurrence due to the dynamic operation of HVAC systems. The number of windows W_n was calculated using the following equation:

$$W_n = \frac{S_n - W_l}{W_s} + 1 \quad (5.2)$$

where S_n is the number of data samples within the data sequence, W_l is the window length, W_s is the window sliding stride. The candidates of the window length considered in a parametric study were 6, 12 and 24 and also used for the window stride correspondingly, which means that there is no overlapping between adjacent windows. The resulting dataset sizes under each season for the simulated dataset are presented in Table 5.6.

Table 5.6 Sliding window specifications for the simulated dataset.

Window Length	Window stride	Number of windows in each one-season dataset
6	6	8960
12	12	4480
24	24	2240

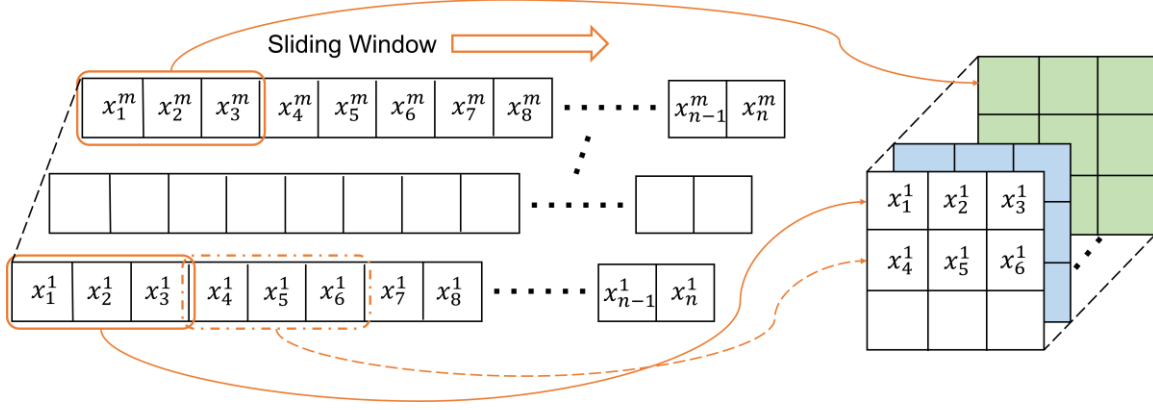


Figure 5.4 The 1D-sliding window method for time-series data with a dimension of $n \times m$.

5.5 Dataset Splitting

Overfitting is a common phenomenon in the application of machine learning to real-world problems. It occurs when the model learns its parameters from a dataset and is tested on the same dataset. As a result, this model would have perfect performance scores on the seen data, but would perform poorly on the unseen data. To prevent overfitting the FDD model, the transformed HVAC time series data from the input dataset was split into three subsets, namely, training, validation and testing sets for the purpose of cross-validation. The training set was used to optimize its model parameters during the FDD model training. The validation set was employed to evaluate model performances during the training to indicate whether the training needs to continue or not. The testing set was utilized to assess the final model performances after the training. The training data size ratio was selected from the candidates of 80%, 70% and 60% and will be determined based on a general hyperparameter analysis in section 6.2. Subsequently, the validation and testing data size were defined as 50% of the remaining dataset, respectively. The stratified shuffle split method (Pedregosa et al., 2011) was used as the cross-validation strategy, which splits the dataset by preserving the original proportion of data samples from each class in the split sets. The split was repeated ten times, and thus the architecture was trained, validated and tested on ten different data distributions.

5.6 Model Architecture and Hyperparameters

The proposed model architecture combines both convolutional and recurrent layers, as illustrated in Figure 5.5. Compared with the conventional shallow machine learning algorithms, such as SVM and fully-connected neural networks, the combined CNN-RNN architecture has the following advantages: (1) the convolutional part of the architecture can extract discriminative features from HVAC time series data, which contains a large number of

sensor signal sequences and complex dynamic fault patterns (Cheng et al., 2021), particularly in the present study; (2) the recurrent part can learn the temporal patterns in HVAC operational datasets; (3) the CNN-RNN architecture results in an end-to-end learning, which mitigates the potential performance degradation due to the separation of feature extraction and classification (Canizo et al., 2019). The preliminary theory explanations for CNN and RNNs are provided in section 5.6.1. The structure of the proposed architecture and its hyperparameters are described in section 5.6.2.

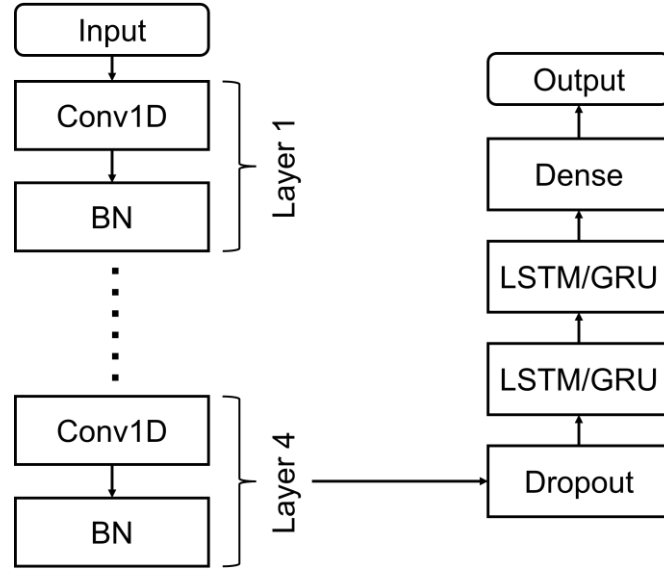


Figure 5.5 The proposed combined CNN-RNN architecture (the recurrent layers were selected from LSTM and GRU).

5.6.1 Preliminary Theory for the Architecture

In this subsection, the theory behind the convolutional neural network and recurrent neural network will be explained in subsection 5.6.1.1 and 5.6.1.2, respectively.

5.6.1.1 Convolutional Neural Network

Convolution layers are able to capture insightful features from input data through the convolutional kernel. As a result, the output feature map is generated by sliding the kernel over the input data or the feature map generated by the preceding convolution layer, as illustrated in Figure 5.6. The output of a 1D convolution can be usually denoted as (Cheng et al., 2021):

$$(I * K)_i = \sum_{j=0}^{N-1} I_j K_{i-j} \quad (5.3)$$

where I represents the input time series sequence of length N and K is the convolutional kernel.

There are several key hyperparameters regarding the convolutional layer. The number of kernels represents the number of output feature maps from the current convolutional layer. The kernel size defines the scale of the convolution operation, i.e., the length of the 1D convolutional window in this case (green windows in Figure 5.6). The stride is the step that the kernel slides over the input feature map. The padding parameter can be either ‘valid’ or ‘same’. ‘Valid’ means that no padding is applied to input feature map(s). ‘Same’ leads to adding zeros around the edges of each input feature map so that the dimensionality of each output feature map can be kept the same as that of the corresponding input feature map. Apart from these, the activation function, which is used to transform the weighted sum of the input into the output of each node from each layer, for the convolutional layer is usually rectified linear activation function (ReLU), as expressed in Eqn. 5.4.

$$f(x) = \max(0, x) \quad (5.4)$$

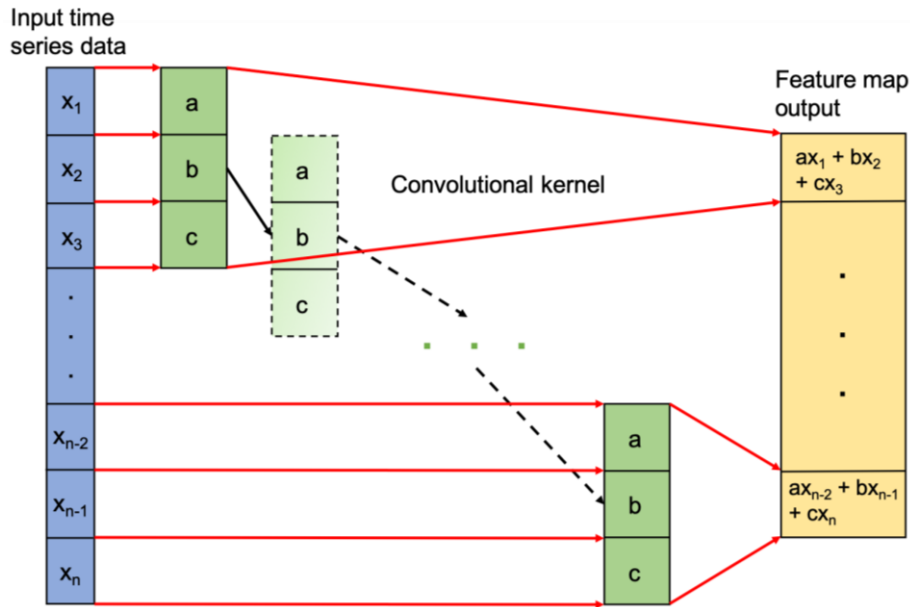


Figure 5.6 An example of 1D convolution operation.

5.6.1.2 Recurrent Neural Network

Recurrent layers are capable of storing useful information from previous time steps and employing them for future predictions. As presented in Figure 5.7a, a recurrent layer consists of a single memory cell that repeatedly computes the output Y_t for each time step. At each time step, the memory cell takes the input time series (starting from X_0 to X_t) at that time step and generates an output (starting from Y_0 to Y_t) for that time step. Besides, a hidden state vector (starting from H_0 to H_t) is generated at each time step and fed as an extra input to the memory cell at the next time step. However, vanilla RNNs usually suffer from vanishing gradient issue, and therefore two variants of the vanilla RNN, GRU and LSTM, have been adopted as solutions to this issue (Figure 5.7b).

The LSTM was proposed by Hochreiter and Schmidhuber (1997) to resolve the vanishing gradient issue. Instead of using an activation function of the hidden state in the memory cells, three gates were employed in the memory cells: the input, forget and output gates. Each gate uses a sigmoid activation function to control the values passing through the gate. The forget gate f_t determines the degree of the information from the previous state to be forgotten. The output gate o_t controls which part of the cell is output to the hidden state h_t . The mathematical operations for the input i_t , forget f_t and output gates o_t , the candidate cell $\tilde{c}_t$, the updated cell state c_t and the output of the current hidden unit h_t are expressed as follows:

$$i_t = \sigma(W_i[x_t + h_{t-1}] + b_i) \quad (5.5)$$

$$f_t = \sigma(W_f[x_t + h_{t-1}] + b_f) \quad (5.6)$$

$$o_t = \sigma(W_o[x_t + h_{t-1}] + b_o) \quad (5.7)$$

$$\tilde{c}_t = \tanh(W_c[x_t + h_{t-1}] + b_c) \quad (5.8)$$

$$c_t = c_{t-1} \times f_t + \tilde{c}_t \times i_t \quad (5.9)$$

$$h_t = o_t \times \tanh(c_t) \quad (5.10)$$

where W_i , W_f , W_o and W_c are the weighting vectors, b_i , b_f , b_o , b_c are the bias vectors, the subscripts i , f , o and c denote the input gate, forget gate, output gate and the memory cell, respectively. σ denotes the sigmoid function and $\tanh$ is the hyperbolic tangent function.

GRU was proposed by Cho, van Merriënboer and Bahdanau (2014). Compared with LSTM, the GRU has only two gates and less parameters, and thus the training of GRU is more

efficient and requires less computational power. The GRU consists of two gates, i.e., the update gate and the reset gate. Similar to the gates used in LSTM, the sigmoid activation function is used for each gate. The update gate determines the degree of the cell state to be updated with the candidate state. The reset gate determines whether the previous hidden state should be kept or not. The mathematical operations for the update gate, reset gate, the candidate cell and the output of the current hidden unit are represented as follows:

$$z_t = \sigma(W_z[x_t + h_{t-1}] + b_z) \quad (5.11)$$

$$r_t = \sigma(W_r[x_t + h_{t-1}] + b_r) \quad (5.12)$$

$$\tilde{h}_t = \tanh(W_h[x_t + r_t \times h_{t-1}] + b_h) \quad (5.13)$$

$$h_t = (1 - z_t) \times h_{t-1} + z_t \times \tilde{h}_t \quad (5.14)$$

where W_z , W_r and W_h are the weighting vectors, b_z , b_r and b_h are the bias vectors, the subscripts z , r , and h denote the update gate, the reset gate and the hidden unit, respectively.

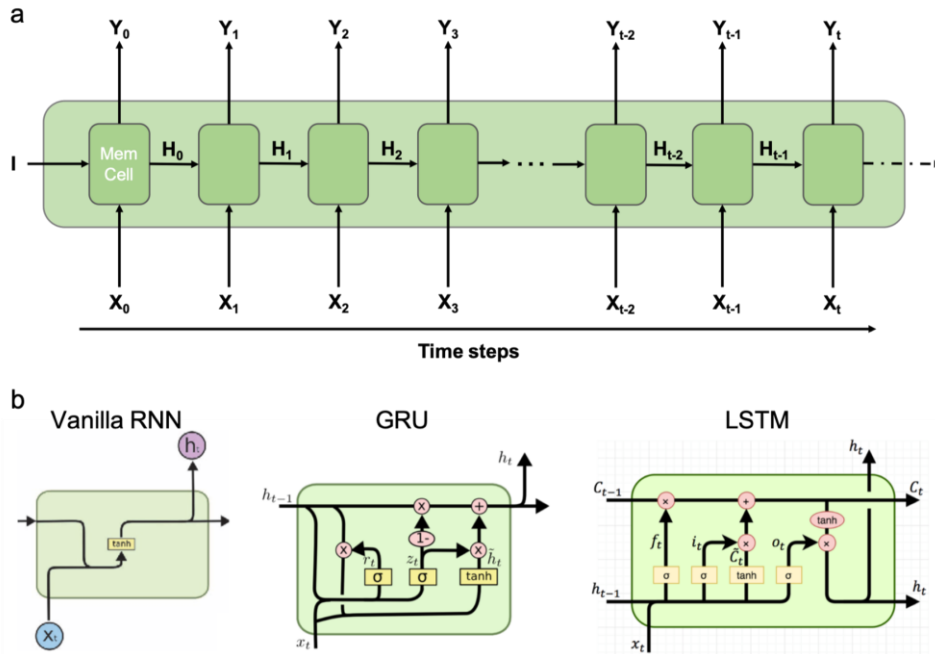


Figure 5.7 (a) A typical recurrent layer; (b) The basic structures for vanilla RNN, GRU and LSTM.

5.6.2 Proposed Architecture Structure and Hyperparameters

The proposed architecture consists of four stacked 1D convolutional layers and two stacked recurrent layers, as presented in Figure 5.5. Each convolutional layer is followed by a batch

normalization (BN) layer (Ioffe and Szegedy, 2015) to accelerate the architecture training. In addition, the only dropout layer (Srivastava et al., 2014) is added after all convolutional layers and BN layers to avoid the overfitting issue and the disharmony between the dropout and BN techniques (Li et al., 2019). The dropout rates of 0.15, 0.2 and 0.25 were attempted in this study. Besides, ReLu (Goodfellow, Bengio and Courville, 2016) is applied to each convolutional layer as the activation function, and the ‘same’ padding (Goodfellow, Bengio and Courville, 2016) is employed for the convolutions. The candidates for the number of convolutional kernels in each layer and the kernel size for the experiments using the simulated dataset are shown in Table 5.7. As for the recurrent layers, two types of cutting edge RNNs are employed in this part of the architecture, namely, LSTMs and GRUs. Both types of recurrent layers are able to learn long-term temporal correlations. ‘tanh’ activation function (Goodfellow, Bengio and Courville, 2016) is adopted for the recurrent layers. The regularization of the recurrent layers is implemented by applying ‘recurrent dropout’ to the second recurrent layer, which drops the connections between the recurrent units (Gal and Ghahramani, 2016), and the dropout rate is set to 0.1. Apart from this, the candidates for the number of recurrent units in each layer are 64, 128, 256 and 512. At last, a dense layer with the activation function of ‘softmax’ (Goodfellow, Bengio and Courville, 2016) is used after the recurrent layers to produce the probability for each class.

The architecture was trained using back-propagation (Rumelhart, Hinton and Williams, 1986) to update the weights and biases of the neurons. The back-propagation method propagates backward the loss between the predicted output and the actual output from the output layer to the input layer. The gradient of the loss is then calculated with respect to the weight and bias of each neuron, as follows:

$$\frac{\partial \mathcal{L}}{\partial W^{[l]}} = \frac{\partial \mathcal{L}}{\partial z^{[l]}} \cdot [a^{[l-1]}]^T \quad (5.15)$$

$$\frac{\partial \mathcal{L}}{\partial b^{[l]}} = \frac{\partial \mathcal{L}}{\partial z^{[l]}} \quad (5.16)$$

$$z^{[l]} = W^{[l]} \cdot a^{[l-1]} + b^{[l]} \quad (5.17)$$

where $\mathcal{L}$ is the loss function, $W^{[l]}$ is the weight vector of the l th layer, $b^{[l]}$ is the bias vector of the l th layer, $z^{[l]}$ is the weighted sum of the activations from the previous layer with a bias, $a^{[l-1]}$ is the activations from the previous layer. The gradient descent algorithm updates the weights and biases of each neuron in order to minimize the loss. In this study, the stochastic

gradient descent (SGD) algorithm (Ruder, 2017) was used as the optimizer since it performs a parameter update for each training sample and can reach fast convergence, as represented by Eqn. 5.18:

$$\theta = \theta - \eta \cdot \nabla_{\theta} J(\theta; x^{(i)}; y^{(i)}) \quad (5.18)$$

where θ is the model parameters, η is the learning rate, $J(\theta; x^{(i)}; y^{(i)})$ is the cost function for each training sample $x^{(i)}$ and its corresponding label $y^{(i)}$.

Table 5.7 Proposed FDD architecture hyperparameters for fault data experiments.

Hyperparameters	Values
Number of kernels in the input convolutional layer $K_{n,1}$	112, 150, 200, 250 and 300
Number of kernels in the second convolutional layer $K_{n,2}$	300, 350, 400, 450, 500, 550 and 600
Number of kernels in the third convolutional layer $K_{n,3}$	$2K_{n,2}$
Number of kernels in the third convolutional layer $K_{n,4}$	$2K_{n,3} \sim 3K_{n,3}$
Kernel size	3, 5, 7 and 9
Stride	1
Padding	‘same’
Convolutional dropout rate	0.15, 0.2 and 0.25
Number of units in recurrent layers	64, 128, 256 and 512
Recurrent dropout rate	0.1
Window length	6, 12 and 24
Batch sizes	32, 56, 64, 112 and 128
Training data size ratio	80%, 70% and 60%
Cross validation strategy and number of resplitting iterations	Stratified shuffle split for 10 times
Optimizer	SGD
Epochs	500 (with early stopping)
Learning rate	0.0001–0.002 (exact value depending on experiments)

Due to the multi-class classification nature of the FDD problem, categorical cross entropy was used as the loss function. To speed up the training and optimize the model performance, a set of batch sizes, including 32, 56, 64, 112 and 128, were tested for finding the best batch size. Besides, the maximum number of epochs was set to 500 and the early stopping technique was employed to stop the training if the model performances converged before the defined

maximum number of epochs. The best learning rate for each experiment was found by setting a learning rate scheduler that gradually grows from 0.00001 to 0.1 during a training practice. The whole model tuning process was conducted by attempting various combinations of the aforementioned hyperparameter candidates, and an overview of all the hyperparameters is presented in Table 5.7. It should be noted that all the experiments were conducted using a workstation with an AMD Ryzen 9 3900X 12-core processor, a Nvidia RTX 3080 graphic card and 32 GB RAM.

5.7 Model Evaluation

Due to the imbalanced multi-class classification nature of the FDD problem, some machine learning evaluation metrics, such as accuracy, may produce misleading high scores that incorrectly reflect good model performances, as data-driven FDD models have a strong capability of identifying the majority class, HVAC normal conditions. Therefore, F1 score (Sasaki, 2007) and Cohen's Kappa (CK) score (Cohen, 1960) were used to evaluate the general classification capability of the proposed FDD architectures. The average metric scores of the ten-times re-splitting iterations were used to evaluate the architectures' performances in different data experiments. The F1 score for each class is calculated using the precision and recall scores for each class as follows:

$$F1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \quad (5.19)$$

where precision and recall are calculated using the following equations:

$$Precision = \frac{TP}{TP + FP} \quad (5.20)$$

$$Recall = \frac{TP}{TP + FN} \quad (5.21)$$

Where TP represents the number of true positive predictions, *FP* represents the number of false positive predictions, *FN* represents the number of false negative predictions (Frank, et al., 2019). The Cohen's Kappa score *k* is calculated by the following equation:

$$k = \frac{p_o - p_e}{1 - p_e} \quad (5.22)$$

Where p_o represents the observed proportional agreement of the true and predicted labels, p_e represents the probability of random agreement of both true and predicted labels.

In addition, the fault detection rate, fault diagnosis rate, false positive rate and false negative rate were calculated based on the confusion matrices. The fault detection rate is defined as the ratio of the amount of detected faults to the total amount of faults. The fault diagnosis rate represents the ratio of the amount of correctly diagnosed faults (the fault type is correctly predicted) to the total amount of detected faults. The false positive rate is the ratio of the amount of the actually negative (normal conditions) samples that are predicted as positive samples (faults) to the total amount of negative samples. The false negative rate is the ratio of the amount of the actually positive (faulty conditions) samples that are predicted as negative samples (normal conditions) to the total amount of positive samples.

Chapter 6 Fault Detection and Diagnosis Evaluation Experiments

6.1 Introduction

To evaluate the FDD performance of the proposed architecture in Chapter 5, FDD evaluation experiments designed based on potential fault data scenarios in reality were conducted and the experimental results were presented in this chapter. Section 6.2 describes the potential practical fault data scenarios that were used to design FDD evaluation experiments. Section 6.3 shows FDD evaluation experiments designed based on the simulated fault dataset and two real fault datasets. Section 6.4 presents the results for a general hyperparameter analysis based on three general hyperparameters, i.e., window length, training data size and batch size. Section 6.5 discusses the results for four evaluation experiments based on the simulated fault dataset and the identified novel practical data issue. Section 6.6 discusses the experimental results for verifying the identified practical data issue based on two real fault datasets. Section 6.7 summarizes the crucial findings from these FDD evaluation experiments and recommends the potential future work. It should be noted that this chapter has been published in the paper ‘Fault data seasonal imbalance and insufficiency impacts on data-driven heating, ventilation and air-conditioning fault detection and diagnosis performances for energy-efficient building operations’ (Zhong, Calautit and Wu, 2023) and hence the content from the paper is reused in this chapter.

6.2 Potential Practical Fault Data Scenarios

There are two main data issues that have already been identified in previous literature (Fan et al., 2021a), i.e., the imbalanced distribution between normal and fault data, and the fault data insufficiency, as mentioned in section 1.2.1. Apart from the aforementioned two issues, there is another issue that has not yet been realized. According to Zhong et al. (2022), HVAC fault impacts vary with changing climate conditions, as HVAC system operating conditions change with outdoor weather conditions. Lu et al. (2021) indicated that HVAC fault impact patterns are correlated to system seasonal working conditions. However, the acquisition of fault data under all the seasons from real buildings is often difficult due to the frequency and the span of occurrences for different faults. This issue results in several possible data scenarios in reality: (1) the fault data under some classes only cover a certain season but is not available under other seasons (Fig. 6.1a); (2) the fault data under some classes covers two or more seasons, but is still lacking in remaining seasons (Fig. 6.1b); (3) the fault data under some classes

cover all the seasons, and the total amount of data seems to be sufficient, but the amount of data under each season is limited (Fig. 6.1c). These possible data scenarios may induce potential performance degradations for HVAC FDD applications in real buildings, as a sufficient amount of fault data under one or several but not all the seasons cannot guarantee that the FDD models can fully learn different fault patterns in other season(s). Thus, the potential impact of this issue on HVAC FDD performances was investigated in this study.

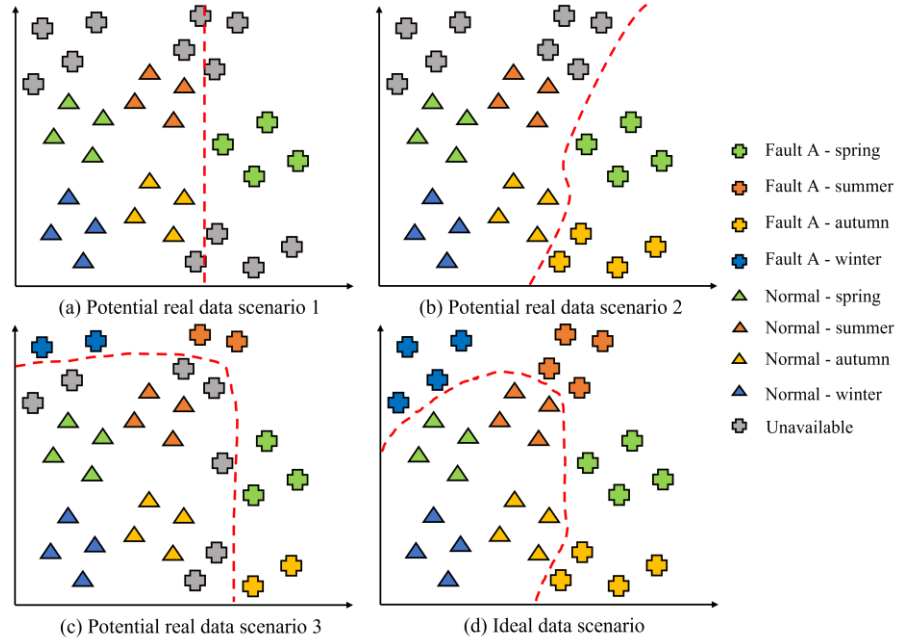


Figure 6.1 Potential examples of FDD classification for practical data scenarios (a-c) and the ideal data scenario (d) (the red dash line in each case represents the possible separation boundary predicted by the FDD model to differentiate the normal data from the fault A data).

6.3 Evaluation Experiments Setup

The experiments for assessing the FDD performances under potential practical data scenarios consist of three parts: (1) some initial experiments were conducted to find the optimal values for window length, training data size ratio and batch size, as presented in subsection 6.3.1; (2) four data experiments, considering potential practical data scenarios, were implemented based on the simulated fault dataset, as shown in subsection 6.3.2; (3) similar experiments to those based on the simulate dataset were carried out using two real fault datasets, in order to verify the findings obtained from experiments in part 2, as discussed in subsection 6.3.3.

6.3.1 Initial Data Experiments for Tuning General Hyperparameters

A hyperparameter analysis was firstly conducted to determine the values of some general parameters, including window length, training data size and batch size. The candidates for these hyperparameters, as listed in Table 5.7, were selected for evaluating the resultant FDD performances. The objective of this analysis was to fix these general hyperparameters to the optimal values, in order to save the computational resource for searching the remaining deep neural network hyperparameters for the evaluation experiments from subsection 6.3.2 and 6.3.3.

6.3.2 Practical Data Scenarios using Simulated Dataset

The simulated dataset includes four equal-sized subsets for four seasons. The sample amount of each fault class is the same in each subset. Based on the simulated dataset, four data experiments were proposed to assess the influence of the practical issue of limited fault data under one or multiple seasons on the proposed FDD architecture's performances:

- (1) The architecture was trained on a certain season's training set and tested on each season's full-size dataset, as illustrated in Figure 6.2a.
- (2) The architecture was trained on a mixed training set that contains half data from each of two different seasons and is also of the same total size as the one-season training set in experiment 1. The trained architecture was tested on each season's full-size dataset, as illustrated in Figure 6.2b.
- (3) The architecture was trained on a mixed training set that includes a quarter of data from each season and is also of the same total size as the one-season training set in experiment 1. The trained architecture was tested on each season's full-size dataset, as illustrated in Figure 6.2c.
- (4) The architecture was trained on a mixed training set that contains full-size data from each season and is four times the size of the one-season training set in experiment 1. The trained architecture was tested on each season's full-size dataset, as illustrated in Figure 6.2d.

It should be noted that the amount of normal and fault data in each scenario is stratified, which means that the imbalanced ratio of each fault class to the normal class is maintained, and the proportion of each fault class in each mixed dataset remains the same as 0.02 as that in each one-season's dataset. It should also be noted that the effect of imbalanced ratios on the FDD performances is not evaluated in this thesis since this has already been studied in (Fan et al., 2021a).

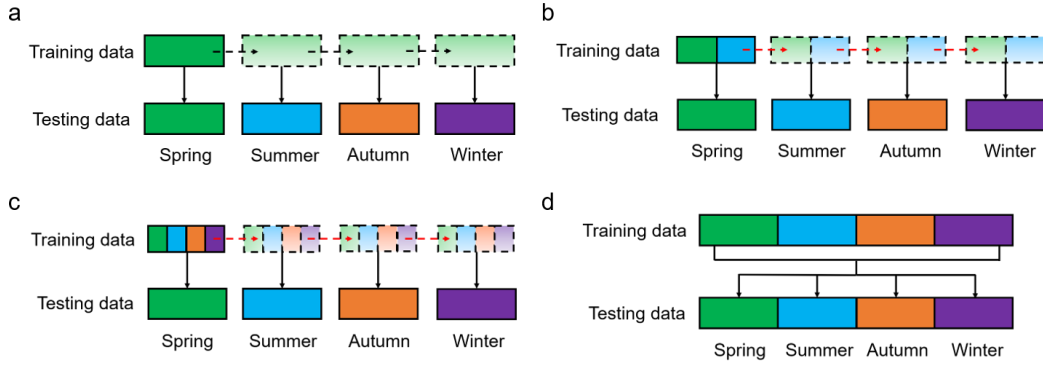


Figure 6.2 The proposed data experiments under the simulated fault dataset: (a) Training data: a one-season dataset; testing data: each season's dataset (spring as an example here). (b) Training data: a mixed dataset of two seasons of half size; testing data: each season's dataset (a mixed dataset of spring and summer as an example here). (c) Training data: a mixed dataset of four seasons of quarter size; testing data: each season's dataset. (d) Training data: a mixed dataset of four seasons of full size; testing data: each season's dataset.

6.3.3 Verification Experiments using Real Building Datasets

To verify the findings from the previous experiments detailed in subsection 6.3.2, similar data scenarios were proposed according to the ASHRAE RP-1312 dataset. The only difference between this set of experiments and the previous set of experiments in section 6.3.2 lies in the coverage of seasons. The seasons considered in this set of experiments are only spring, summer and winter due to the data availability in the dataset, as illustrated in Figure 6.3. The details of this set of experiments were explained as follows:

- (5) The architecture was trained on a certain season's training set and tested on each season's full-size dataset, as illustrated in Figure 6.3a.
- (6) The architecture was trained on a mixed training set that contains half data from each of two different seasons and is also of the same total size as the one-season training set in experiment 5. The trained architecture was tested on each season's full-size dataset, as illustrated in Figure 6.3b.
- (7) The architecture was trained on a mixed training set that includes a third of data from each season and is also of the same total size as the one-season training set in experiment 5. The trained architecture was tested on each season's full-size dataset, as illustrated in Figure 6.3c.
- (8) The architecture was trained on a mixed training set that contains full-size data from each season and is three times the size of the one-season training set in experiment 5. The trained architecture was tested on each season's full-size dataset, as illustrated in Figure 6.3d.

In addition, the NREL's dataset is fully used to produce three data scenarios, including two one-season training scenarios using the summer and winter dataset, respectively, and a mixed dual-seasons scenario of summer and winter datasets, as presented in Figure 6.4.

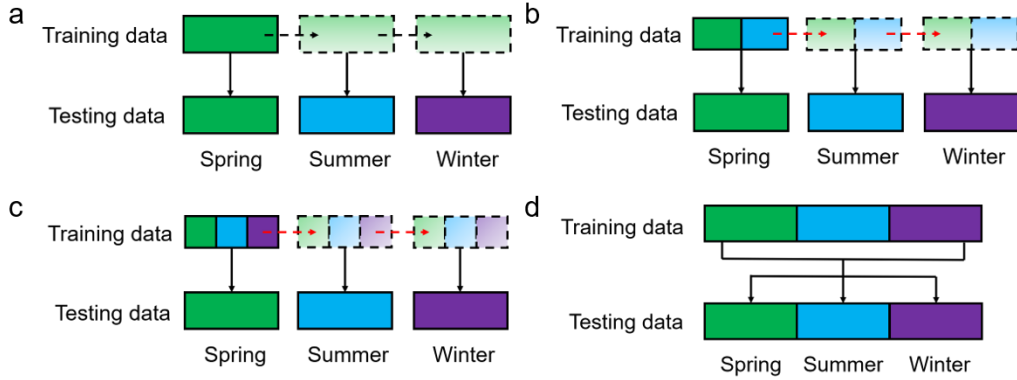


Figure 6.3. The proposed data experiments under the ASHRAE RP1312 fault dataset: (a) Training data: a one-season dataset; testing data: each season's dataset (spring as an example here). (b) Training data: a mixed dataset of two seasons of half size; testing data: each season's dataset (a mixed dataset of spring and summer as an example here). (c) Training data: a mixed dataset of three seasons of quarter size; testing data: each season's dataset. (d) Training data: a mixed dataset of three seasons of full size; testing data: each season's dataset.

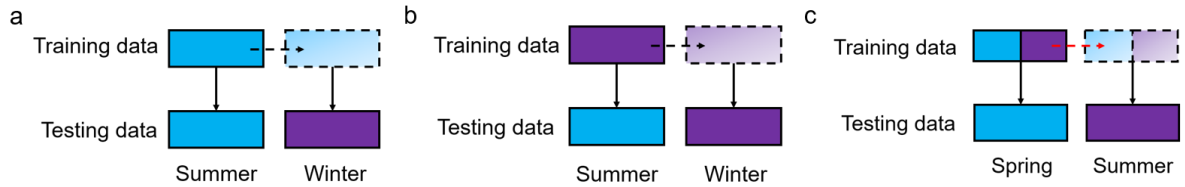


Figure 6.4. The proposed data experiments under the NREL's fault dataset: (a) Training data: the summer dataset; testing data: each season's dataset. (b) Training data: the winter dataset; testing data: each season's dataset. (c) Training data: a mixed dataset of summer and winter of half size from each; testing data: each season's dataset.

6.4 General Hyperparameter Analysis Results

This analysis was implemented using the combined CNN-LSTM architecture and the combined CNN-GRU architecture, respectively. The baseline configurations for these two architectures are presented in Table 6.1. The FDD performances, in terms of F1 and CK scores (F1 and CK scores are metrics suitable for evaluating machine learning model performance under class-imbalanced data distributions), for window lengths, training data

size ratios, and batch sizes, are illustrated in Figure 6.5. The window length of 6 contributed to the highest scores, while the window lengths of 12 and 24 led to lower and significantly lower scores. This indicates that smaller window length is more robust in capturing the fault patterns, as the smaller window length leads to a larger number of windows fed into the FDD model training when there is a limited amount of fault data available for training. As for the training data size ratio, the performance scores reduced with the decrease in the ratio. Besides, an adequate amount of data needs to be left for validating and testing the model, and thus the training data size ratio is set to 0.8 without considering higher ratios. Regarding the batch size, the peak of the performance scores can be observed at the batch size of 112. On one hand, the batch size of 128 included a higher variability in feature patterns and a higher number of windows for each class, but failed to outperform the batch size of 112, which may be due to the less times of model weights adjustments per epoch when the batch size of 128 is used. On the other hand, the smaller batches than 112 indeed increased the times for updating the model weights per epoch, however, they were likely to contain lower variability in feature patterns and some of them might not be able to take into account sufficient amount of data windows for each class. To conclude, the window length of 6, the training data size ratio of 0.8 and the batch size of 112 are used for further experiments.

Table 6.1 Baseline architecture configurations used in the general hyperparameter analysis.

Hyperparameters	CNN-LSTM	CNN-GRU
Number of kernels for Conv 1	150	200
Number of kernels for Conv 2	500	450
Number of kernels for Conv 3	1000	900
Number of kernels for Conv 4	2000	1800
Kernel size	3	3
Convolutional dropout rate	0.15	0.2
Number units for Recurrent 1	512	512
Number units for Recurrent 2	256	256
Recurrent dropout rate	0.1	0.1

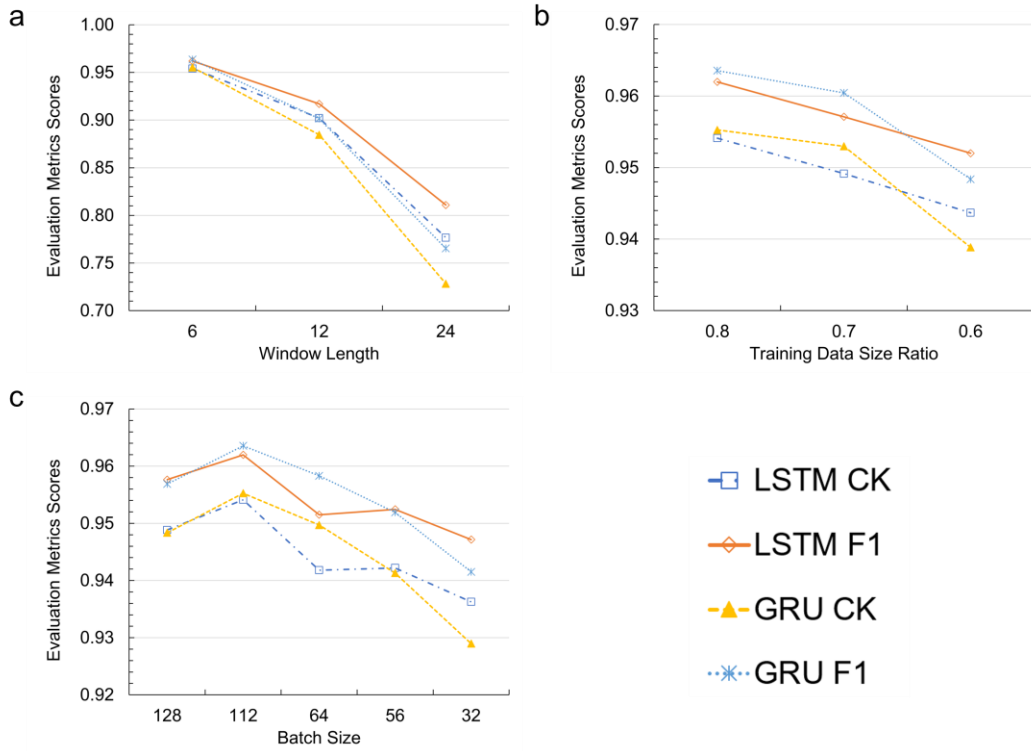


Figure 6.5. FDD performance comparisons among different (a) window lengths; (b) training data size ratios; (c) batch sizes.

6.5 Evaluation Experiments Results based on Simulated Dataset

The results obtained from the four data experiments are discussed in subsections 6.5.1-6.5.4. The highest F1 and CK scores tested under each season for both architectures, CNN-LSTM and CNN-GRU, from each experiment are displayed in Table 6.2. F1 and CK scores are metrics suitable for evaluating machine learning model performance under class-imbalanced data distributions. Based on Table 6.2, the higher F1 and CK scores between the two architectures are shown in Figure 6.6. In addition, the average fault detection, fault diagnosis, false positive and false negative rates among ten cross-validation iterations for the best architecture, trained on the spring only, mixed spring and summer, full-seasons quarter-size and full-seasons full-size dataset scenarios, are shown in Table 6.3 and illustrated in Figure 6.7.

Table 6.2 FDD highest test metrics scores under each season for the architectures trained on each dataset. (The values in italics are the larger scores between the two architectures, and the values in bold are the highest scores under this test season for each scenario, respectively).

Training data	Architecture	Spring		Summer		Autumn		Winter	
		F1	CK	F1	CK	F1	CK	F1	CK
Spring	CNN-LSTM	0.975	0.971	<i>0.081</i>	<i>0.020</i>	<i>0.074</i>	0.037	0.072	0.027
	CNN-GRU	0.975	0.970	0.079	0.009	0.072	<i>0.041</i>	<i>0.078</i>	<i>0.048</i>
Summer	CNN-LSTM	0.068	0.010	0.972	0.966	0.082	0.014	0.067	-0.007
	CNN-GRU	<i>0.075</i>	<i>0.027</i>	0.968	0.961	<i>0.089</i>	<i>0.049</i>	<i>0.079</i>	<i>0.028</i>
Autumn	CNN-LSTM	0.080	0.018	<i>0.059</i>	<i>0.015</i>	0.983	0.980	0.064	0.007
	CNN-GRU	<i>0.087</i>	<i>0.030</i>	0.057	0.006	0.984	0.982	<i>0.070</i>	<i>0.011</i>
Winter	CNN-LSTM	0.078	0.023	0.054	0.016	0.079	0.023	0.979	0.975
	CNN-GRU	<i>0.082</i>	<i>0.026</i>	<i>0.055</i>	0.016	<i>0.084</i>	<i>0.033</i>	0.982	0.979
Spring&Summer	CNN-LSTM	0.937	0.925	0.895	0.873	0.099	0.055	0.081	0.037
	CNN-GRU	0.942	0.933	<i>0.901</i>	<i>0.880</i>	<i>0.109</i>	<i>0.067</i>	<i>0.087</i>	<i>0.046</i>
Summer&Winter	CNN-LSTM	<i>0.122</i>	0.104	0.908	0.890	0.115	0.092	0.948	0.940
	CNN-GRU	0.118	0.104	0.902	0.884	0.123	0.101	0.949	0.941
Quarter size	CNN-LSTM	<i>0.862</i>	<i>0.841</i>	<i>0.790</i>	<i>0.759</i>	<i>0.862</i>	<i>0.842</i>	<i>0.858</i>	<i>0.836</i>
	CNN-GRU	0.850	0.827	0.771	0.737	0.846	0.823	0.845	0.821
Full size	CNN-LSTM	0.977	0.973	0.951	0.940	0.972	0.968	0.973	0.969
	CNN-GRU	0.978	0.974	0.958	0.948	0.977	0.973	0.978	0.974

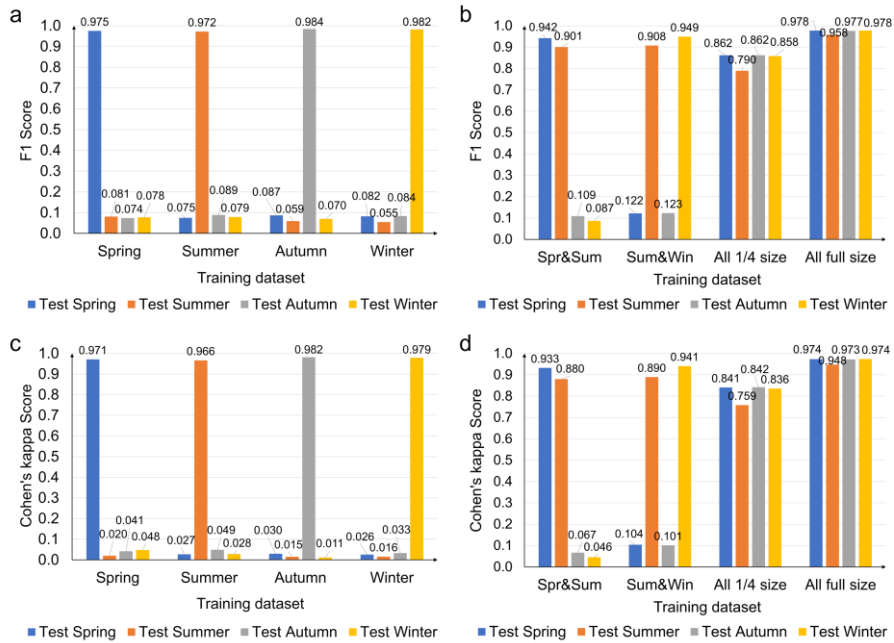


Figure 6.6. Highest F1 scores between architectures trained on (a) one-season, (b) dual-seasons and full-seasons datasets; highest CK scores between architectures trained on (c) one-season, (d) dual-seasons and full-seasons datasets ('All 1/4 size' and 'All full size' represent the full-season quarter size and full-season full size scenarios, respectively).

Table 6.3 Average fault detection, fault diagnosis, false positive and false negative rates among ten cross-validation times for the best architecture in experiments 1–4 using the simulated dataset.

Test dataset	Fault detection rate	Fault diagnosis rate	False positive rate	False negative rate
<i>Exp 1: CNN-LSTM trained on the spring dataset</i>				
Spring	0.999	1.000	0.009	0.001
Summer	0.208	0.144	0.178	0.792
Autumn	0.176	0.128	0.117	0.824
Winter	0.091	0.111	0.058	0.909
<i>Exp 2: CNN-GRU trained on the mixed spring-summer dataset</i>				
Spring	0.970	0.999	0.019	0.030
Summer	0.947	1.000	0.033	0.053
Autumn	0.122	0.411	0.066	0.878
Winter	0.071	0.254	0.029	0.929
<i>Exp 3: CNN-LSTM trained on the full-seasons quarter-size mixed dataset</i>				
Spring	0.902	0.998	0.034	0.098
Summer	0.829	1.000	0.046	0.164
Autumn	0.909	0.993	0.029	0.091
Winter	0.900	0.996	0.037	0.100
<i>Exp 4: CNN-GRU trained on the full-seasons full-size mixed dataset</i>				
Spring	0.999	1.000	0.008	0.001
Summer	0.996	1.000	0.017	0.004
Autumn	0.998	1.000	0.009	0.002
Winter	0.994	1.000	0.007	0.006

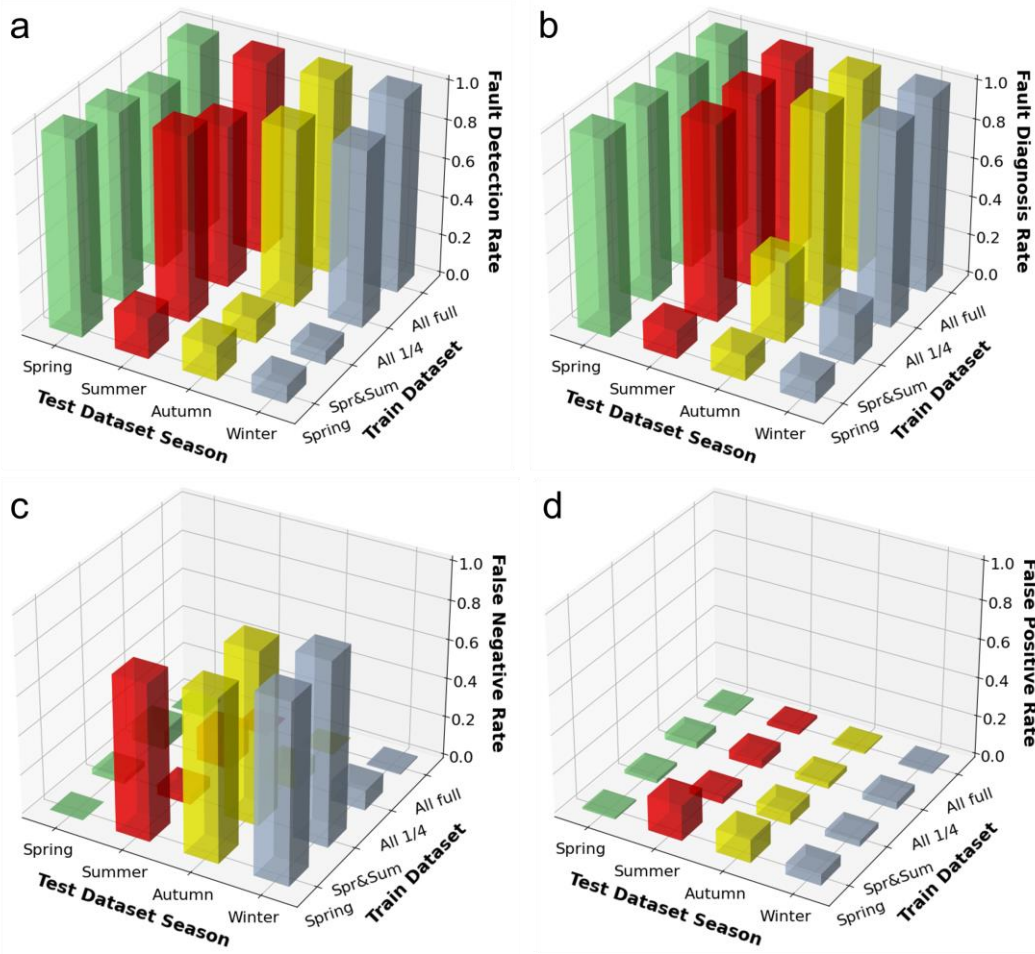


Figure 6.7 Fault detection, fault diagnosis, false negative and false positive rates tested under each season for experiments 1–4 using the simulated dataset (‘All 1/4’ represents the full-season quarter size scenario; ‘All full’ represents the full-season full size scenario).

6.5.1 Experiment 1 Results: One-Season Dataset Evaluation

Experiment 1 was conducted based on each one-season dataset. Comparing the test performances of the two architectures under the season same as the training data’s in Table 6.2, the CNN-GRU architecture outperformed another architecture CNN-LSTM under autumn and winter, but indicated slightly poorer scores under spring and summer. In general, the FDD architectures only indicated good test metrics under the seen season from the training, but showed significantly poor scores under every unseen season (Fig. 6.6a, 6.6c). The best architecture CNN-LSTM indicated significantly high fault detection and diagnosis rates (Fig. 6.7a-6.7b), and low false positive and negative rates when evaluated under the spring full-size dataset (Fig. 6.7c-6.7d), from which the training data was extracted. However, the fault detection and diagnosis rates tested on the datasets of other three seasons are considerably lower than those tested on the spring dataset (Fig. 6.7a-6.7b).

The t-stochastic neighbour embedding (t-SNE) dimensionality reduction method was used to visualize the 2D distribution of hidden features extracted by this architecture. The good separation of features of different classes indicates a robust FDD ability of the architecture. As shown in Figure 6.8a, the fault classes, labelled as 1 to 14 (see Table 5.1), are separated into clusters with obvious boundaries among them when tested under the spring full-size dataset, and only a small amount of features from fault classes overlapped with those from the normal class (labelled as 0). As for the hidden features from the test datasets of the other three seasons (Fig. 6.8b-d), there are no obvious boundaries among different classes. Under these three seasons, the data points from each fault class are scattered, and mixed with the data points from the normal class. These findings confirm the poor generalization ability of the FDD architecture, trained on sufficient fault data from a certain season, in identifying the faults under the unseen seasons. This conclusion agrees with the finding from previous fault impact analysis studies (Zhong, Calautit and Wu, 2022; Lu et al., 2021) that HVAC fault impacts vary with different climatic conditions. In addition, the fault detection and diagnosis rates decreased along with the season transitions from summer, to autumn, and to winter (Fig. 6.7a-b). Meanwhile, the false positive rate decreased (Fig. 6.7d) and the false negative rate increased with the season transitions (Fig. 6.7c). This phenomenon indicates that the model is becoming more biased towards the majority class, i.e., the normal class, from summer to winter.

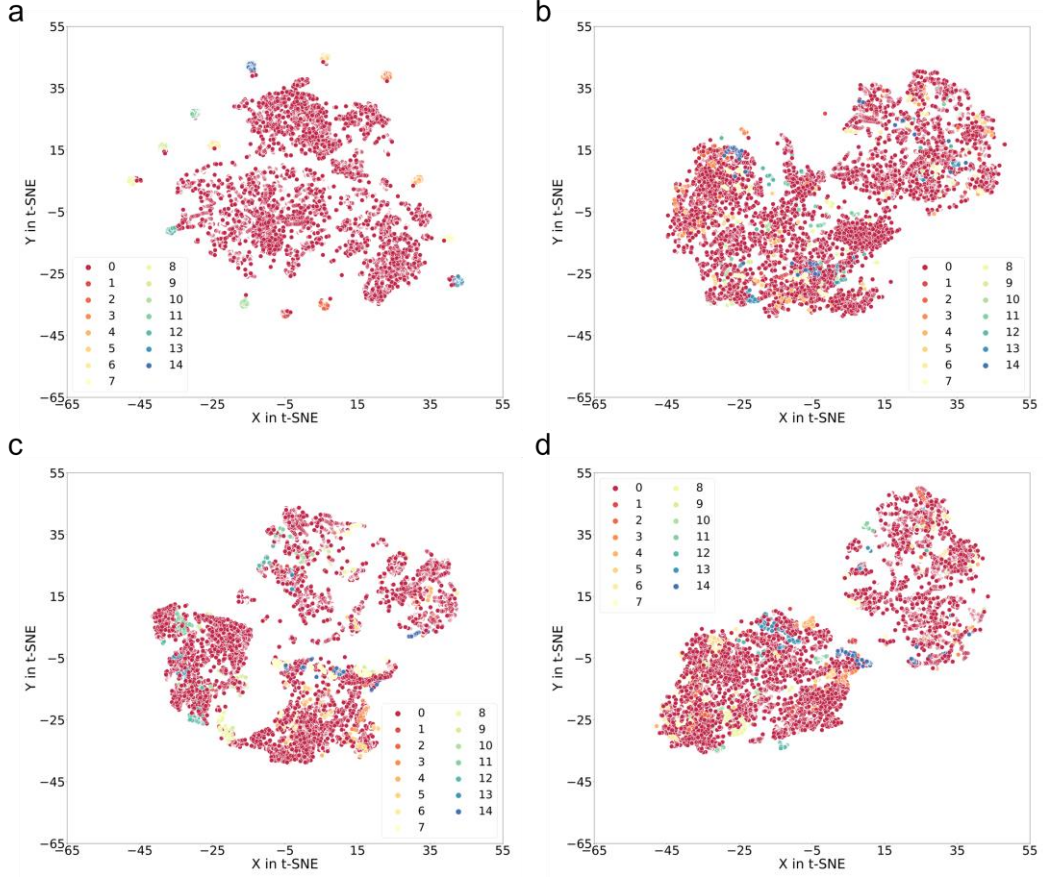


Figure 6.8 Scatterplots for the hidden features extracted by the best architecture CNN-LSTM tested under (a) spring; (b) summer; (c) autumn; (d) winter; in experiment 1 using the t-SNE method.

6.5.2 Experiment 2 Results: Dual-Seasons Dataset Evaluation

Experiment 2 was carried out on two mixed datasets of dual seasons, i.e., spring and summer, and summer and winter. Comparing the performances of the two architectures in Table 6.2, the CNN-GRU architecture trained on the mixed dataset of spring and summer outperformed the CNN-LSTM architecture when tested under each season. Besides, the CNN-GRU architecture trained on the mixed dataset of summer and winter also indicated higher performance scores when tested under autumn and winter, but showed slightly poorer scores than the CNN-LSTM architecture when tested under summer. The highest performance scores, tested under the seen seasons from training, decreased slightly in experiment 2 (Fig. 6.6b, 6.6d), compared with those from experiment 1 (Fig. 6.6a, 6.6c). This can be attributed to the less amount of fault data under each season contained in each dual-seasons' dataset from experiment 2 than that included in each one-season's dataset from experiment 1.

As shown in Table 6.3, the fault detection and diagnosis rates are higher than 0.94 and around 1, respectively, when tested under seen seasons, i.e., spring and summer, but are lower than 0.12 and 0.41, respectively, when tested under unseen seasons, i.e., autumn and winter. The considerable performance differences between seen and unseen seasons can be attributed to the differences in scatter point distributions in Figure 6.9. The feature clusters for most fault classes tested under seen seasons are separated from each other and also from the normal class (Fig. 6.9a-6.9b). In contrast, the feature clusters for the classes tested under unseen seasons are mixed up (Fig. 6.9c-6.9d). This conforms with the finding from experiment 1 that the FDD architecture has limited ability to detect and diagnose the faults from the unseen seasons. Compared with the performance indicators for the CNN-LSTM architecture trained on the spring only dataset in experiment 1, both fault detection and diagnosis rates tested on the summer full-size dataset increased substantially for this CNN-GRU architecture trained on the mixed spring and summer dataset (Fig. 6.7a-b) while the false negative rate tested on the same season's dataset decreased considerably (Fig. 6.7c).

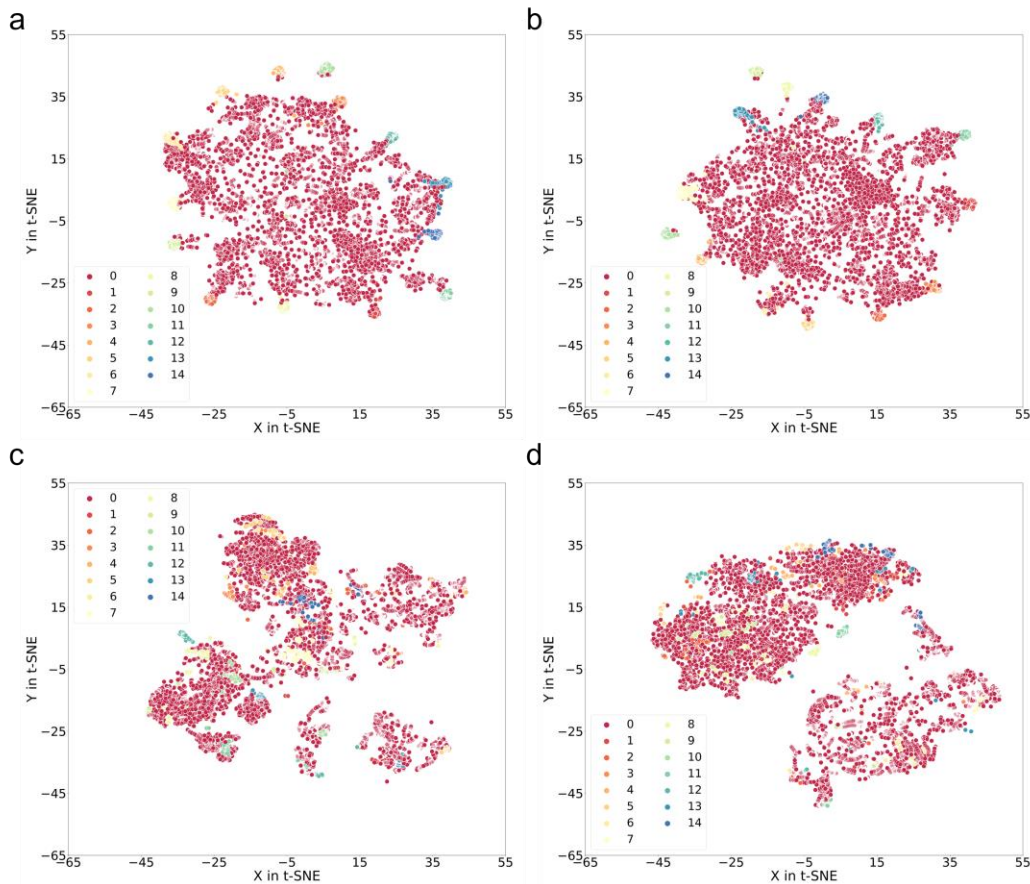


Figure 6.9 Scatterplots for the hidden features extracted by the best architecture CNN-GRU tested under (a) spring; (b) summer; (c) autumn; (d) winter; in experiment 2 using the t-SNE method.

6.5.3 Experiment 3 Results: Full-Season Quarter-Size Dataset Evaluation

Experiment 3 was implemented on a mixed dataset that contains a quarter of each season's dataset. Comparing the performances of two architectures in Table 6.2, the CNN-LSTM architecture outperformed the CNN-GRU architecture when tested under each season. Compared with the highest performance scores from previous two experiments, the scores for the unseen seasons from the previous two experiments increased substantially in experiment 3, while the scores for the seen seasons from previous experiments decreased slightly in experiment 3 due to the gradual decrease in the data amount of each seen season, i.e. full-size to half-size, and to quarter size across experiments 1-3. For example, the highest F1 score tested under spring decreased from 0.975 (Fig. 6.6a) to 0.942 (Fig. 6.6b), and further to 0.862 (Fig. 6.6b) for each best architecture trained on the spring, mixed spring and summer, and full-seasons quarter-size datasets, respectively.

By including the fault data under unseen seasons from experiments 1 and 2 into the training set of experiment 3, the architecture CNN-LSTM showed significant improvements in increasing fault detection and diagnosis rates (Fig. 6.7a-b), and mitigating false negative rate (Fig. 6.7c), when tested under the autumn and winter full-size datasets. In addition, the reduction of data amount for the seen seasons from previous experiments slightly degraded fault detection rates (Fig. 6.7a) but had no prominent impact on fault diagnosis rates (Fig. 6.7b). According to the t-SNE visualization results, the boundaries among fault classes tested under autumn and winter (Fig. 6.10c-6.10d) become apparent in comparison with those tested under the same seasons in experiment 2 (Fig. 6.9c-6.9d). In addition, some fault clusters overlapped with the normal cluster tested under spring and summer, but most fault clusters are still separated from each other (Fig. 6.10a-6.10b). This pattern explains the reason why there are slight decreases in fault detection rates under these two seasons compared with those from previous experiments. In conclusion, the coverage of fault data under different seasons is more crucial in enhancing the FDD performances than the amount of fault data used for training under each single season.

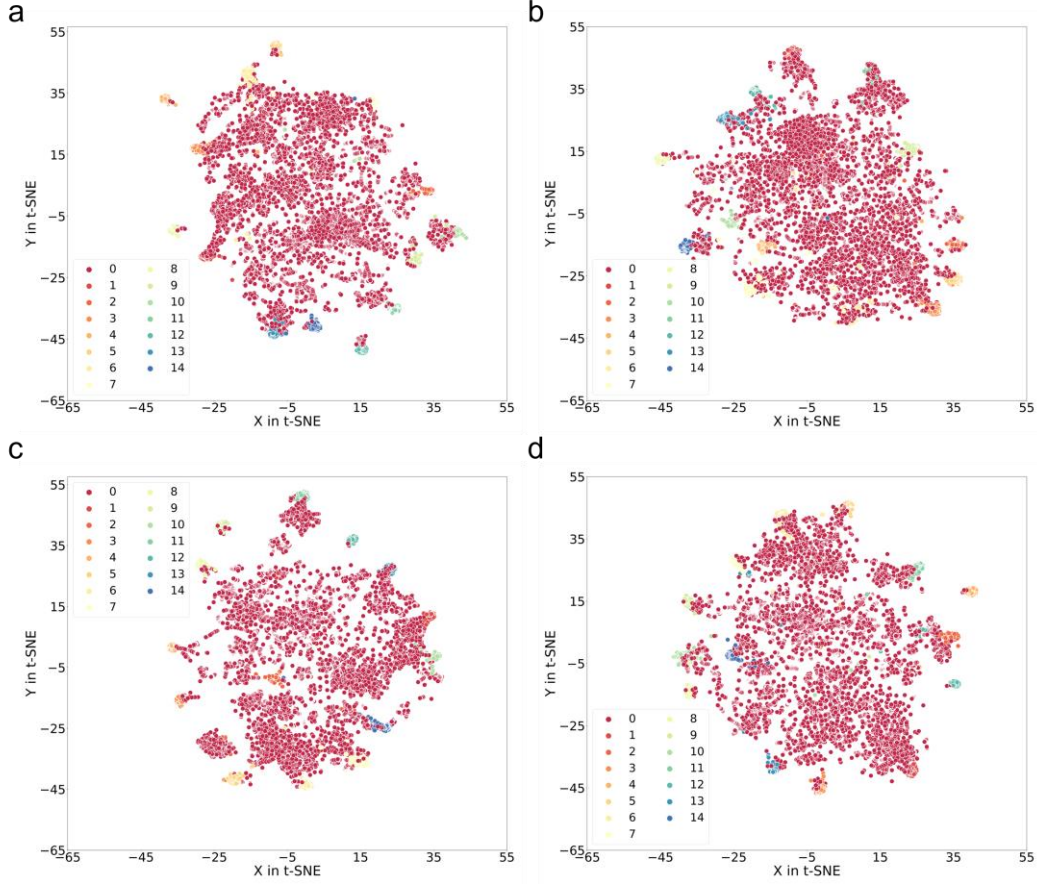


Figure 6.10 Scatterplots for the hidden features extracted by the best architecture CNN-LSTM tested under (a) spring; (b) summer; (c) autumn; (d) winter; in experiment 3 using the t-SNE method.

6.5.4 Experiment 4 Results: Full-Season Full-Size Dataset Evaluation

Experiment 4 was performed on a mixed dataset that contains full-size datasets of each season. As presented in Table 6.2, the CNN-GRU architecture slightly outperformed the CNN-LSTM architecture when tested under each season. The highest F1 scores tested under each season are within the range of 0.96 to 0.98 (Fig. 6.6b), which demonstrates excellent FDD performances under all the seasons. Compared with the highest F1 scores from experiment 3, the F1 scores under this experiment increased by around 0.12 to 0.17 (Fig. 6.6b). With the increased amount of fault data under each season compared with experiment 3, the best architecture CNN-GRU in this experiment is capable of capturing the fault patterns under all the seasons. As shown in Table 6.3, the architecture indicated robust performances when tested under all the seasons, as the fault detection and diagnosis rates under each season are both around 1. Besides, both false negative and positive rates are less than 0.02 under all seasons. The t-SNE visualization results (Fig. 6.11) also demonstrate that the boundaries

among fault classes are considerably apparent and all the fault classes' clusters are separated from the normal class cluster.

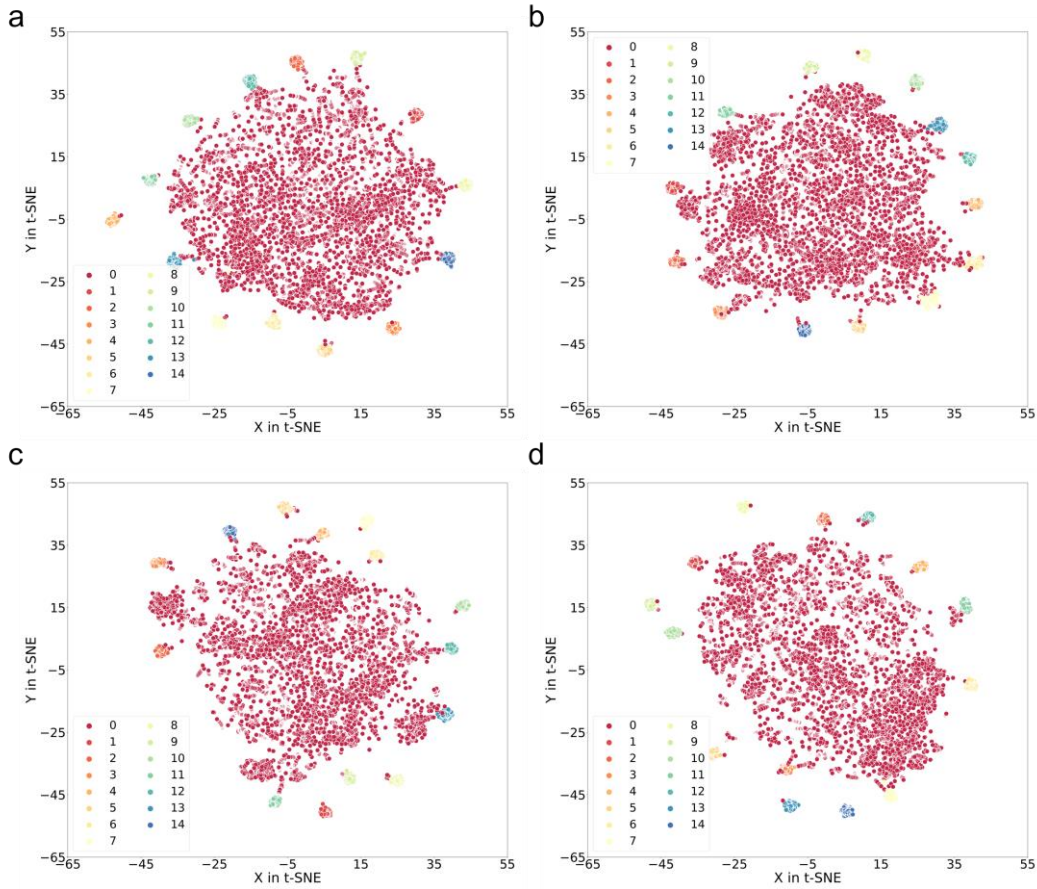


Figure 6.11 Scatterplots for the hidden features extracted by the best architecture CNN-GRU tested under (a) spring; (b) summer; (c) autumn; (d) winter; in experiment 4 using the t-SNE method.

6.6 Practical Issue Verification Experiments Results

To verify the findings obtained from the data experiments using the simulated fault dataset, data experiments 5-8 based on the ASHRAE RP1312 dataset and data experiments 9-11 based on the NREL's dataset were evaluated by the CNN-LSTM and CNN-GRU architectures. The same performance indicators as those used in experiments 1-4 are presented in Tables 6.4 and 6.5, respectively. Comparing the fault detection and fault diagnosis rates tested on each season's full-size dataset in experiment 5, the best architecture CNN-LSTM trained on the spring dataset performs well on the spring full-size dataset with both indicators of around 1, but shows lower values of both indicators tested under summer and winter (Fig. 6.12a-6.12b). Similarly, in experiments 9 and 10, the fault detection and fault diagnosis rates tested under

the unseen season are averagely 0.17 and 0.77 lower than those tested under the seen season. Apart from this, experiment 6 demonstrates that the fault detection rate tested under the unseen season, i.e., winter, is averagely 0.04 lower than that tested under the seen seasons, i.e., spring and summer, and the fault diagnosis rate under the unseen season is averagely 0.23 lower than that tested under the seen seasons. Therefore, the limited FDD ability for the faults from the unseen seasons is also observed in real fault data experiments.

However, a different observation in experiments 5–8 from experiments 1–4 is the variations in each performance indicator with the changes of fault data amount under spring are slight, as shown from the green bars in Figure 6.12a–6.12d. Besides, the decrease of fault data amount in the training set from full-size (exp. 9 and 10) and to half-size (exp. 11) led to fault detection rate reductions of 0.04 and 0.01 and fault diagnosis rate reductions of 0.07 and 0.03 under summer and winter, respectively. Therefore, the FDD performances can be improved by increasing the amount of fault data under a season, but the improvements are also limited by other factors, such as the number of faults and the ratio of fault to normal data included in the training set. If the number of faults considered for the FDD task is low or the ratio of fault to normal data is high, the increase in the amount of fault data under a season may have a limited effect on the FDD performance enhancement.

Moreover, the FDD performance enhancement led by expanding the seasonal coverage of fault data is also prominent in real fault data experiments. As presented in Figure 6.12a-6.12b, the fault detection and fault diagnosis rates tested under the unseen seasons, i.e., summer and winter, for the best architecture trained on the spring only dataset (exp. 5) are lower than those tested under the same seasons for the best architecture trained on the full-seasons quarter dataset (exp. 7). Similarly, the best architecture trained on the mixed summer and winter dataset indicated higher fault detection and diagnosis rates (Fig. 6.13a-6.13b) and lower false negative and positive rates (Fig. 6.13c-6.13d) than those tested under the unseen season, i.e., winter in experiment 9 and summer in experiment 10, for the best architecture trained on the summer only, and winter only dataset, respectively.

To summarize, as demonstrated by the simulated and real fault data experiments, the FDD architectures have limited generalization ability to detect and identify the faults from unseen seasons, and including fault data from a wider coverage of seasons into the training set has a

significant impact on improving the FDD performances. In addition, the FDD performances can be improved by increasing the fault data amount under a season, but the improvements are affected by the number of faults and the ratio of fault to normal data included in the training set.

Table 6.4 Average fault detection, fault diagnosis, false positive and false negative rates among ten cross-validation times for the best architecture in experiments 5–8 using the ASHRAE RP-1312 dataset.

Test dataset	Fault detection rate	Fault diagnosis rate	False positive rate	False negative rate
<i>Exp 5: CNN-LSTM trained on the spring dataset</i>				
Spring	0.9975	0.9994	0.0003	0.0025
Summer	0.6388	0.6297	0.0868	0.3612
Winter	0.8152	0.7877	0.2154	0.1848
<i>Exp 6: CNN-GRU trained on the mixed spring-summer dataset</i>				
Spring	0.9910	0.9980	0.0039	0.0090
Summer	0.9978	0.9980	0.0015	0.0022
Winter	0.9520	0.7659	0.3073	0.0480
<i>Exp 7: CNN-LSTM trained on the full-seasons quarter-size mixed dataset</i>				
Spring	0.9909	0.9978	0.0035	0.0091
Summer	0.9955	0.9974	0.0018	0.0045
Winter	0.9977	0.9992	0.0007	0.0023
<i>Exp 8: CNN-LSTM trained on the full-seasons full-size mixed dataset</i>				
Spring	0.9953	0.9991	0.0013	0.0047
Summer	0.9998	0.9998	0.0009	0.0002
Winter	0.9999	1.0000	0.0005	0.0001

Table 6.5 Average fault detection, fault diagnosis, false positive and false negative rates among ten cross-validation times for the best architecture in experiments 9–11 using the NREL’s dataset.

Test dataset	Fault detection rate	Fault diagnosis rate	False positive rate	False negative rate
<i>Exp 9: CNN-LSTM trained on the summer dataset</i>				
Summer	0.9979	0.9925	0.0240	0.0021
Winter	0.8208	0.2158	0.4958	0.0021
<i>Exp 10: CNN-LSTM trained on the winter dataset</i>				
Summer	0.8317	0.2267	0.7892	0.1683
Winter	0.9983	0.9996	0.0153	0.0017
<i>Exp 11: CNN-GRU trained on the mixed summer and winter dataset</i>				
Summer	0.9554	0.9206	0.0608	0.0446
Winter	0.9933	0.9664	0.0250	0.0067

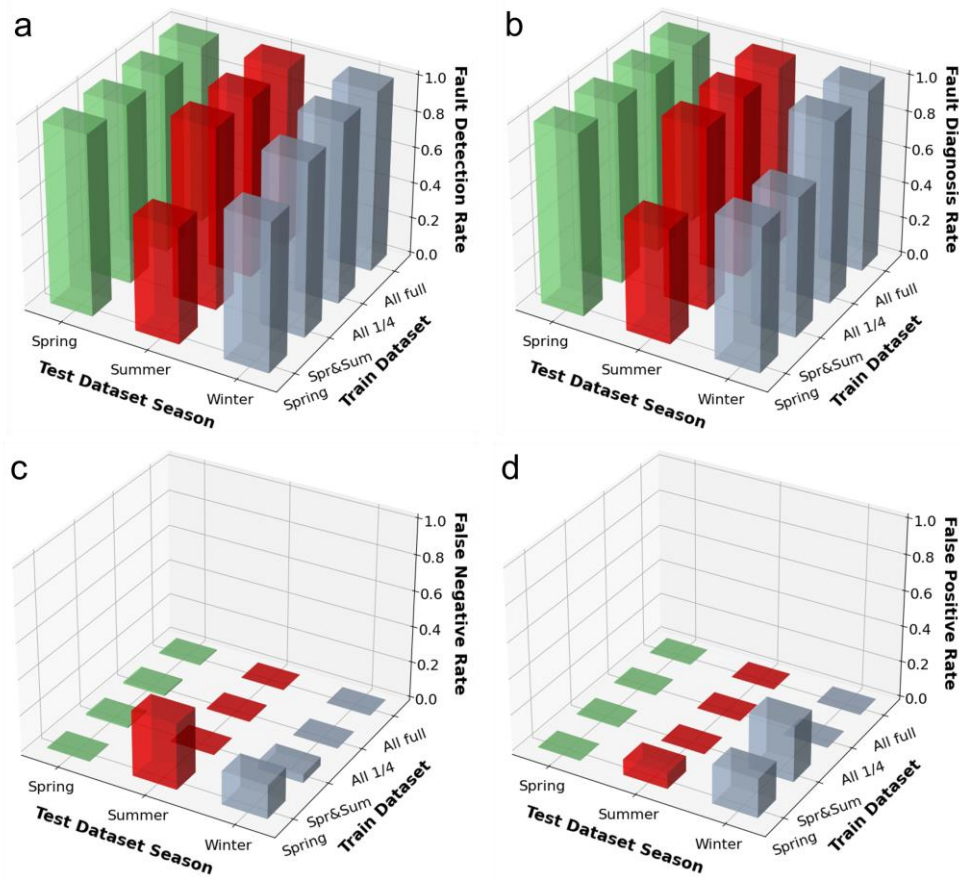


Figure 6.12 Fault detection rate, fault diagnosis rate, false negative rate and false positive rate tested under each season for experiments 1–4 using the ASHRAE RP1312 dataset (‘All 1/4’ represents the full-season quarter size scenario; ‘All full’ represents the full-season full size scenario).

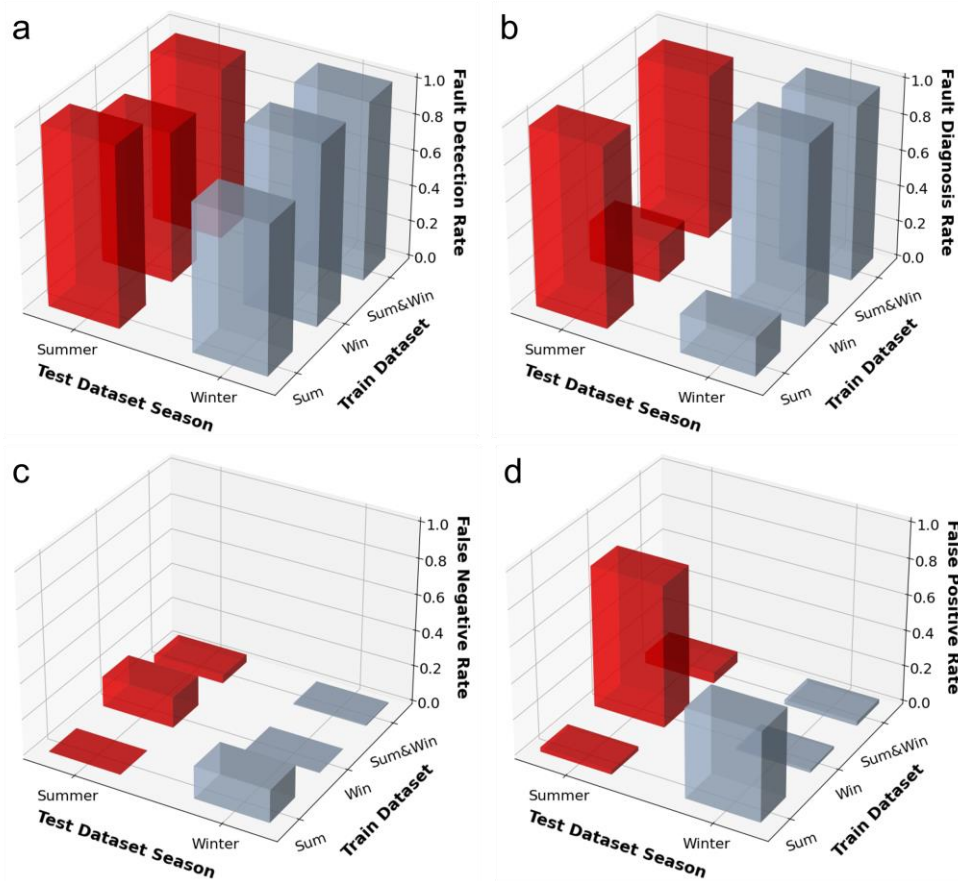


Figure 6.13 Fault detection rate, fault diagnosis rate, false negative rate and false positive rate tested under each season for experiments 1–4 using the NREL’s dataset.

6.7 Fault Detection and Diagnosis Evaluation Summary

This chapter evaluated FDD performance differences of the proposed architecture under fault data seasonal insufficient and imbalanced scenarios and an ideal fault data scenario covering climatic conditions from four seasons. The fault and normal data were gathered from fault simulations using a verified prototype building EnergyPlus model and two fault datasets from real buildings. Four evaluation experiments using the simulated fault dataset were proposed based on the practical issue related to the fault data season coverage and the amount of fault data within each season, in order to assess FDD performances under different data scenarios. Moreover, two sets of further experiments based on each of the real fault datasets were implemented to validate the practical issue. The main findings of this chapter are as follows: (1) The smaller window length is more robust in capturing the fault patterns, as the smaller window length leads to a larger number of windows fed into the FDD model training when there is a limited amount of fault data available for training.

- (2) The FDD architectures trained on sufficient fault data under a certain season(s) show poor generalization ability to detect and identify the faults under other unseen seasons.
- (3) The FDD performance under a season can be enhanced by increasing the amount of fault data under this season, but the enhancement is limited by other factors, such as the number of faults and the imbalance ratio of fault to normal data included in the training set.
- (4) The coverage of fault data under different seasons is more crucial in enhancing the FDD performances than the amount of fault data under a certain season(s).

Based on the findings of this study, the importance of fault data covered under different seasonal climatic conditions is highlighted. The results of this study will help guide FDD researchers to develop new data-driven FDD methods or revise existing methods considering this practical issue, which is highly intended to lead to FDD performance gaps between development tests and applications in real buildings. Future work can be expanded to a broader topic, evaluating the influence of the fault data intra-class imbalance under different system working conditions on FDD performances. In addition, data-driven generative techniques can be employed to produce fault data under the seasons when the fault data is unavailable.

Chapter 7 Integration of Generative Modeling with Data-driven Fault Detection and Diagnosis

7.1 Introduction

Based on the findings obtained from FDD evaluation experiments in Chapter 6, the absent fault data under certain season(s) leads to the fault data intra-class imbalance among different seasons, thus resulting in FDD performance degradations under FDD model's unseen seasons. To address this practical issue, different variants of GANs, a type of cutting-edge deep generative networks, have been explored in this chapter to produce synthetic fault data for seasons when fault data are insufficient or absent. Section 7.2 introduces different variants of GANs, namely, conditional GAN (CGAN), WGAN and the combination of them, CWGAN. Section 7.3 explains the proposed multi-label CWGAN (ML-CWGAN)'s workflow and the integration of a fault data generation module with an FDD module as an entire framework. Section 7.4 describes the settings for GAN evaluation experiments. Section 7.5 shows the baseline FDD performance of the FDD classifiers. Section 7.6 presents GAN-aided FDD performances by the proposed framework and other GAN-based frameworks for the comparison purpose. This section also discusses the FDD performance enhancement by each framework and the performance gap between GAN-aided FDD using the aggregated real and synthetic fault samples, and baseline FDD using the real fault samples of the same amount as those used in GAN-aided FDD.

7.2 Generative Adversarial Network Introduction

GAN (Goodfellow et al., 2014) is a robust unsupervised deep learning algorithm that can implicitly learn the complex patterns of high-dimensional data samples and utilize the captured patterns to generate new data samples. The basic mechanism of GANs works as follows: there are typically two components in the GAN framework, i.e., a generator and a discriminator, as illustrated in Figure 7.1. The generator is used to produce synthetic samples based on input random noises, and the role of discriminator is to distinguish between the synthetic samples (fake samples) and the real samples, outputting the validity of each input sample as real or fake. The objective of the discriminator is to effectively identify the synthetic samples from the real samples. The objective of the generator is to fool the discriminator by generating samples similar to real samples. Hence, the two components work

in an adversarial mode to improve the quality of synthetic samples. The objective function for training the original GAN is as follows:

$$\min_G \max_D V(D, G) = E_{x \sim P_{data(x)}} [\log D(x)] + E_{z \sim P_z(z)} [\log (1 - D(G(z)))] \quad (7.1)$$

Where $V(D, G)$ is the objective function, D is the function of the discriminator and G is the function of the generator. x is a real sample and z is a random noise input to the generator. $x \sim P_{data(x)}$ represents the sample x is extracted from the real sample dataset and $z \sim P_z(z)$ represents the noise input is extracted from a random distribution. The training process of GANs ideally ends until the synthetic samples generated by the generator cannot be identified from the real samples by the discriminator.

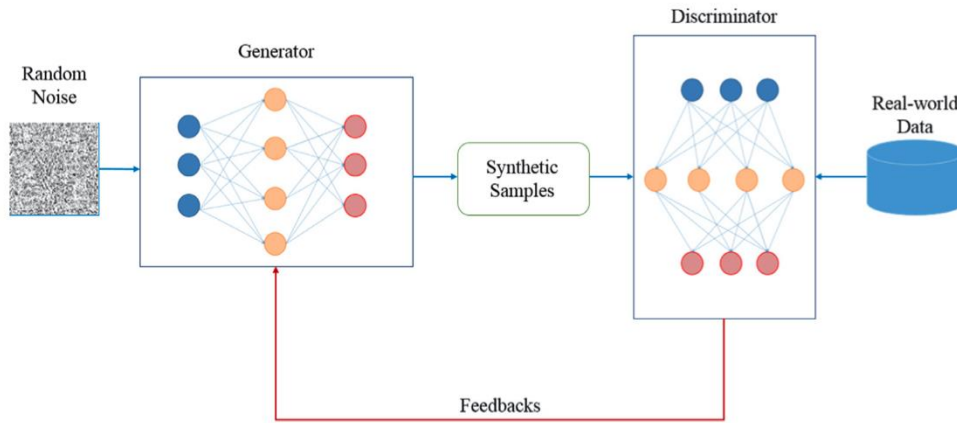


Figure 7.1 The diagram of a typical original GAN framework (Yan et al., 2020).

Since the first GAN was developed in 2014 by Goodfellow et al. (2014), a series of extended architectures to the original GAN was proposed. Two basic GAN extensions, CGAN and WGAN, and their combination have been widely utilized in data-driven HVAC FDD studies (Zhong et al., 2019; Yan et al., 2020). Thus, these extended architectures will be specifically introduced in this section. Subsection 7.2.1 describes the CGAN and its objective function. Subsection 7.2.2 introduces the WGAN and its objective function. Subsection 7.2.3 explains the benefits of combining the CGAN and the WGAN into a CWGAN and the CWGAN's performance on mitigating the impact of imbalanced data class issues on classification performances.

7.2.1 Conditional Generative Adversarial Network (CGAN)

CGAN (Mirza and Osindero, 2014) is a variant of the original GAN framework which allows the class label to be input into the generator and discriminator. Unlike the original GAN which neglects the class label and requires generating samples under each class separately, the CGANs can generate samples of different classes simultaneously by taking into account the label information, as presented in Figure 7.2. Hence, the CGANs are suitable for synthesizing fault samples for fault classes that have limited amount of samples. The training objective function for the CGAN is as follows:

$$\min_G \max_D V(D, G) = E_{x \sim P_{data}(x)} [\log D(x|c)] + E_{z \sim P_z(z)} [\log (1 - D(G(z|c)))] \quad (7.2)$$

where c is the class label. The generator takes the noises and class labels as inputs and outputs the synthetic samples. The discriminator takes synthetic and real samples, and their corresponding class labels as inputs, and predicts the validity of each sample. By incorporating label information into the GAN framework, CGANs can provide more accurate data sample generations than the original GAN.

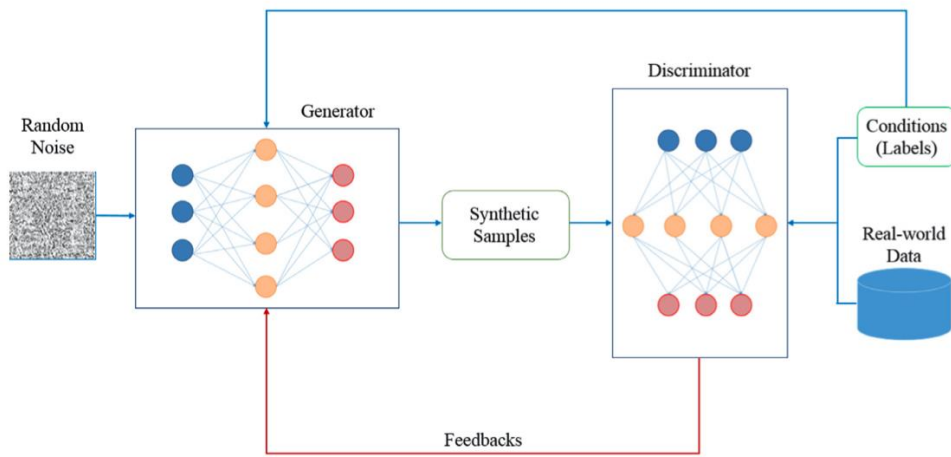


Figure 7.2 The diagram of a CGAN framework (Yan et al., 2020).

7.2.2 Wasserstein Generative Adversarial Network (WGAN)

WGAN (Arjovsky, Chintala and Bottou, 2017) is another variant of the original GAN which aims to stabilize the training of the original GAN. Wasserstein distance, which measures the discrepancy between the synthetic samples and real samples, was applied to modify the original GAN's loss function. The WGAN replaces the discriminator in the original GAN

framework with a critic which predicts the realness of a sample using a score. The training objective function for the WGAN is as follows:

$$L = E_{x \sim P_{data}}[D(x)] - E_{x \sim P_G}[D(x)] \quad (7.3)$$

$$V_G = -E_{x \sim P_G}[D(x)] \quad (7.4)$$

$$V_C = E_{x \sim P_G}[D(x)] - E_{x \sim P_{data}}[D(x)] \quad (7.5)$$

Where L is the overall objective function, V_G and V_C are the objective functions for the generator and critic, respectively.

7.2.3 Conditional Wasserstein Generative Adversarial Network (CWGAN)

By integrating the Wasserstein distance with the CGAN, the CWGAN is able to learn the patterns of real samples in a stable training process and to produce accurate synthetic samples for each class. Zheng et al. (2020) compared the performance of CWGAN on addressing imbalanced data issue in machine learning classification tasks, and indicated the CWGAN outperformed WGAN, CGAN and the original GAN in enhancing the classification performances under imbalanced data scenarios. Hence, the CWGAN is a suitable solution to address the imbalanced data issue in multi-class classification problems, which is ideal for dealing with the fault and normal data class imbalance (Yan, Chong and Mo, 2020; Yan et al., 2020), and one of the objectives of this thesis, addressing fault data intra-class imbalance among different seasons for data-driven HVAC FDD tasks.

7.3 Generative Adversarial Network (GAN)-aided FDD Framework

To tackle the fault data intra-class imbalance among different seasons, a ML-CWGAN was proposed in this thesis to support effective FDD performances under fault data seasonal intra-class imbalanced scenarios. Multi-label GAN frameworks have only been applied into a few studies including electronic health record data generation (Choi et al., 2017; Lu et al., 2023) and food image generation (Liu, Niu and He, 2023), and have not been used for synthetic fault data generation. As shown in Figure 7.3a, the proposed ML-CWGAN incorporates both fault and season labels into its training process, which aims to generate fault samples under each specific season. The fault and season label information for each sample is one-hot encoded together into a label vector. The label vector is fed into the generator with the noises sampled from a random distribution, in order to produce synthetic fault samples. The synthetic fault samples and their corresponding labels are then input into the critic with the

real samples and their corresponding labels. After a training iteration of the critic, the critic provides the realness score for each sample and these scores are then used to update the weights of the generator and the critic for the next iteration.

The training of the ML-CWGAN is ideally ended until the convergence of the generator loss. However, it is considerably difficult and computational costly to reach a convergence for CWGAN training. In this thesis, the maximum number of training epochs is set and the quality of synthetic fault samples is not assessed depending purely on the generator loss, but evaluated by the baseline FDD classifier developed in chapter 5. Hence, the entire GAN-aided FDD framework consists of a fault data generation module and a fault data quality evaluation module based on the FDD classifier performance assessment, as shown in Figure 7.3b.

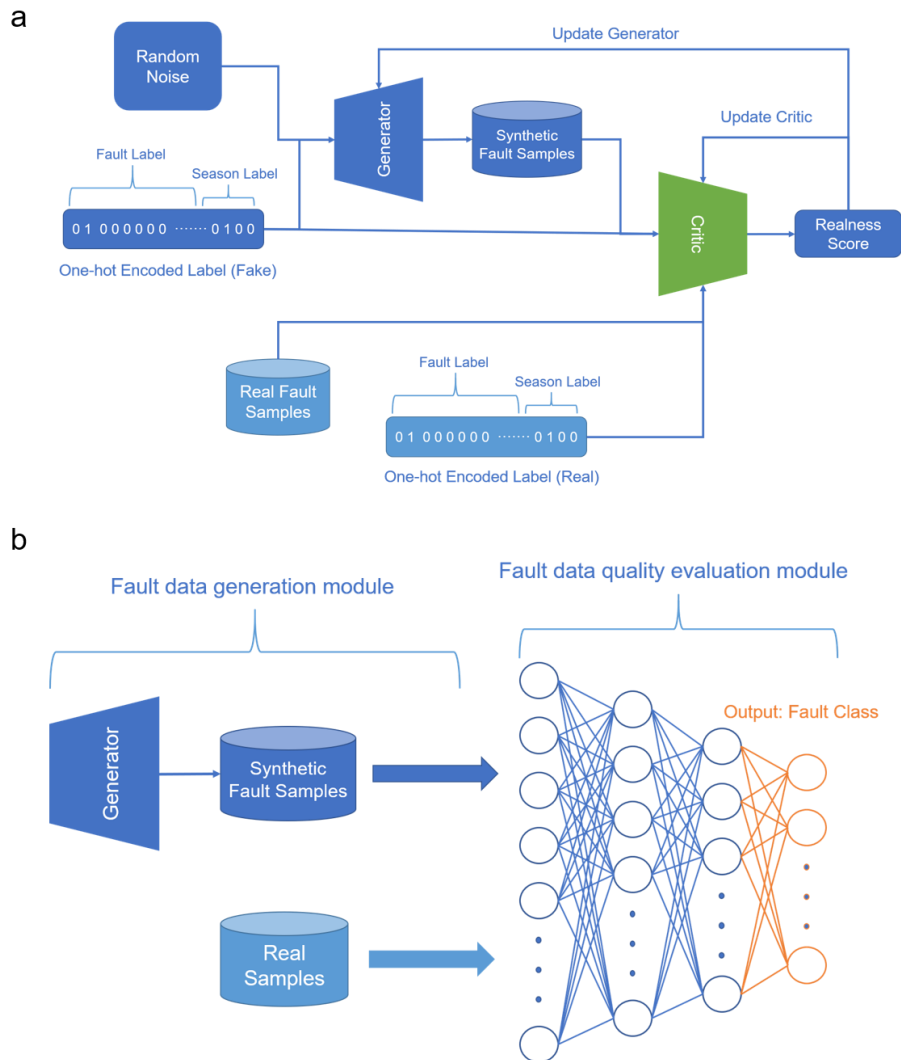


Figure 7.3 (a) The diagram of a Multi-label CWGAN (ML-CWGAN) framework; (b) The diagram of the entire GAN-aided FDD Framework.

7.4 Generative Adversarial Network Evaluation Experiments Setup

The evaluation experiments for the proposed ML-CWGAN consist of three stages: (1) Initial experiments were carried out to evaluate the influences of the data amount per fault class and the normal to fault data imbalance ratio on the baseline FDD performance, in order to determine the values for these two factors which were held constant for the further experiments; (2) Baseline FDD performance evaluation experiments that were assessed under 14 fault data seasonal imbalance scenarios and 2 seasonal balanced but insufficient scenarios, and an ideal benchmark scenario representing sufficient real fault samples under each season; (3) GAN-aided FDD performance evaluation experiments under the same data scenarios to the baseline's for GAN-based architectures: single-label CWGAN (SL-CWGAN), multi-label CGAN (ML-CGAN) and ML-CWGAN. The setup for each stage of evaluation experiments is explained in subsections 7.4.1-7.4.3, respectively.

7.4.1 Fault Data Amount and Imbalance Ratio Selection

To determine the values for the controlled variables, i.e., the data amount per fault class and the normal to fault data imbalance ratio, for further evaluation experiments of the proposed framework, a set of initial experiments with different pairs of these two controlled variables were implemented. The data amount per fault class was held constant to 250 at first and the influences of the normal to fault data imbalance ratios, containing 28, 21, 14, 7 and 1, on the baseline FDD performance were tested. It should be noted that the normal to fault data imbalance ratio of 28 is the maximum imbalance ratio in the simulated dataset. As shown in Table 7.1, the imbalance ratio of 28 led to the lowest test F1 scores (F1 score is a metric suitable for evaluating machine learning model performance under class-imbalanced data distributions) for the baseline FDD classifiers, CNN-GRU and CNN-LSTM architectures. Hence, the imbalance ratio of 28, representing the data scenario with the most significant impact on the FDD performance, was selected. After fixing the imbalance ratio to 28, the effects of different data amounts per class, ranging from 50 to 250, on FDD performances were evaluated. As a result, the data amount of 50 per class induced the lowest test F1 scores, and therefore was selected.

Table 7.1 Initial experiments for assessing the influences of the data amount per fault class and the normal to fault data imbalance ratio on the baseline FDD performance.

Data amount for each fault class	Normal/fault imbalance ratio	Testing F1 score (GRU)	Testing F1 score (LSTM)
250	28	0.985	0.982
250	21	0.987	0.984
250	14	0.988	0.988
250	7	0.991	0.990
250	1	0.988	0.987
200	28	0.985	0.981
150	28	0.981	0.976
100	28	0.977	0.970
50	28	0.951	0.927

7.4.2 Baseline FDD Performance Evaluation Experiments

The baseline FDD performance was evaluated under a series of fault data seasonal-imbalanced and seasonal-balanced but insufficient scenarios described in subsection 7.4.2.1 and the setup for the evaluation experiments was explained in subsection 7.4.2.2.

7.4.2.1 Baseline FDD Evaluation Data Scenarios

The baseline data scenarios for assessing baseline FDD performances were proposed in Table 7.2. The ideal benchmark scenario consists of 250 real fault samples for each class per season. The scenarios 1 and 16 contain 50 and 25 real fault samples for each class per season, respectively, resulting in two seasonal balanced but insufficient scenarios. The remaining seasonal imbalanced scenarios contain real fault samples ranging from 25 to 50 for each class per season.

Table 7.2 The seasonal-imbalanced and seasonal-balanced but insufficient data scenarios used for evaluating the baseline FDD performances.

Scenario	Real fault data amount per class				Total fault data amount
	Spring	Summer	Autumn	Winter	
Benchmark	250	250	250	250	14000
1	50	50	50	50	2800
2	50	50	50	25	2450
3	50	50	25	50	2450
4	50	25	50	50	2450
5	25	50	50	50	2450
6	50	50	25	25	2100
7	50	25	50	25	2100
8	50	25	25	50	2100
9	25	25	50	50	2100
10	25	50	25	50	2100
11	25	50	50	25	2100
12	25	25	25	50	1750
13	25	25	50	25	1750
14	25	50	25	25	1750
15	50	25	25	25	1750
16	25	25	25	25	1400

7.4.2.2 Baseline FDD Evaluation Experiments Setup

The CNN-GRU and CNN-LSTM classifiers proposed in section 5.6 were used as the baseline FDD classifiers. The hyperparameters for the two classifiers were searched based on the candidates provided in Table 5.6. After the search of optimal hyperparameters, two sets of hyperparameters were determined for the two classifiers, respectively, as listed in Table 7.3. The input data batch size and convolutional kernel size were set to 64 and 3, respectively, for both classifiers. The number of convolutional kernels per layer was defined as 112, 400, 800, 2400 for the CNN-GRU classifier, and 112, 500, 1000, 3000 for the CNN-LSTM classifier. The dropout rate for the dropout layer after all convolutional layers and BN layers was set to 0.2 and 0.15 for CNN-LSTM and CNN-GRU, respectively. The number of recurrent units for each layer was determined as 256 and 128 for both classifiers. The dropout rate for the recurrent units was defined as 0.1. Next, both classifiers with the optimal set of hyperparameters were trained using fault data scenarios presented in Table 7.2 and tested

under each season's full-size dataset including real samples only. The test F1 score per season under each data scenario was used for evaluating baseline FDD performances.

Table 7.3 The optimal set of hyperparameters for CNN-LSTM and CNN-GRU, respectively.

Hyperparameters	CNN-LSTM	CNN-GRU
Batch size	64	64
Number of convolutional kernels for each layer	112, 500, 1000, 3000	112, 400, 800, 2400
Convolutional kernel size	3	3
Convolutional dropout rate	0.2	0.15
Number of recurrent units for each layer	256, 128	256, 128
Recurrent dropout rate	0.1	0.1

7.4.3 GAN-aided FDD Evaluation Experiments

The GAN-aided FDD performances were evaluated under the data scenarios described in subsection 7.4.3.1. The structure of the GANs is shown and explained in subsection 7.4.3.2. The candidates for hyperparameter tuning of the GAN-based frameworks are provided in subsection 7.4.3.3. The setup for the GAN-aided FDD evaluation experiments is presented in subsection 7.4.3.4.

7.4.3.1 GAN-aided Data Scenarios

The proposed framework and other GAN-based frameworks were employed to synthesize fault samples based on the real fault data scenarios provided in Table 7.2. As presented in Table 7.4, using the synthetic fault samples, the GANs aim to level up the fault data amount per class for each season to 250 for the GAN-aided scenarios 1–16, the same as the ideal benchmark scenario. Hence, the total fault data amount for all the GAN-aided scenarios and the benchmark scenario is the same.

Table 7.4 The GAN-aided data scenarios by aggregating the real fault samples with the synthetic fault samples (in round bracket) for evaluating GAN-aided FDD performances.

Scenario	Aggregated fault data amount per class				Total fault data amount
	Spring	Summer	Autumn	Winter	
Benchmark	250	250	250	250	14000
1	50 + (200)	50 + (200)	50 + (200)	50 + (200)	14000
2	50 + (200)	50 + (200)	50 + (200)	25 + (225)	14000
3	50 + (200)	50 + (200)	25 + (225)	50 + (200)	14000
4	50 + (200)	25 + (225)	50 + (200)	50 + (200)	14000
5	25 + (225)	50 + (200)	50 + (200)	50 + (200)	14000
6	50 + (200)	50 + (200)	25 + (225)	25 + (225)	14000
7	50 + (200)	25 + (225)	50 + (200)	25 + (225)	14000
8	50 + (200)	25 + (225)	25 + (225)	50 + (200)	14000
9	25 + (225)	25 + (225)	50 + (200)	50 + (200)	14000
10	25 + (225)	50 + (200)	25 + (225)	50 + (200)	14000
11	25 + (225)	50 + (200)	50 + (200)	25 + (225)	14000
12	25 + (225)	25 + (225)	25 + (225)	50 + (200)	14000
13	25 + (225)	25 + (225)	50 + (200)	25 + (225)	14000
14	25 + (225)	50 + (200)	25 + (225)	25 + (225)	14000
15	50 + (200)	25 + (225)	25 + (225)	25 + (225)	14000
16	25 + (225)	25 + (225)	25 + (225)	25 + (225)	14000

7.4.3.2 Generative Adversarial Network Structure

The GAN structure for the proposed ML-CWGAN and other GAN-based architectures is illustrated in Figure 7.4. The generator consists of three blocks of dense, leakyReLU and BN layers, an output dense layer and a reshape layer that reshapes the output from the last dense layer into the same shape of real fault samples. The generator takes the sampled noises from a random distribution and an embedding layer incorporating the class label information as inputs, and produces synthetic fault samples as outputs. The critic (for CWGAN) or discriminator (for CGAN) contains an input dense layer, a leakyReLU layer, two blocks of dense, leakyReLU and dropout layers, and an output dense layer. The first dense layer in the critic/discriminator takes synthetic fault samples, i.e., the output of the generator, real fault samples and their corresponding labels as inputs for the critic/discriminator. The last dense layer outputs a binary value indicating real or fake for the discriminator or a sample realness score for the critic. In addition, it should be noted that the leakyReLU layer was used as the

activation layer in the GAN frameworks, rather than the ReLU layer due to the vanishing-gradient issue of the ReLU layer.

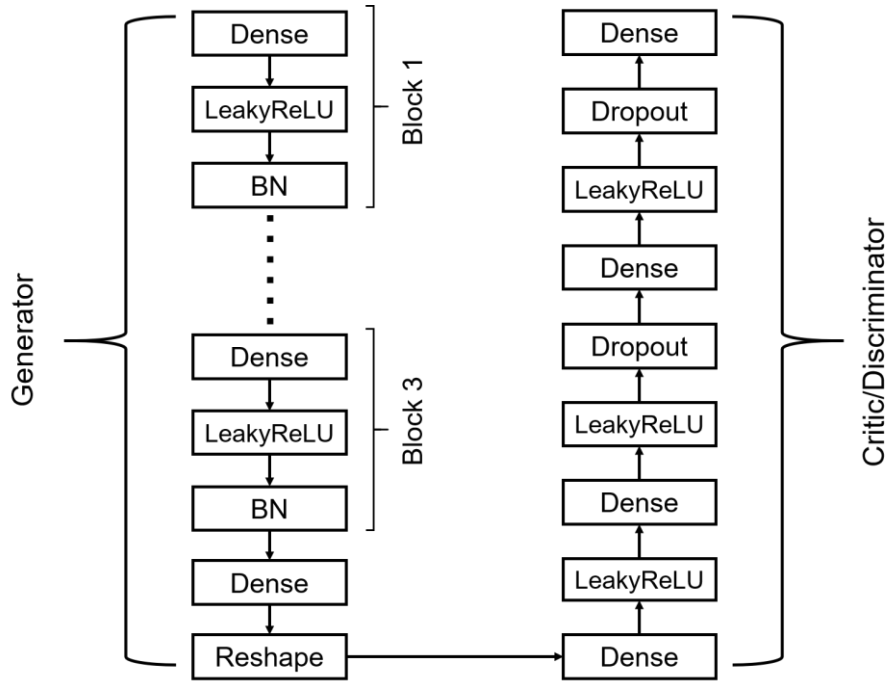


Figure 7.4 The GAN structure for the proposed ML-CWGAN and other GAN-based architectures.

7.4.3.3 Generative Adversarial Network Hyperparameters Tuning

The candidates for hyperparameter tuning of the GAN-based frameworks, i.e., SL-CWGAN, ML-CGAN and ML-CWGAN, are provided in Table 7.5. The number of unit for the first dense layer in the generator was selected from 32, 64, 128, 256 and 512. The number of unit for the second and third dense layer in the generator was all set to double the number of unit from the preceding layer. The number of unit for the first dense layer in the discriminator was set to the same as that of the third dense layer in the generator. The number of unit for the second and third dense layer in the discriminator was all defined as half the number of unit from the preceding layer. As for the CWGANs, the number of critics was determined from 5, 8, 10 and 15 and the gradient clipping constraint was selected from 0.0005, 0.001, 0.005 and 0.01. Apart from these, the dimension of the latent space in the GANs was searched from the range between 30 and 120. Besides, the alpha value in the LeakyReLU activation layer, the momentum of batch normalization and the dropout rate for the dropout layers in the critic/discriminator were all set to the constants of 0.2, 0.8 and 0.4, respectively.

Regarding the training of GANs, the Adam optimizer (Kingma and Ba, 2014) was used for CGANs and the RMSprop optimizer (Ruder, 2017) was applied to CWGANs. The batch size for feeding both synthetic and real fault samples to the critic/discriminator was determined from 128, 256 and 512. The maximum number of training epochs was selected from 1000, 2000, 3000 and 4000 and the learning rate was searched from the range between 0.00001 and 0.0002.

Table 7.5 CGAN and CWGAN architecture hyperparameters.

Hyperparameters	Values
Number of units for the 1st dense layer in generator G_1	32, 64, 128, 256, 512
Number of units for the 2nd dense layer in generator G_2	$2G_1$
Number of units for the 3rd dense layer in generator G_3	$2G_2$
Number of units for the 1st dense layer in discriminator/critic D_1	G_3
Number of units for the 2nd dense layer in discriminator/critic D_2	$\frac{1}{2}D_1$
Number of units for the 3rd dense layer in discriminator/critic D_3	$\frac{1}{2}D_2$
Number of critics (only for CWGAN)	5, 8, 10, 15
Gradient clipping constraint (only for CWGAN)	0.0005, 0.001, 0.005, 0.01
Latent dimension	30 ~ 120
LeakyReLU alpha value	0.2
Batch normalization momentum	0.8
Dropout rate for the critic/discriminator	0.4
Batch sizes	128, 256, 512
Optimizer	Adam for CGAN; RMSprop for CWGAN
Maximum epochs	1000, 2000, 3000, 4000
Learning rate	0.00001–0.0002

7.4.3.4 GAN-aided FDD Evaluation Experiment Setup

The proposed ML-CWGAN framework was used to produce synthetic fault samples under 14 seasonal-imbalanced fault data scenarios and two seasonal-balanced but insufficient fault data scenarios, in order to level up the fault data amount per class for each season to 250, the same as the ideal benchmark scenario's. Then, the FDD classifier with superior baseline FDD performance in section 7.5 was selected for evaluating the quality of synthetic fault samples generated by the proposed framework. The synthetic fault samples were aggregated with the real samples, and the aggregated samples were used for training the FDD classifier. Next, the

trained FDD classifier was tested under each season's full-size dataset including real samples only. Based on the test results, the performance of the ML-CWGAN-aided FDD framework under each data scenario was compared with the corresponding baseline FDD performance, and also compared with those led by other GAN-based architectures, i.e., SL-CWGAN and ML-CGAN, as presented in Table 7.6.

Table 7.6 The candidates for the fault data generation module and the FDD classifier module.

Fault Data Generation	FDD Classifier
ML-CWGAN	CNN-GRU or CNN-LSTM (the classifier of superior baseline FDD performance was selected)
ML-CGAN	
SL-CWGAN	

7.5 Baseline FDD Evaluation Experiments Results

The baseline FDD performances for the CNN-GRU and CNN-LSTM classifiers were evaluated under the data scenarios proposed in Table 7.2. The F1 score (F1 score is a metric suitable for evaluating machine learning model performance under class-imbalanced data distributions) tested under each season and the average F1 score of all the seasons for both classifiers under each data scenario were provided in Table 7.7. Only the benchmark scenario achieved test F1 scores of at least 0.95 per season, and thus 250 samples per season for each fault was considered as the minimum data amount per fault for training a baseline FDD model with an F1 score of at least 0.95. Based on Table 7.7, the variations of each season's F1 score among data scenarios for both classifiers were shown in Figure 7.5. According to the overall variation trends for F1 scores under the four seasons, both classifiers indicated the best FDD performance under winter and the worst performance under autumn. In addition, the F1 score under each season fluctuated along with different data scenarios representing different seasonal intra-class imbalanced ratios. For instance, from scenario 7 to 8, the fault sample amount under autumn decreased from 50 to 25 while the fault sample amount under winter increased from 25 to 50. Therefore, the imbalanced ratios among data samples under each season varied from 2:1:2:1 to 2:1:1:2, resulting in an increase in F1 score under winter and a substantial decrease in F1 score under autumn.

Table 7.7 Test F1 score per season and season-average F1 score for the baseline FDD classifiers under each data scenario.

Scenario	Test F1 Scores									
	Spring		Summer		Autumn		Winter		Average	
	CNN-GRU	CNN-LSTM	CNN-GRU	CNN-LSTM	CNN-GRU	CNN-LSTM	CNN-GRU	CNN-LSTM	CNN-GRU	CNN-LSTM
Benchmark	0.959	0.946	0.972	0.960	0.947	0.937	0.973	0.965	0.963	0.952
1	0.908	0.913	0.924	0.929	0.878	0.860	0.939	0.937	0.912	0.910
2	0.909	0.911	0.931	0.928	0.874	0.880	0.935	0.930	0.912	0.912
3	0.912	0.902	0.930	0.923	0.852	0.817	0.942	0.943	0.909	0.896
4	0.907	0.888	0.901	0.861	0.889	0.841	0.941	0.937	0.909	0.882
5	0.893	0.852	0.933	0.907	0.896	0.868	0.954	0.935	0.919	0.890
6	0.915	0.898	0.916	0.920	0.849	0.828	0.915	0.908	0.899	0.888
7	0.898	0.901	0.896	0.878	0.896	0.840	0.917	0.915	0.902	0.883
8	0.907	0.889	0.881	0.868	0.850	0.776	0.943	0.936	0.895	0.867
9	0.889	0.873	0.901	0.878	0.883	0.840	0.946	0.928	0.905	0.879
10	0.882	0.897	0.929	0.938	0.826	0.859	0.938	0.941	0.894	0.909
11	0.880	0.855	0.916	0.913	0.857	0.861	0.921	0.903	0.893	0.883
12	0.860	0.897	0.889	0.903	0.831	0.856	0.934	0.943	0.878	0.900
13	0.880	0.865	0.889	0.869	0.877	0.815	0.922	0.916	0.892	0.866
14	0.877	0.865	0.927	0.915	0.854	0.809	0.923	0.912	0.895	0.875
15	0.901	0.897	0.872	0.872	0.831	0.780	0.915	0.925	0.880	0.868
16	0.867	0.883	0.879	0.871	0.822	0.803	0.908	0.917	0.869	0.869

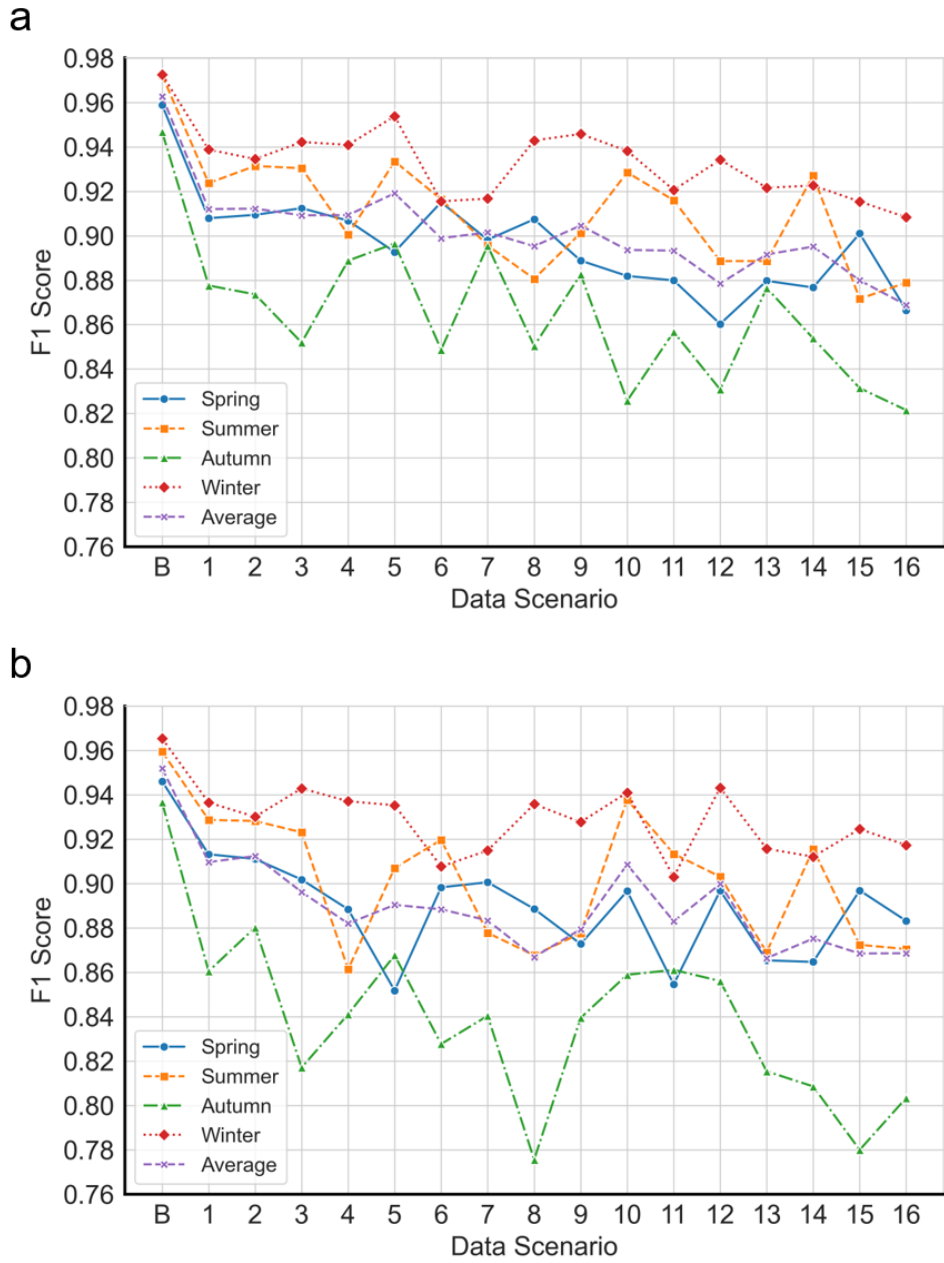


Figure 7.5 The test F1 score per season for (a) CNN-GRU classifier; (b) CNN-LSTM classifier under each data scenario.

The comparison of the average F1 score of all the seasons between both classifiers was shown in Figure 7.6. Under most data scenarios, the average F1 scores for the CNN-GRU classifier were higher than those for the CNN-LSTM classifier, and thus the CNN-GRU classifier was selected as the FDD classifier for evaluating the quality of synthetic fault samples generated by GAN frameworks. The baseline F1 scores for the CNN-GRU classifier under each data scenario were also used as baseline performances for comparing with the GAN-aided FDD performances in section 7.6.

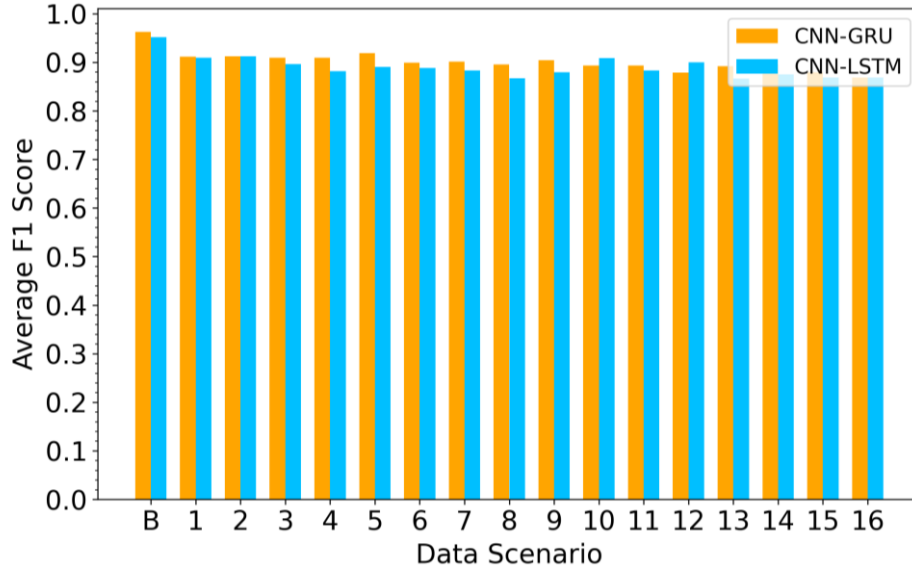


Figure 7.6 Season-average F1 scores for CNN-GRU and CNN-LSTM classifiers under each data scenario.

7.6 GAN-aided FDD Performance Results

The GAN-aided FDD performances under each data scenario for the proposed framework and other comparing GAN-based frameworks were evaluated and presented in this section. The variations in the average F1 score (F1 score is a metric suitable for evaluating machine learning model performance under class-imbalanced data distributions) of all the seasons from baseline performances under each data scenario for each framework are presented in Table 7.8. Section 7.6.1 describes the proposed ML-CWGAN-aided FDD performance variation from the baseline performance of the FDD classifier under each data scenario. Section 7.6.2 compares the FDD performance variation based on the ML-CWGAN with that based on the ML-CGAN. Section 7.6.3 compares the FDD performance variation based on the ML-CWGAN with that based on the SL-CWGAN. Section 7.6.4 quantifies the performance gap between the GAN-aided FDD trained on the aggregated real and synthetic fault samples, and the baseline FDD trained on the real fault samples of the same amount as those used in GAN-aided FDD.

Table 7.8 Comparisons of season-average F1 score difference between each GAN-aided FDD performance and the baseline performance under each data scenario.

Data Scenario	ML-CWGAN	ML-CGAN	Difference between ML-CWGAN and ML-CGAN	SL-CWGAN	Difference between ML-CWGAN and SL-CWGAN
1	0.032	0.033	-0.001	0.018	0.015
2	0.026	0.025	0.001	-0.008	0.033
3	0.017	0.028	-0.011	0.011	0.017
4	0.031	0.024	0.007	0.021	0.003
5	0.014	0.009	0.005	0.016	-0.007
6	0.030	0.019	0.011	0.019	0.000
7	0.030	0.031	-0.001	0.028	0.003
8	0.033	0.033	0.000	0.025	0.008
9	0.025	0.023	0.002	0.021	0.002
10	0.031	0.039	-0.008	0.030	0.009
11	0.043	0.035	0.008	0.031	0.004
12	0.036	0.039	-0.003	0.042	-0.003
13	0.032	0.028	0.004	0.027	0.001
14	0.028	0.032	-0.004	0.018	0.014
15	0.050	0.047	0.003	0.038	0.009
16	0.045	0.043	0.002	0.047	-0.004
Mean	0.032	0.031	0.001	0.024	0.008
Maximum	0.050	0.047	0.003	0.047	0.000
Minimum	0.014	0.009	0.005	-0.008	0.017

7.6.1 ML-CWGAN-aided FDD Performance

The hyperparameter tuning for the ML-CWGAN was conducted and the ML-CWGAN-aided FDD performance for each set of hyperparameters was tested under each data scenario. The optimal configuration of hyperparameters for the ML-CWGAN was provided in Table 7.9. Using the optimal configuration of the ML-CWGAN, the synthetic fault samples were generated for each data scenario, and the aggregated synthetic and real fault samples were used for training the FDD classifier. The trained FDD classifier was then tested under each season's full-size dataset including real samples only and each season's test F1 score variation from the baseline performance under the data scenarios was illustrated in Figure 7.7a. Notably, the overall F1 score trends for all the seasons indicate the highest F1 score increase under autumn for the majority of data scenarios and the lowest F1 score increase under

winter, which is opposite to the baseline absolute F1 scores under those seasons in Figure 7.5a. Although the amount of aggregated real and synthetic samples under each data scenario is same, the F1 score for each season is different among the data scenarios (Fig.7.7b). The higher portion of real fault samples within the total of 250 fault samples for a season under a data scenario leads to higher F1 score under this season. For instance, comparing data scenario 8 with scenario 9, scenario 8 contains 50 real fault samples out of the total of 250 fault samples under spring, while scenario 9 only contains 25 fault samples out of the total under the same season. Consequently, the F1 score under spring decreased by more than 0.02 from scenario 8 to scenario 9.

Table 7.9 ML-CWGAN architecture hyperparameters.

Hyperparameters	Values
Number of unit for the first dense layer in generator G_1	256
Number of unit for the second dense layer in generator G_2	512
Number of unit for the third dense layer in generator G_3	1024
Number of units for the first dense layer in critic D_1	1024
Number of units for the second dense layer in critic D_2	512
Number of units for the third dense layer in critic D_3	256
Number of critics	10
Gradient clipping constraint	0.005
Latent dimension	50
LeakyReLU alpha value	0.2
Batch normalization momentum	0.8
Dropout rate for the critic/discriminator	0.4
Batch sizes	256
Optimizer	RMSprop
Maximum epochs	2000
Learning rate	0.00001

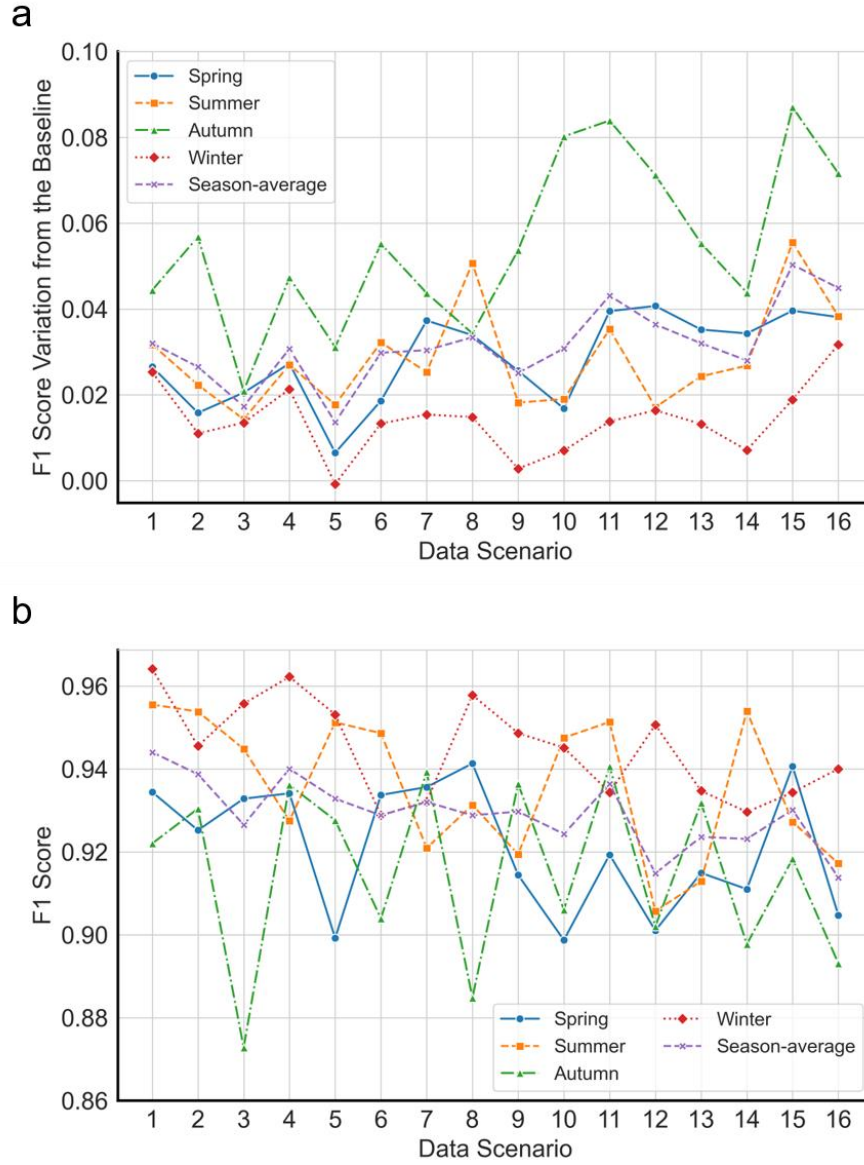


Figure 7.7 (a) The each season's F1 score variation between the ML-CWGAN-aided performance and the baseline performance under each data scenario; (b) The F1 score per season for the ML-CWGAN-aided FDD under each data scenario.

7.6.2 FDD Performance Enhancement: ML-CGAN vs ML-CWGAN

The season-average F1 score variations from baseline performances led by the ML-CGAN and the ML-CWGAN were compared and shown in Figure 7.8. Both ML-CGAN-aided FDD and ML-CWGAN-aided FDD strategies improved F1 scores from baseline FDD performances. According to Table 7.8, the ML-CGAN largely enhanced the FDD performance by up to 0.047 from the baseline and the ML-CWGAN considerably improved the baseline F1 score by up to 0.050. As presented in Figure 7.8, the ML-CWGAN-aided FDD outperformed the ML-CGAN-aided FDD under 10 of the 16 scenarios and the mean of

season-average F1 scores among data scenarios for the ML-CWGAN-aided FDD was 0.001 higher than that of the ML-CGAN-aided FDD. Overall, both GAN architectures can produce similar quality of synthetic fault samples for improving FDD performances and the ML-CWGAN have a better adaptability to a wider range of data scenarios.

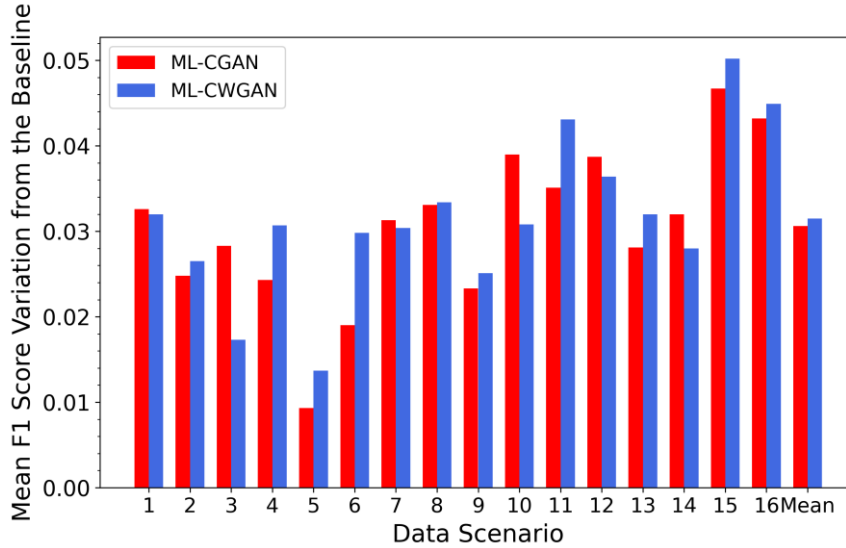


Figure 7.8 Season-average F1 scores for ML-CGAN-aided FDD and ML-CWGAN-aided FDD under each data scenario.

7.6.3 FDD Performance Enhancement: ML-CWGAN vs SL-CWGAN

The season-average F1 score variations from baseline performances led by the ML-CWGAN and the SL-CWGAN were compared and shown in Figure 7.9a. Each season's test F1 score variation from baseline performances under each data scenario for the SL-CWGAN-aided FDD was illustrated in Figure 7.9b. As shown in Figure 7.9a, the ML-CWGAN-aided FDD outperformed the SL-CWGAN-aided FDD under most data scenarios. The mean of season-average F1 scores among data scenarios for the ML-CWGAN-aided FDD was 0.008 higher than that for the SL-CWGAN-aided FDD as presented in Table 7.8. Noteworthily, the performances of SL-CWGAN-aided FDD were unsteady among the data scenarios. In particular, the SL-CWGAN-aided FDD led to a F1 score lower than the baseline F1 score under the scenario 2 (Fig. 7.9a), which means that the feature patterns of synthetic fault samples produced by the SL-CWGAN substantially deviated from those of real fault samples. In addition, this phenomenon can be verified by each season's F1 score variation from the baseline under the scenario 2 (Fig.7.9b). The F1 scores tested under three out of the four seasons were lower than the baseline performances. In general, the ML-CWGAN is more

robust in generating high quality fault samples under fault data seasonal intra-class imbalanced scenarios.

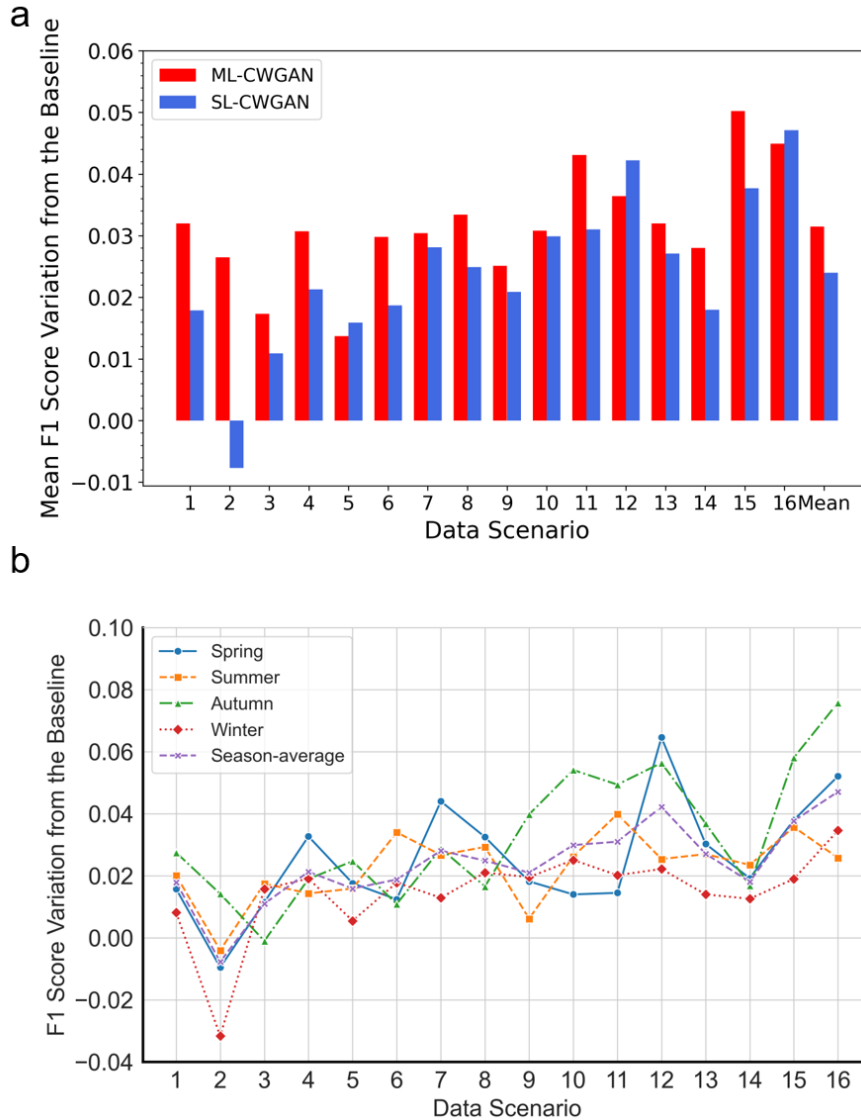


Figure 7.9 (a) Season-average F1 scores for ML-CWGAN -aided FDD and SL-CWGAN-aided FDD under each data scenario; (b) The each season's F1 score variation between the SL-CWGAN-aided performance and the baseline performance under each data scenario.

7.6.4 FDD Performance Gap: Aggregated Real and Synthetic Samples vs Real Samples

The ML-CWGAN-aided FDD performance, trained on the aggregated real and synthetic fault samples, was further compared with the ideal benchmark performance based on real fault samples of the same amount as aggregated fault samples', and the performance gap between them for each data scenario was visualized in Figure 7.10. The largest season-average F1 score gap from the ideal benchmark performance reached 4.9% and the lowest gap reached

1.9%. The highest F1 score gap reached over 0.07 for autumn under scenario 3 and the lowest F1 score gap was around 0.01 under autumn and winter for several scenarios. According to the amount of real fault samples for each season under the data scenarios in Table 7.4, the peak of F1 score gap trend for each season mostly occurred when the amount of real fault samples for this season under this scenario was reduced from 50 to 25, which indicated that the feature patterns of synthetic fault samples produced by the ML-CWGAN still deviated from those of real fault samples. Hence, a high-quality synthetic fault sample selection protocol using machine learning anomaly detection technique(s) needs to be proposed in the future work, in order to control the synthetic fault sample quality fed into the training of FDD classifiers.

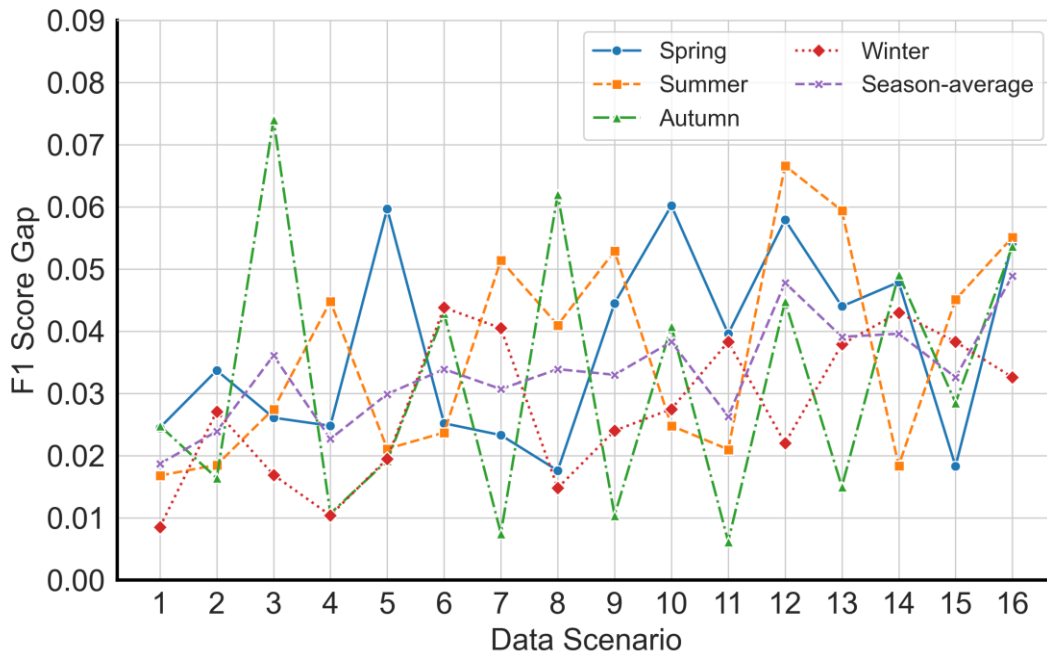


Figure 7.10 The F1 score gap between the ML-CWGAN -aided FDD performance and the ideal benchmark performance based on real fault samples only.

7.7 Key Findings and Potential Future Work

This chapter proposed a ML-CWGAN to support effective FDD performances under fault data seasonal intra-class imbalanced scenarios, in order to tackle the challenge in baseline FDD performance degradation under these scenarios. 16 baseline data scenarios, incorporating 14 fault data seasonal imbalanced scenarios and 2 seasonal balanced but insufficient scenarios, and an ideal benchmark scenario were proposed. The ML-CWGAN architecture contributed to the increase of fault data amount per season to 250, same as the

fault data amount per season under the benchmark scenario, using synthetic fault samples produced by the architecture. Based on the aggregated synthetic and real fault samples under each scenario, ML-CWGAN-aided FDD performances for all the seasons were compared with baseline performances, ML-CGAN-aided FDD performances and SL-CWGAN-aided FDD performances. In addition, the ML-CWGAN-aided FDD performances were also compared with FDD performances under the ideal benchmark scenario, in order to identify the performance gap between the FDD classifier trained on aggregated real and synthetic samples and that trained on real samples of the same amount. The key findings for this chapter are summarized as below:

- (1) Under different data scenarios, the ML-CWGAN-aided FDD considerably improved the baseline F1 score by up to 0.050. The higher portion of real fault samples within the total of 250 fault samples for a season led to higher F1 score under this season.
- (2) Both ML-CGAN and ML-CWGAN architectures can produce similar quality of synthetic fault samples for improving FDD performances and the ML-CWGAN have a better adaptability to a wider range of data scenarios.
- (3) Compared with the SL-CWGAN, the ML-CWGAN is more robust in generating high quality fault samples under fault data seasonal intra-class imbalanced scenarios.
- (4) The feature patterns of synthetic fault samples produced by the ML-CWGAN still deviated from those of real fault samples. The largest season-average F1 score gap from the ideal benchmark performance reached 4.9% and the lowest gap reached 1.9%.
- (5) Only the benchmark scenario achieved test F1 scores of at least 0.95 per season, and thus 250 samples per season for each fault was considered as the minimum data amount per fault for training a baseline FDD model with an F1 score of at least 0.95.

The future work should explore integrating other generative modeling algorithms, such as auxiliary classifier GAN that trains the discriminator to predict not only the data sample is real or fake, but also the probability of the data sample belonging to each fault class, with the baseline FDD classifier. In addition, a high-quality synthetic fault sample selection protocol using machine learning anomaly detection technique(s) needs to be incorporated into the fault data generation module, in order to control the synthetic fault sample quality fed into the training of FDD classifiers. Moreover, a quantitative investigation of the minimum data amount required for training the ML-CWGAN is worth to be performed considering different imbalanced ratios in future.

Chapter 8 Conclusions

The main contributions and key findings of this thesis are summarized in section 8.1. The limitations still existing in the research are provided in section 8.2. The future research directions based on the topic of this thesis are recommended in section 8.3.

8.1 Summary of Main Contributions and Key Findings

This thesis provides responses to the three research questions proposed in section 1.3: (1) the impacts of HVAC single and multiple faults under different climate periods and under different seasons, in terms of system energy usages, and occupant thermal comfort levels, were evaluated through fault simulations using EnergyPlus; (2) the data-driven HVAC FDD performances under different fault data intra-class seasonal imbalanced scenarios were assessed; (3) a ML-CWGAN-aided FDD framework was proposed to enhance HVAC FDD performances under fault data seasonal imbalanced scenarios.

To conduct the fault impact analysis under different climatic conditions, a fault modeling and simulation framework was proposed. A series of typical faults in VAV systems were modelled using Python scripts and a widely-used building performance simulation tool, EnergyPlus. The fault impact simulations were performed by introducing developed fault models into a verified prototype office building model. Based on fault simulations, the fault impacts under different seasons and climate periods, in terms of system energy consumptions and occupant thermal comfort, were assessed. The results prove that the fault impacts are affected by climatic condition variations and the seasonal patterns are presented in fault impacts.

Based on the insights gathered from fault impact analyses, the FDD performances for a combined convolutional and recurrent neural network classifier were assessed under fault data seasonal insufficient and imbalanced scenarios and an ideal fault data scenario covering climatic conditions from all the seasons. The developed FDD model was trained under four simulated data scenarios: (1) one-season dataset only containing fault data from a season in each dataset; (2) dual-seasons dataset which has a half of fault data from each of the two seasons; (3) full-seasons dataset which has a quarter of fault data from each of the four seasons; (4) full-seasons dataset which has the whole size of fault data from each season. Each trained model was then tested using all the fault data from each season, in order to

assess the impacts of fault data seasonal insufficiency and imbalance on FDD performances. According to the evaluation results, the FDD classifiers trained on sufficient fault data under a certain season(s) indicated FDD performance degradations tested under unseen season(s). Moreover, two sets of further experiments based on each real fault dataset were implemented to validate this practical issue.

To deal with the challenge in FDD performance degradation led by the fault data intra-class imbalance among different seasons, a ML-CWGAN architecture was integrated with the baseline FDD classifier to support effective FDD performances under fault data seasonal imbalanced scenarios. 16 baseline data scenarios, including 14 fault data seasonal imbalanced scenarios and 2 seasonal balanced but insufficient scenarios, and an ideal benchmark scenario were proposed. The ML-CWGAN architecture was used to increase fault data amount per season to 250, the fault data amount per season under the benchmark scenario, using synthetic fault samples generated by the architecture. Based on the aggregated synthetic and real fault samples under each scenario, ML-CWGAN-aided FDD performances for all the seasons were compared with the baseline performances, ML-CGAN-aided FDD performances and SL-CWGAN-aided FDD performances. In addition, the ML-CWGAN-aided FDD performances were also compared with FDD performances under the ideal benchmark scenario, in order to identify the performance gap between the FDD classifier trained on aggregated real and synthetic samples and that trained on real samples of the same amount.

The key findings from this thesis are summarized as follows:

- (1) Changing climate conditions can have a significant influence on single fault impacts and multiple faults' interactions including synergetic and antagonistic effects.
- (2) HVAC fault impact seasonal patterns were exhibited by a large number of faults and most faults kept seasonal patterns along with changing climates.
- (3) HVAC FDD classifiers trained on sufficient fault data under a certain season(s) show poor generalization ability to detect and identify the faults under other unseen seasons.
- (4) The HVAC FDD performance under a season can be enhanced by increasing the amount of fault data under this season, but the enhancement is limited by other factors, such as the number of faults and the imbalance ratio of fault to normal data included in the training set.
- (5) The coverage of fault data under different seasons is more crucial in enhancing HVAC FDD performances than the amount of fault data under a certain season(s).

- (6) Under different data scenarios, the ML-CWGAN-aided FDD considerably improved the baseline F1 score by up to 0.050. The higher portion of real fault samples within the total of 250 fault samples for a season led to higher F1 score under this season.
- (7) Both ML-CGAN and ML-CWGAN architectures can produce similar quality of synthetic fault samples for improving the FDD performance and the ML-CWGAN have a better adaptability to a wider range of data scenarios.
- (8) Compared with the SL-CWGAN, the ML-CWGAN is more robust in generating high quality fault samples under fault data seasonal imbalanced scenarios.
- (9) The feature patterns of synthetic fault samples produced by ML-CWGAN still deviated from those of real fault samples. The largest season-average F1 score gap from the ideal benchmark performance reached 4.9% and the lowest gap reached 1.9%.
- (10) Only the benchmark scenario achieved test F1 scores of at least 0.95 per season, and thus 250 samples per season for each fault was considered as the minimum data amount per fault for training a baseline FDD model with an F1 score of at least 0.95.

8.2 Research Limitations

As HVAC fault impact analyses under changing climate conditions are novel to the HVAC fault impact analysis field and HVAC FDD performance enhancement using data generative modeling techniques is still a growing topic in the data-driven HVAC FDD field, the thesis represents an exploration on this topic. Hence, there are still a number of limitations in this study:

- (1) The HVAC fault modeling approach used in this thesis only took into account constant fault intensities for developing fault models, while the intensities of some faults may evolve along with the time in reality. Therefore, the fault impacts evaluated in this work can only represent the situation of static faults and a more complex situation of evolving faults has not been considered.
- (2) The HVAC fault impact analyses were performed based on an ideal assumption that the faults were undiscovered for a year for the purpose of easily comparing the impacts of different faults under same climatic conditions. Due to the lack of fault occurrence frequency data during the period of completing this work, fault impact evaluations considering the occurrence frequency of each fault were unable to implement.
- (3) A VAV system with an RTU was used as the case study for the thesis and the findings from fault impact analyses have not been verified in other types of HVAC systems.

- (4) The HVAC fault impact analyses in the thesis were conducted under the marine climate only and fault impact evaluations under other climate zones have not been performed.
- (5) This study only evaluated the effect of the fault data intra-class seasonal imbalance on HVAC FDD performances, which can be expanded to a broader topic, evaluating the influence of the fault data intra-class imbalance under different system working conditions on FDD performances.
- (6) Training a robust GAN architecture for generating reliable data samples relies on domain experience and sophisticated tuning techniques. The feature patterns of synthetic fault samples produced by the proposed ML-CWGAN in this thesis still deviated from those of real fault samples. Therefore, the proposed ML-CWGAN still requires further optimizations or the incorporation of a high-quality synthetic fault sample selection protocol.
- (7) This work only considered the fault detection and diagnosis task for HVAC single faults due to limited real fault datasets containing multiple simultaneous faults.

8.3 Future Work

While exploring the research questions of the thesis, a series of research directions, in terms of HVAC fault impact analysis and data-driven HVAC FDD, are raised and worth investigating in future:

- (1) The developed HVAC fault models only consider a situation of static faults with constant fault intensities. A more complex situation, faults with evolving intensities, needs to be modelled in future, and corresponding fault experiments or real fault measurements are required to validate these fault models.
- (2) With the increasingly available HVAC fault occurrence frequency data, future work can take into account this characteristic to quantify fault impacts in a more realistic way.
- (3) The fault impacts from other types of HVAC systems, such as fan coil systems, are worth exploring.
- (4) Future work can explore HVAC fault impacts under different climate zones, such as hot climates, to compare the differences in fault impacts under different climate zones.
- (5) HVAC fault data under different system working conditions, such as chiller fault data under different working conditions, can be divided into specific groups. Hence, evaluating the impact of the fault data intra-class imbalance among different working conditions on HVAC FDD performances can be considered as a future direction. The clustering algorithms can be utilized to identify different groups of system working conditions, apart from depending on experts' domain knowledge.

- (6) Future studies are recommended to integrate a high-quality synthetic fault sample selection protocol, such as using a variational auto-encoder to select high-quality fault samples generated by GANs.
- (7) Other variants of GANs, such as auxiliary classifier GAN that trains the discriminator to predict not only the data sample is real or fake, but also the probability of the data sample belonging to each fault class, can be attempted to generate synthetic fault samples in future.
- (8) Real HVAC fault datasets incorporating multiple simultaneous faults need to be constructed in future, in order to evaluate and test frameworks possessing multiple-simultaneous faults detection and diagnosis capability.
- (9) A quantitative investigation of the minimum data amount required for training the ML-CWGAN is worth to be performed considering different imbalanced ratios in future.

References

- Albayati, M.G., Faraj, J., Thompson, A., Patil, P., Gorthala, R. and Rajasekaran, S. (2022) ‘Semi-Supervised Machine Learning for Fault Detection and Diagnosis of a Rooftop Unit’, *Big Data Mining and Analytics*, 6(2), pp. 170 – 184.
- Allen, W.H., Rubaai A. and Chawla R. (2016) ‘Fuzzy Neural Network-Based Health Monitoring for HVAC System Variable-Air-Volume Unit’, *IEEE Transactions on Industry Applications*, 52, no. 3.
- American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE). (2014) *ASHRAE Guideline 14-2014, Measurement of Energy, Demand, and Water Savings*, Atlanta, United States: ASHRAE.
- American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE). (2017) *ASHRAE Weather Data 2017*, Atlanta, United States: ASHRAE.
- Arjovsky, M., Chintala, S. and Bottou, L. (2017) ‘Wasserstein GAN’, arXiv:1701.07875.
- Basarkar, M., Pang, X., Wang, L., Haves, P. and Hong, T. (2011) *Modeling and simulation of HVAC faults in EnergyPlus*. Berkeley: Lawrence Berkeley National Laboratory. Available at: <https://escholarship.org/uc/item/9ps43482> (Accessed: 20 June 2020).
- Bell, N.O., Bilbao, J.I., Kay, M. and Sproul, A.B. (2022) ‘Future climate scenarios and their impact on heating, ventilation and air-conditioning system design and performance for commercial buildings for 2050’, *Renewable and Sustainable Energy Reviews*, 162, no. 112363.
- Bellanco, I., Belío, F., Vallés, M., Gerber, R. and Salom, J. (2022) ‘Common fault effects on a natural refrigerant, variable-speed heat pump’, *International Journal of Refrigeration*, 133, pp. 259–266.

- Björsell, N., Bring, A., Eriksson, L., Grozman, P., Lindgren, M., Sahlin, P., Shapovalov, A. and Vuolle, M. (1999) 'IDA indoor climate and energy', *International Building Simulation Conference 1999*. Kyoto, Japan, 13-15 September.
- Bonvini, M., Sohn, M., Granderson, J., Wetter, M. and Piette, M. (2014) 'Robust on-line fault detection diagnosis for HVAC components based on nonlinear state estimation techniques', *Applied Energy*, 124, pp. 156-166.
- Bourguignon, M., Silva, R.B. and Cordeiro, G.M. (2014) 'The Weibull-G family of probability distributions', *Journal of Data Science*, 12(1), pp. 53-68.
- Buckberry, H.L. and Bhandari, M. (2012) 'ORNL MAXLAB Flexible Research Platforms', *Proceedings of 2012 ACEEE Summer Study on Energy Efficiency in Buildings*, Pacific Grove, United States, 12-17 August.
- Bushby, S.T., Castro, N., Galler, M.A. and Park, C. (2001) *Using the Virtual Cybernetic Building Testbed and FDD Test Shell for FDD Tool Development*. Gaithersburg: National Institute of Standards and Technology. Available at: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.106.2695> (Accessed: 6 August 2020).
- Breuker, M.S. and Braun, J.E. (1998) 'Common faults and their impacts for rooftop air conditioners', *HVAC&R Research*, 4(3), pp. 303-318.
- Capozzoli, A., Lauro, F. and Khan, I. (2015) 'Fault detection analysis using data mining techniques for a cluster of smart office buildings', *Expert Systems with Applications*, 42, pp. 4324-4338.
- Canizo, M., Triguero, I., Conde, A. and Onieva, E. (2019) 'Multi-head CNN-RNN for multi-time series anomaly detection: An industrial case study', *Neurocomputing*, 363, pp. 246-260.
- Catrini, P. and Piacentino, A. (2023) 'Experimental performance characterization of variable-speed packaged rooftop units with fouled evaporator', *Applied Thermal Engineering*, 233, no. 121159.

Chen, Y., Huang, S. and Vrabie, D. (2018) ‘A simulation based approach for impact assessment of physical faults: Large commercial building HVAC case study’, *Proceedings of 2018 Building Performance Modeling Conference and SimBuild*, Chicago, United States, 26-28 September. Available at: <https://www.ashrae.org/File%20Library/Conferences/Specialty%20Conferences/2018%20Building%20Performance%20Analysis%20Conference%20and%20SimBuild/Papers/C113.pdf> (Accessed: 5 October 2021).

Chen, J., Zhang, L., Li, Y., Shi, Y., Gao, X. and Hu, Y. (2022) ‘A review of computing-based automated fault detection and diagnosis of heating, ventilation and air conditioning systems’, *Renewable and Sustainable Energy Reviews*, 161, no. 112395.

Chen, Z., O’Neill, Z., Wen, J., Pradhan, O., Yang, T., Lu, X., Lin, G., Miyata, S., Lee, S., Shen, C., Chiosa, R., Piscitelli, M.S., Capozzoli, A., Hengel, F., Kühner, A., Pritoni, M., Liu, W., Clauß, J., Chen, Y. and Herr, T. (2023) ‘A review of data-driven fault detection and diagnostics for building HVAC systems’, *Applied Energy*, 339, no. 121030.

Chen, Z., Xiao, F. and Guo, F. (2023) ‘Similarity learning-based fault detection and diagnosis in building HVAC systems with limited labeled data’, *Renewable and Sustainable Energy Reviews*, 185, no. 113612.

Cheng, H., Chen, H., Li, Z. and Cheng, X. (2020) ‘Ensemble 1-D CNN diagnosis model for VRF system refrigerant charge faults under heating condition’, *Energy and Buildings*, 224, no. 110256.

Cheng, F., Cai, W., Zhang, X., Liao, H. and Cui C. (2021) ‘Fault detection and diagnosis for Air Handling Unit based on multiscale convolutional neural networks’, *Energy and Buildings*, 236, no. 110795.

Cheung, H. and Braun, J.E. (2013a) ‘Simulation of fault impacts for vapor compression systems by inverse modeling. Part I: Component modeling and validation’, *HVAC&R Research*, 19(7), pp. 892-906.

Cheung, H. and Braun, J.E. (2013b) ‘Simulation of fault impacts for vapor compression systems by inverse modeling. Part II: System modeling and validation’, *HVAC&R Research*, 19(7), pp. 907-921.

Cheung, H. and Braun, J. E. (2015) *Development of fault models for hybrid fault detection and diagnostics algorithm*. Golden, United States: National Renewable Energy Laboratory. doi:10.2172/1235409.

Cheung, H. and Braun, J.E. (2016) ‘Empirical modeling of the impacts of faults on water-cooled chiller power consumption for use in building simulation programs’, *Applied Thermal Engineering*, 99, pp. 756-764.

Cho, K., van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H. and Bengio, Y. (2014) ‘Learning Phrase Representations using RNN Encoder–Decoder for Statistical Machine Translation’, arXiv:1406.1078v3.

Choi, E., Biswal, S., Malin, B., Duke, J., Stewart, W.F. and Sun, J. (2017) ‘Generating multi-label discrete patient records using generative adversarial networks’, *Proceedings of Machine Learning Research*, 68, pp. 286-305.

Christidis, N., Stott, P.A., Brown, S., Karoly, D.J. and Caesar, J. (2007) ‘Human contribution to the lengthening of the growing season during 1950–99’, *Journal of Climate*, 20, pp. 5441–5454.

Cohen, J. (1960) ‘A Coefficient of Agreement for Nominal Scales’, *Educational and Psychological Measurement*, 20, pp. 37-46.

Cordeiro, G.M., Ortega, E.M. and Ramires, T.G. (2015) ‘A new generalized Weibull family of distributions: mathematical properties and applications’, *Journal of Statistical Distributions and Applications*, 2, no. 13.

- Crowe, E., Chen, Y., Granderson, J., Reeve, H., Troup, L., Yuill, D. and Chen, Y. (2022) *What We Learned From Analyzing 18 Million Rows of Commercial Buildings' HVAC Fault Data*. Berkeley, United States: Lawrence Berkeley National Laboratory. Available at: https://eta-publications.lbl.gov/sites/default/files/what_we_learned_from_analyzing_18_0.pdf (Accessed: 7 June 2023).
- Deng, Q., Chen, Z., Tang, P., Li, X., Yang, C. and Yang, X. (2023) 'A novel semi-supervised fault diagnosis method for chillers based on neighbor-optimized graph convolutional network', *Energy and Buildings*, 301, no. 113703.
- Deru, M., Field, K., Studer, D., Benne, K., Griffith, B., Torcellini, P., Liu, B., Halverson, M., Winiarski, D., Rosenberg, M., Yazdanian, M., Huang, J. and Crawley, D. (2011) *U.S. Department of Energy commercial reference building models of the national building stock*. Golden, United States: National Renewable Energy Laboratory. doi:10.2172/1009264
- Dey, D., Dong, B. and Li, Z. (2016) 'A probabilistic framework to diagnose faults in air handling units', *Proceedings of 4th International High Performance Buildings Conference*, West Lafayette, United States, 11-14 July. Available at: <https://docs.lib.purdue.edu/ihpbc/168/> (Accessed: 30 July 2020).
- Downey, T. and Proctor, J. (2002) 'What Can 13,000 Air Conditioners Tell Us?', 2002 *American Council for an Energy-Efficient Economy (ACEEE) summer study*. Pacific Grove, United States, 18-23 August.
- Du, Z. and Jin, X. (2007a) 'Detection and diagnosis for sensor fault in HVAC systems', *Energy Conversion and Management*, 48, pp. 693–702.
- Du, Z. and Jin, X. (2007b) 'Tolerant control for multiple faults of sensors in VAV systems', *Energy Conversion and Management*, 48, pp. 764–777.
- Du, Z. and Jin, X. (2007c) 'Detection and diagnosis for multiple faults in VAV systems', *Energy and Buildings*, 39, pp. 923–934.

- Du, Z., Jin, X. and Wu, L. (2007) 'PCA-FDA-Based Fault Diagnosis for Sensors in VAV Systems', *HVAC&R Research*, 13, no. 2.
- Du, Z. and Jin, X. (2008) 'Multiple faults diagnosis for sensors in air handling unit using Fisher discriminant analysis', *Energy Conversion and Management*, 49, pp. 3654–3665.
- Du, Z., Jin, X. and Yang, Y. (2009) 'Fault diagnosis for temperature, flow rate and pressure sensors in VAV systems using wavelet neural network', *Applied Energy*, 86(9), pp. 1624–1631.
- Du, Z., Fan, B., Jin, X. and Chi, J. (2014a) 'Fault detection and diagnosis for buildings and HVAC systems using combined neural networks and subtractive clustering analysis', *Building and Environment*, 73, pp. 1–11.
- Du, Z., Fan, B., Chi, J. and Jin, X. (2014b) 'Sensor fault detection and its efficiency analysis in air handling unit using the combined neural networks', *Energy and Buildings*, 72, pp. 157–166.
- Du, Z., Domanski, P.A. and Payne, W.V. (2016) 'Effect of common faults on the performance of different types of vapor compression systems', *Applied Thermal Engineering*, 98, pp. 61–72.
- Eames, M., Kershaw, T. and Coley, D. (2011) 'On the creation of future probabilistic design weather years from UKCP09', *Building Services Engineering Research & Technology*, 32(2), pp. 127–142.
- Ebrahimifakhar, A., Kabirikopaei, A. and Yuill, D. (2020) 'Data-driven fault detection and diagnosis for packaged rooftop units using statistical machine learning classification methods', *Energy and Buildings*, 225, no. 110318.

Elnour, M. and Meskin, N. (2020) ‘Actuator Fault Diagnosis in Multi-Zone HVAC Systems using 2D Convolutional Neural Networks’, *Proceedings of 2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT)*, Doha, Qatar, 2-5 February. Available at: <https://doi.org/10.1109/ICIoT48696.2020.9089508> (Accessed: 30 July 2020).

Elnour, M., Meskin, N. and Al-Naemi, M. (2020) ‘Sensor data validation and fault diagnosis using Auto-Associative Neural Network for HVAC systems’, *Journal of Building Engineering*, 27, no. 100935.

EnergyPlus. (2020) Engineering Reference. EnergyPlus 9.3.

Energy Information Administration (EIA). (2016) *Commercial building energy consumption survey (CBECS) 2012*. Washington, D.C., United States: EIA. Available at: <https://www.eia.gov/consumption/commercial/data/2012/> (Accessed: 27 September 2020).

Energy Information Administration (EIA). (2018) *Residential energy consumption survey (RECS) 2015*. Washington, D.C., United States: EIA. Available at: <https://www.eia.gov/consumption/residential/data/2015/> (Accessed: 27 September 2020).

Eom, Y.H., Yoo, J.W., Hong, S.B. and Kim, M.S. (2019) ‘Refrigerant charge fault detection method of air source heat pump system using convolutional neural network for energy saving’, *Energy*, 187, no. 115877.

Fan, Y., Cui, X., Han, H. and Lu, H. (2019) ‘Chiller fault diagnosis with field sensors using the technology of imbalanced data’, *Applied Thermal Engineering*, 159, no. 113933.

Fan, C., Li, X., Zhao, Y. and Wang, J. (2021a) ‘Quantitative assessments on advanced data synthesis strategies for enhancing imbalanced AHU fault diagnosis performance’, *Energy and Buildings*, 252, no. 111423.

Fan, C., Liu, X., Xue, P. and Wang, J. (2021b) ‘Statistical characterization of semi-supervised neural networks for fault detection and diagnosis of air handling units’, *Energy and Buildings*, 234, no. 110733.

Fan, C., Liu, Y., Liu, X., Sun, Y. and Wang, J. (2021c) ‘A study on semi-supervised learning in enhancing performance of AHU unseen fault detection with limited labeled data’, *Sustainable Cities and Society*, 70, no. 102874.

Fan, C., He, W., Liu, Y., Xue, P. and Zhao, Y. (2022) ‘A Tun: From multi-source data integration to knowledge sharing strategies’, *Energy and Buildings*, 262, no. 111995.

Fan, C., Lei, Y., Sun, Y. and Mo, L. (2023a) ‘Novel transformer-based self-supervised learning methods for improved HVAC fault diagnosis performance with limited labeled data’, *Energy*, 278, no. 127972.

Fan, C., Lin, Y., Piscitelli, M.S., Chiosa, R., Wang, H., Capozzoli, A. and Ma, Y. (2023b) ‘Leveraging graph convolutional networks for semi-supervised fault diagnosis of HVAC systems in data-scarce contexts’, *Building Simulation*, 16, pp. 1499–1517.

Fanger, P.O. (1970) *Thermal Comfort: Analysis and Applications in Environmental Engineering*. Copenhagen, Denmark: Danish Technical Press.

Frank, S., Heaney, M., Jin, X., Robertson, J., Cheung, H., Elmore, R. and Henze, G. (2016) ‘Hybrid model-based and data- driven fault detection and diagnostics for commercial buildings’, Proceedings of the 2016 ACEEE Summer Study on Energy Efficiency in Buildings, Pacific Grove, United States, 21-26 August. Available at: https://www.aceee.org/files/proceedings/2016/data/papers/12_437.pdf (Accessed: 24 November 2019).

Frank, S., Lin, G., Jin, X., Singla, R., Farthing, A. and Granderson, J. (2019) ‘A performance evaluation framework for building fault detection and diagnosis algorithms’, *Energy and Buildings*, 192, pp. 84–92.

Gal, Y. and Ghahramani, Z. (2016) 'A Theoretically Grounded Application of Dropout in Recurrent Neural Networks', *NIPS'16: Proceedings of the 30th International Conference on Neural Information Processing Systems*. Barcelona, Spain, 5-10 December. pp. 1027–1035.

Gao, D., Wang, S., Gang, W. and Xiao, F. (2016) 'A model-based adaptive method for evaluating the energy impact of low delta-T syndrome in complex HVAC systems using support vector regression', *Building Services Engineering Research and Technology*, 37(5), pp. 573–596.

Gao, J., Han, H., Ren, Z. and Fan, Y. (2021) 'Fault diagnosis for building chillers based on data self-production and deep convolutional neural network', *Journal of Building Engineering*, 34, no. 102043.

Ginestet, S., Marchio, D. and Morisot, O. (2008) 'Evaluation of faults impacts on energy consumption and indoor air quality on an air handling unit', *Energy and Buildings*, 40(1), pp. 51-57.

Goodfellow, I.J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y. (2014) 'Generative Adversarial Nets', arXiv:1406.2661.

Goodfellow, I., Bengio, Y. and Courville, A. (2016) *Deep Learning*. Cambridge, Massachusetts, United States: the MIT Press.

Gou, W., Ren, Z., Chen, H., Xing, L., Zhou, Z., Xia, X. and Shi, J. (2023) 'Experimental research on the performance and parameters sensitivity analysis of variable refrigerant flow system with common faults imposed in heating mode', *Energy and Buildings*, 278, no. 112624.

Gunay, B., Shen, W., Huchuk, B., Yang, C., Bucking, S. and O'Brien, W. (2018) 'Energy and comfort performance benefits of early detection of building sensor and actuator faults', *Building Services Engineering Research Technology*, 39(6), pp. 652–666.

Gunay, B., Hobson, B.W., Darwazeh, D. and Bursill, J. (2023) ‘Estimating energy savings from HVAC controls fault correction through inverse greybox model-based virtual metering’, *Energy and Buildings*, 282, no. 112806.

Guo, Y., Li, G., Chen, H., Wang, J., Guo, M., Sun, S. and Hu, W. (2017) ‘Optimized neural network-based fault diagnosis strategy for VRF system in heating mode using data mining’, *Applied Thermal Engineering*, 125, pp. 1402–1413.

Guo, Y., Li, G., Chen, H., Hu, Y., Li, H., Xing, L. and Hu, W. (2017) ‘An enhanced PCA method with Savitzky-Golay method for VRF system sensor fault detection and diagnosis’, *Energy and Buildings*, 142, pp. 167–178.

Guo, Y., Li, G., Chen, H., Hu, Y., Li, H., Liu, J., Hu, M. and Hu, W. (2017) ‘Modularized PCA method combined with expert-based multivariate decoupling for FDD in VRF systems including indoor unit faults’, *Applied Thermal Engineering*, 115, pp. 744–755.

Guo, Y., Li, Y. and Li, W. (2023) ‘On-site fault experiment and diagnosis research of the carbon dioxide transcritical heat pump system for energy saving’, *Energy*, 274, no. 127405.

Han, H., Gu, B., Wang, T. and Li, Z.R. (2011a) ‘Important sensors for chiller fault detection and diagnosis (FDD) from the perspective of feature selection and machine learning’, *International Journal of Refrigeration*, 34(2), pp. 586-599.

Han, H., Gu, B., Hong, Y. and Kang, J. (2011b) ‘Automated FDD of multiple-simultaneous faults (MSF) and the application to building chillers’, *Energy and Buildings*, 43(9), pp. 2524-2532.

Han, L., Deng, Y., Chen, H., Wei, G., Sheng, K. and Shi, J. (2022a) ‘A robust VRF fault diagnosis method based on ensemble BiLSTM with attention mechanism: Considering uncertainties and generalization’, *Energy and Buildings*, 269, no. 112243.

Han, S., Shao, H., Huo, Z., Yang, X. and Cheng, J. (2022b) ‘End-to-end chiller fault diagnosis using fused attention mechanism and dynamic cross-entropy under imbalanced datasets’, *Building and Environment*, 212, no. 108821.

Hassan, A.S., Hemeda, S.E., Maiti, S.S. and Pramanik, S. (2017) ‘The generalized additive Weibull-G family of distributions’, *International Journal of Statistics and Probability*, 6(5), pp. 65-83.

Haves, P. (1997) ‘Fault modelling in component-based HVAC simulation’, *Proceedings of the International IBPSA Building Simulation Conference 1997*, Prague, Czech Republic, 8-10 September. Available at: http://www.ibpsa.org/proceedings/BS1997/BS97_P101.pdf (Accessed: 22 July 2020).

Heinemeier, K. and Cherniack, M. (2012) *Rooftop HVAC Fault Detection and Diagnostics: Capabilities of New Technologies for Energy Reduction*. Portland: New Buildings Institute. Available at: https://newbuildings.org/sites/default/files/RooftopHVAC_FaultDetectionDiagnosisMarket.pdf (Accessed: 30 July 2020).

Hochreiter, S. and Schmidhuber, J. (1997) ‘Long Short-Term Memory’, *Neural Computing*, 9, pp. 1735–1780.

Hu, Y., Yuill, D.P., Ebrahimifakhar, A. and Rooholghodos, A. (2021a) ‘An experimental study of the behavior of a high efficiency residential heat pump in cooling mode with common installation faults imposed’, *Applied Thermal Engineering*, 184, no. 116116.

Hu, Y., Yuill, D.P., Rooholghodos, S.A., Ebrahimifakhar, A. and Chen, Y. (2021b) ‘Impacts of simultaneous operating faults on cooling performance of a high efficiency residential heat pump’, *Energy and Buildings*, 242, no. 110975.

Hu, Y. and Yuill, D.P. (2022a) ‘Impacts of common faults on an air conditioner with a microtube condenser and analysis of fault characteristic features’, *Energy and Buildings*, 254, no. 111630.

Hu, Y. and Yuill, D.P. (2022b) ‘Multiple simultaneous faults’ impacts on air-conditioner behavior and performance of a charge diagnostic method’, *Applied Thermal Engineering*, 215, no. 119015.

Hu, Y., Chen, Y. and Yuill, D. P. (2023) ‘Experimental study of liquid line temperature drop as a fault feature for detecting severe liquid line restrictions in residential air-conditioning systems’, *International Journal of Refrigeration*, 152, pp. 293–302.

Huang, S., Zuo, W., Vrabie, D. and Xu, R. (2021) ‘Modelica-based system modeling for studying control-related faults in chiller plants and boiler plants serving large office buildings’, *Journal of Building Engineering*, 44, no. 102654.

Huang, J., Wen, J., Yoon, H., Pradhan, O., Wu, T., O’Neill, Z., Candan, K.S. (2022) ‘Real vs. simulated: Questions on the capability of simulated datasets on building fault detection for energy efficiency from a data-driven perspective’, *Energy and Buildings*, 259, no. 111872.

Huang, S., Ye, Y., Wu, D. and Zuo, W. (2021) ‘An assessment of power flexibility from commercial building cooling systems in the United States’, *Energy*, 221, no. 119571.

International Energy Agency (IEA). (2022) *World Energy Outlook 2022*. Paris, France: IEA. Available at: <https://www.iea.org/reports/world-energy-outlook-2022> (Accessed: 2 June 2023).

Integrated Environmental Solutions (IES). (2016) *ApacheHVAC user guide*. Glasgow, United Kingdom: IES.

Intergovernmental Panel on Climate Change (IPCC). (2021) ‘Summary for Policymakers’, in Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S.L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M.I., Huang, M., Leitzell, K., Lonnoy, E., Matthews, J.B.R., Maycock, T.K., Waterfield, T., Yelekçi, O., Yu, R., and Zhou, B. (eds.) *Climate Change 2021: The Physical Science Basis*. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press, pp. SPM-5. doi:10.1017/9781009157896.001.

Intergovernmental Panel on Climate Change (IPCC). (2013) ‘Summary for Policymakers’, in Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.) *Climate Change 2013: The Physical Science Basis*. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press, pp. 19. Available at: https://www.ipcc.ch/site/assets/uploads/2018/02/WG1AR5_SPM_FINAL.pdf (Accessed: 6 April 2021).

Intergovernmental Panel on Climate Change (IPCC). (2007) ‘Summary for Policymakers’, in Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller (eds.) *Climate Change 2007: The Physical Science Basis*. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press, pp. 18. Available at: <https://www.ipcc.ch/site/assets/uploads/2018/02/ar4-wg1-spm-1.pdf> (Accessed: 7 April 2021).

Intergovernmental Panel on Climate Change (IPCC). (2000) ‘Summary of Policymakers’, *IPCC Special Report Emissions Scenarios*. Geneva, Switzerland: IPCC. Available at: <https://www.ipcc.ch/site/assets/uploads/2018/03/sres-en.pdf> (Accessed: 7 April 2021).

Ioffe, S. and Szegedy, C. (2015) ‘Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift’, *ICML'15: the 32nd International Conference on Machine Learning*. Lille, France, July 7-9. pp. 448-456.

Ji, J., Sun, Y., Gao, D. and Cui, Z. (2023) ‘Systematic analysis method of sensor and actuator faults propagation and impacts for optimally controlled complex building chilled water systems’, *Journal of Building Engineering*, 76, no. 107272.

Jiang, K., Gao, X., Gao, H., Han, H. and Qi, Y. (2023) ‘VIT-GADG: A Generative Domain-Generalized Framework for Chillers Fault Diagnosis Under Unseen Working Conditions’, *IEEE Transactions on Instrumentation and Measurement*, 72, no. 3527413.

- Jones, P.D., Kilsby, C.G., Harpham C., Glenis V. and Burton A. (2010) *UK Climate Projections science report: Projections of future daily climate for the UK from the weather generator*, Newcastle, United Kingdom: University of Newcastle. Available at: <https://core.ac.uk/download/pdf/299302528.pdf> (Accessed: 25 July 2020)
- Johnson, J.M. and Khoshgoftaar, T.M. (2019) ‘Survey on deep learning with class imbalance’, *Journal of Big Data*, 6, no. 27.
- Karami, M. and Wang, L. (2018) ‘Fault detection and diagnosis for nonlinear systems: A new adaptive Gaussian mixture modeling approach’, *Energy and Buildings*, 166, pp. 477–488.
- Katipamula, S., Pratt, R.G., Chassin, D.P., Taylor, Z.T., Gowri, K. and Brambley, M.R. (1999) ‘Automated fault detection and diagnostics for outdoor-air ventilation systems and economizers: methodology and results from field testing’, *ASHRAE Transactions*, 105(1), pp. 555–567.
- Katipamula, S. and Brambley, M.R. (2005) ‘Methods for Fault Detection, Diagnostics, and Prognostics for Building Systems— A Review, Part I’, *HVAC&R Research*, 11(1), pp. 3-25.
- Kelly, G., Park, C., Clark, D. and May, W. (1984) ‘HVACSIM+ a dynamic building-HVAC-control systems simulation program.’, Paper presented at the Workshop on HVAC controls, Modeling, and Simulation.
- Kharseh, M., Altorkmany, L., Al-Khawaj, M. and Hassani, F. (2014) ‘Warming impact on energy use of HVAC system in buildings of different thermal qualities and in different climates’, *Energy Conversion and Management*, 81, pp. 106–111.
- Khire, R. and Trcka, M. (2013) ‘Model based failure mode effect analysis on whole building energy performance’, *Proceedings of 13th Conference of International Building Performance Simulation Association*, Chambéry, France, 26-28 August. Available at: http://www.ibpsa.org/proceedings/BS2013/p_1069.pdf (Accessed: 20 June 2020).

- Kim, M., Payne, W.V., Domanski, P.A., Yoon, S.H. and Hermes, C.J.L. (2009) 'Performance of a residential heat pump operating in the cooling mode with single faults imposed', *Applied Thermal Engineering*, 29(4), pp. 770-778.
- Kim, J., Frank, S., Braun, J.E. and Goldwasser, D. (2019) 'Representing small commercial building faults in EnergyPlus, Part I: Model development', *Buildings*, 9(11), no. 233.
- Kim, J. and Leach, M. (2019) *Curated modeled fault data set*, Golden, United States: National Renewable Energy Laboratory, doi: <https://doi.org/10.25984/1862681>.
- Kim, W. and Lee, J. (2021) 'Fault detection and diagnostics analysis of air conditioners using virtual sensors', *Applied Thermal Engineering*, 191, no. 116848.
- Kingma, D.P. and Welling, M. (2014) 'Auto-Encoding Variational Bayes', *Proceedings of 2nd International Conference on Learning Representations (ICLR2014)*, Banff, Canada, 14-16 April. Available at: <https://arxiv.org/pdf/1312.6114.pdf> (Accessed: 12 Januray 2022)
- Kingma, D.P. and Ba, J. (2014) 'Adam: A Method for Stochastic Optimization', arXiv: 1412.6980.
- Krawczyk, B. (2016) 'Learning from imbalanced data: open challenges and future directions', *Progress in Artificial Intelligence*, 5, pp. 221-232.
- Kuhn, M. and Johnson, K. (2018) *Applied Predictive Modeling*. 2nd edn. Berlin: Springer.
- Li, S. and Wen, J. (2010) 'Development and validation of a dynamic air handling unit model, Part 1', *ASHRAE Transactions*, 116(1), pp. 45-56.
- Li, S., Wen, J., Zhou, X. and Klaassen, C.J. (2010) 'Development and validation of a dynamic air handling unit model, Part 2', *ASHRAE Transactions*, 116(1), pp. 57-73.
- Li, S. and Wen, J. (2014) 'A model-based fault detection and diagnostic methodology based on PCA method and wavelet transform', *Energy and Buildings*, 68, pp. 63-71.

Li, Y. and O'Neill, Z. (2016) 'An EnergyPlus/OpenStudio-based fault simulator for buildings', *Proceedings of 2016 ASHRAE Winter Meeting*, Orlando, United States, 24-28 January. Available at: https://www.techstreet.com/standards/or-16-c065-an-energyplus-openstudio-based-fault-simulator-for-buildings?product_id=1910129 (Accessed: 19 July 2020).

Li, Y. and O'Neill, Z. (2018) 'A critical review of fault modeling of HVAC systems in buildings', *Building Simulation*, 11, pp. 953–975.

Li, G., Chen, H., Hu, Y., Wang, J., Guo, Y., Liu, J., Li, H., Huang, R., Lv, H. and Li, J. (2018) 'An improved decision tree-based fault diagnosis method for practical variable refrigerant flow system using virtual sensor-based fault indicators', *Applied Thermal Engineering*, 129, pp. 1292–1303.

Li, G. and Hu, Y. (2018) 'Improved sensor fault detection, diagnosis and estimation for screw chillers using density-based clustering and principal component analysis', *Energy and Buildings*, 173, pp. 502–515.

Li, G. and Hu, Y. (2019) 'An enhanced PCA-based chiller sensor fault detection method using ensemble empirical mode decomposition based denoising', *Energy and Buildings*, 183, pp. 311–324.

Li, Y. and O'Neill, Z. (2019) 'An innovative fault impact analysis framework for enhancing building operations', *Energy and Buildings*, 199, pp. 311–331.

Li, X., Chen, S., Hu, X. and Yang, J. (2019) 'Understanding the Disharmony between Dropout and Batch Normalization by Variance Shift', *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Long Beach, United States, 15-20 June. no. 19262863.

Li, B., Cheng, F., Zhang, X., Cui, C. and Cai, W. (2021a) 'A novel semi-supervised data-driven method for chiller fault diagnosis with unlabeled data', *Applied Energy*, 285, no. 116459.

- Li, B., Cheng, F., Cai, H., Zhang, X. and Cai, W. (2021b) 'A semi-supervised approach to fault detection and diagnosis for building HVAC systems based on the modified generative adversarial network', *Energy and Buildings*, 246, no. 111044.
- Li, G., Yao, Q., Hu, Y., Fang, X. and Wang, L. (2022a) 'Investigating thermostat sensor offset impacts on operating performance and thermal comfort of three different HVAC systems in Wuhan, China', *Case Studies in Thermal Engineering*, 31, no. 101788.
- Li, D., Xu, Y., Zhou, Y., Gou, C. and Ng, S.K. (2022b) 'Cross Domain Data Generation for Smart Building Fault Detection and Diagnosis', *Mathematics*, 10, no. 3970.
- Li, Y., Im, P., Lee, S., Bae, Y., Yoon, Y. and Lee, S. (2023a) 'Sensor incipient fault impacts on building energy performance: a case study on a multi-zone commercial building', *Buildings*, 13, no. 520.
- Li, G., Chen, L., Liu, J. and Fang, X. (2023b) 'Comparative study on deep transfer learning strategies for cross-system and cross-operation-condition building energy systems fault diagnosis' *Energy*, 263, no. 125943.
- Liang, J. and Du, R. (2007) 'Model-based fault detection and diagnosis of hvac systems using support vector machine method', *International Journal of Refrigeration*, 30, pp. 1104-1114.
- Liang, X., Li, P., Chen, S., Jin, X. and Du, Z. (2022) 'Partial domain adaption based prediction calibration methodology for fault detection and diagnosis of chillers under variable operational condition scenarios', *Building and Environment*, 217, no. 109099.
- Liu, J., Chen, H., Wang, J., Li, G. and Li, H. (2016) 'Fault diagnosis of refrigerant charge based on PCA and decision tree for variable refrigerant flow systems', *Proceedings of 16th International Refrigeration and Air Conditioning Conference*, West Lafayette, United States, 11-14 July. Available at: <https://docs.lib.purdue.edu/iracc/1593/> (Accessed: 9 October 2019).

Liu, J., Liu, J., Chen, H., Yuan, Y., Li, Z. and Huang, R. (2018) ‘Energy diagnosis of variable refrigerant flow (VRF) systems: Data mining technique and statistical quality control approach’, *Energy and Buildings*, 175, pp. 148–162.

Liu, J., Zhang, M., Wang, H., Zhao, W. and Liu, Y. (2019) ‘Sensor Fault Detection and Diagnosis Method for AHU Using 1-D CNN and Clustering Analysis’, *Computational Intelligence and Neuroscience*, 2019, no. 5367217.

Liu, J., Zhang, Q., Li, X., Li, G., Liu, Z., Xie, Y. and Li, K. (2021) ‘Transfer learning-based strategies for fault diagnosis in building energy systems’, *Energy and Buildings*, 250, no. 111256.

Liu, Z., Niu, K. and He, Z. (2023) ‘ML-CookGAN: multi-label generative adversarial network for food image generation’, *ACM Transactions on Multimedia Computing, Communications, and Applications*, 19, no. 85.

Lee, S.H. and Yik, F.W.H. (2010) ‘A study on the energy penalty of various air-side system faults in buildings’, *Energy and Buildings*, 42(1), pp. 2-10.

Lee, W.Y., House, J.M. and Kyong, N.H. (2004) ‘Subsystem level fault diagnosis of a building’s air-handling unit using general regression neural networks’, *Applied Energy*, 77, pp. 153–170.

Lu, X., O’Neill, Z., Li, Y. and Niu, F. (2020) ‘A novel simulation-based framework for sensor error impact analysis in smart building systems: A case study for a demand-controlled ventilation system’, *Applied Energy*, 263, no. 114638.

Lu, X., Fu, Y., O’Neill, Z. and Wen, J. (2021) ‘A holistic fault impact analysis of the high-performance sequences of operation for HVAC systems: Modelica-based case study in a medium-office building’, *Energy and Buildings*, 252, no. 111448.

- Lu, C., Reddy, C.K., Wang, P., Nie, D. and Ning, Y. (2023) ‘Multi-Label Clinical Time-Series Generation via Conditional GAN’, *IEEE Transactions on Knowledge and Data Engineering* (Early Access).
- Luo, X.J., Fong, K.F., Sun, Y.J. and Leung, M.K.H. (2019) ‘Development of clustering-based sensor fault detection and diagnosis strategy for chilled water system’, *Energy and Buildings*, 186, pp. 17–36.
- Mao, Q., Fang, X., Hu, Y. and Li, G. (2018) ‘Chiller sensor fault detection based on empirical mode decomposition threshold denoising and principal component analysis’, *Applied Thermal Engineering*, 144, pp. 21–30.
- Mauro, A.W., Pelella, F. and Viscito, L. (2023) ‘Performance degradation of air source heat pumps under faulty conditions’, *Case Studies in Thermal Engineering*, 45, no. 103010.
- Mattera, C.G., Jradi, M., Skydt, M.R., Engelsgaard, S.S. and Shaker, H.R. (2020) ‘Fault detection in ventilation units using dynamic energy performance models’, *Journal of Building Engineering*, 32, no. 101635.
- Martinez-Viol, V., Urbano, E.M., Rangel, J.E.T., Delgado-Prieto, M. and Romeral, L. (2022) ‘Semi-Supervised Transfer Learning Methodology for Fault Detection and Diagnosis in Air-Handling Units’, *Applied Science*, 12, no. 8837.
- Marigo, M., Maccarini, A., Zarrella, A. and Afshari, A. (2023) ‘Fault impact analysis of ventilation systems in residential buildings: A simulation-based case study in Denmark’, *Energy and Buildings*, 292, no. 113150.
- Mavromatidis, G., Acha, S. and Shah, N. (2013) ‘Diagnostic tools of energy performance for supermarkets using Artificial Neural Network algorithms’, *Energy and Buildings*, 62, pp. 304–314.

Mehrabi, M. and Yuill, D. (2018) 'Generalized effects of faults on normalized performance variables of air conditioners and heat pumps', *International Journal of Refrigeration*, 85, pp. 409–430.

Mirnaghi, M.S. and Haghighat, F. (2020) 'Fault detection and diagnosis of large-scale HVAC systems in buildings using data-driven methods: A comprehensive review', *Energy and Buildings*, 229, no. 110492.

Mirza, M. and Osindero, S. (2014) 'Conditional Generative Adversarial Nets', arXiv: 1411.1784.

Myrefelt, S. (2004) 'Reliability and functional availability of HVAC systems', *Proceedings of the Fourth International Conference for Enhanced Building Operations*, Paris, France, 18-19 October. Available at: <https://core.ac.uk/download/pdf/79624117.pdf> (Accessed: 28 July 2020).

Namburu, S.M., Azam, M.S., Luo, J., Choi, K. and Pattipati, K.R. (2007) 'Data-Driven Modeling, Fault Diagnosis and Optimal Sensor Selection for HVAC Chillers', *IEEE Transactions on Automation Science and Engineering*, 4, no. 3.

Otto, K., Eisenhower, B., O'Neill, Z., Yuan, S., Mezic, I. and Narayanan, S. (2012) 'Prioritizing building system energy failure modes using whole building energy simulation', *Proceedings of the Fifth National Conference of IBPSA-USA*, Madison, Wisconsin, United States, 1-3 August. Available at: <http://ibpsa-usa.org/index.php/ibpusa/article/view/474> (Accessed: 19 February 2020).

Padilla, M. and Choinière, D. (2015) 'A combined passive-active sensor fault detection and isolation approach for air handling units', *Energy and Buildings*, 99, pp. 214-219.

Park, B.J., Kim, Y.H., Min, S.K. and Lim, E.P. (2018) 'Anthropogenic and natural contributions to the lengthening of the summer season in the northern hemisphere', *Journal of Climate*, 31(17), pp. 6803–6819.

- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V. and Thirion, B. (2011) ‘Scikit-learn: Machine Learning in Python’, *Journal of Machine Learning Research*, 12, pp. 2825-2830.
- Pelella, F., Viscito, L. and Mauro, A.W. (2022a) ‘Combined effects of refrigerant leakages and fouling on air-source heat pump performances in cooling mode’, *Applied Thermal Engineering*, 204, no. 117965.
- Pelella, F., Viscito, L. and Mauro, A.W. (2022b) ‘Soft faults in residential heat pumps: Possibility of evaluation via on-field measurements and related degradation of performance’, *Energy Conversion and Management*, 260, no. 115646.
- Pourarian, S., Wen, J., Veronica, D., Pertzborn, A., Zhou, X. and Liu, R. (2017) ‘A tool for evaluating fault detection and diagnostic methods for fan coil units’, *Energy and Buildings*, 136, pp. 151-160.
- Remund, J., Müller, S., Schmutz, M., Barsotti, D., Graf, P. and Cattin, R. (2020a) *Meteonorm Handbook part I: Software*, Bern, Switzerland: METEOTEST.
- Remund, J., Müller, S., Schmutz, M., Barsotti, D., Graf, P. and Cattin, R. (2020b) *Meteonorm Handbook part II: Theory*, Bern, Switzerland: METEOTEST.
- Rosato, A., Guarino, F., Sibilio, S., Entchev, E., Masullo, M. and Maffei, L. (2021) ‘Healthy and Faulty Experimental Performance of a Typical HVAC System under Italian Climatic Conditions: Artificial Neural Network-Based Model and Fault Impact Assessment’, *Energies*, 14(17), no. 5362.
- Rosato, A., Guarino, F., Youssef, M.E., Capozzoli, A., Masullo, M. and Maffei, L. (2022a) ‘Faulty Operation of Coils’ and Humidifier Valves in a Typical Air-Handling Unit: Experimental Impact Assessment of Indoor Comfort and Patterns of Operating Parameters under Mediterranean Climatic Conditions’, *Energies*, 15(18), no. 6781.

Rosato, A., Guarino, F., Youssef, M.E., Capozzoli, A., Masullo, M. and Maffei, L. (2022b) ‘Experimental assessment of ground-truth faults in a typical single-duct dual-fan air-handling unit under Mediterranean climatic conditions: Impact scenarios of sensors’ offset and fans’ failure’, *Energy and Buildings*, 275, no. 112492.

Ruder, S. (2017) ‘An overview of gradient descent optimization algorithms’, arXiv:1609.04747.

Rumelhart, D.E., Hinton, G.E. and Williams, R.J. (1986) ‘Learning representations by back-propagating errors’, *Nature*, 323, pp. 533–536.

Sasaki, Y. (2007) ‘The truth of the F-measure’. CS795. Old Dominion University.
<https://www.cs.odu.edu/~mukka/cs795sum09dm/Lecturenotes/Day3/F-measure-YS-26Oct07.pdf>.

Schein, J. and Bushby, S. (2005) *A simulation study of a hierarchical, rule-based method for system-level fault detection and diagnostics in HVAC systems*. Gaithersburg: National Institute of Standards and Technology. Available at:
https://www.nist.gov/publications/simulation-study-hierarchical-rule-based-method-system-level-fault-detection-and?pub_id=860972 (Accessed: 8 August 2020).

Schein, J., Bushby, S.T., Castro, N.S. and House, J.M. (2006) ‘A rule-based fault detection method for air handling units’, *Energy and Buildings*, 38(12), pp. 1485-1492.

Seem, J.E. (2007) ‘Using intelligent data analysis to detect abnormal energy consumption in buildings’, *Energy and Buildings*, 39(1), pp. 52–58.

Shahnazari, H., Mhaskar, P., House, J.M. and Salsbury, T.I. (2019) ‘Modeling and fault diagnosis design for HVAC systems using recurrent neural networks’, *Computers and Chemical Engineering*, 126, pp. 189–203.

Shen, C., Zhang, H., Meng, S. and Li, C. (2023) ‘Augmented data driven self-attention deep learning method for imbalanced fault diagnosis of the HVAC chiller’, *Engineering Applications of Artificial Intelligence*, 117, no. 105540.

Shi, S., Li, G., Chen, H., Liu, J., Hu, Y., Xing, L. and Hu, W. (2017) ‘Refrigerant charge fault diagnosis in the VRF system using Bayesian artificial neural network combined with ReliefF filter’, *Applied Thermal Engineering*, 112, pp. 698–706.

Shi, S., Li, G., Chen, H., Hu, Y., Wang, X., Guo, Y. and Sun, S. (2018) ‘An efficient VRF system fault diagnosis strategy for refrigerant charge amount based on PCA and dual neural network model’, *Applied Thermal Engineering*, 129, pp. 1252–1262.

Shi, Z. and O'Brien, W. (2019) ‘Development and implementation of automated fault detection and diagnostics for building systems: A review’, *Automation in Construction*, 104, pp. 215-229.

Shorten, C. and Khoshgoftaar, T.M. (2019) ‘A survey on Image Data Augmentation for Deep Learning’, *Journal of Big Data*, 6, no. 60.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. (2014) ‘Dropout: A Simple Way to Prevent Neural Networks from Overfitting’, *Journal of Machine Learning Research*, 15, pp. 1929-1958.

Sterling R., Provan G., Febres J., O’Sullivan D., Struss P., Keane M.M. (2014) ‘Model-based fault detection and diagnosis of air handling units: A comparison of methodologies’, *Energy Procedia*, 62, pp. 686–693.

Sun, K., Li, G., Chen, H., Liu, J., Li, J. and Hu, W. (2016) ‘A novel efficient SVM-based fault diagnosis method for multi-split air conditioning system’s refrigerant charge fault amount’, *Applied Thermal Engineering*, 108, pp. 989–998.

Sun, S., Li, G., Chen, H., Huang, Q., Shi, S. and Hu, W. (2017) 'A hybrid ICA-BPNN-based FDD strategy for refrigerant charge faults in variable refrigerant flow system', *Applied Thermal Engineering*, 127, pp. 718–728.

Sun, Z., Jin, H., Gu, J., Huang, Y., Wang, X. and Shen, X. (2019) 'Gradual fault early stage diagnosis for air source heat pump system using deep learning techniques', *International Journal of Refrigeration*, 107, pp. 63–72.

Taheri, S., Ahmadi, A., Mohammadi-Ivatloo, B. and Asadi, S. (2021) 'Fault detection diagnostic for HVAC systems via deep learning algorithms', *Energy and Buildings*, 250, no. 111275.

Tian, C., Wang, Y., Ma, X., Chen, Z. and Xue, H. (2021) 'Chiller Fault Diagnosis Based on Automatic Machine Learning', *Frontiers in Energy Research*, 9, no. 753732.

Tran, D.A.T., Chen, Y., Ao, H.L., Cam, H.N.T. (2016) 'An enhanced chiller FDD strategy based on the combination of the LSSVR-DE model and EWMA control charts', *International Journal of Refrigeration*, 72, pp. 81–96.

Tran, D.A.T., Chen, Y., Jiang, C. (2016) 'Comparative investigations on reference models for fault detection and diagnosis in centrifugal chiller systems', *Energy and Buildings*, 133, pp. 246–256.

Tun, W., Wong, K.W. and Ling, S.H. (2023) 'Advancing Fault Detection in HVAC Systems: Unifying Gramian Angular Field and 2D Deep Convolutional Neural Networks for Enhanced Performance', *Sensors*, 23, no. 7690.

UK Department for Business, Energy and Industrial Strategy. (2021) *Prices of fuels purchased by non-domestic consumers in the UK*. London, United Kingdom: UK Department for Business, Energy and Industrial Strategy. Available at: <https://www.gov.uk/government/statistical-data-sets/gas-and-electricity-prices-in-the-non-domestic-sector> (Accessed: 4 January 2022).

UK Met Office. (2021) *MIDAS Open: UK hourly weather observation data*, v202107. Chilton, United Kingdom: NERC EDS Centre for Environmental Data Analysis. doi: <https://dx.doi.org/10.5285/3bd7221d4844435dad2fa030f26ab5fd>

United Nations Framework Convention on Climate Change (UNFCCC). (2021) *Glasgow Climate Pact*. New York, United States: UNFCCC. Available at: <https://unfccc.int/documents/310475> (Accessed: 22 November 2021).

U.S. Department of Energy (DOE). (2020a) *EnergyPlus Version 9.3 Input Output Reference*. Washington, D.C., United States: DOE.

U.S. Department of Energy (DOE). (2020b) *EnergyPlus Version 9.3 Engineering Reference*. Washington, D.C., United States: DOE.

van Every, P.M., Rodriguez, M., Jones, C.B., Mammoli, A.A. and Martínez-Ramón, M. (2017) ‘Advanced detection of HVAC faults using unsupervised SVM novelty detection and Gaussian process models’, *Energy and Buildings*, 149, pp. 216–224.

van Rossum, G. and Drake, F.L. (2019) *Python 3.6.9 Reference Manual*. Scotts Valley, United States: CreateSpace.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L. and Polosukhin, I. (2017) ‘Attention Is All You Need’, *Proceedings of 31st Conference on Neural Information Processing Systems (NIPS 2017)*, Long Beach, United States, 4 – 9 December. Available at: https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf (Accessed: 10 January 2022).

Venkatasubramanian, V., Rengaswamy, R., Yin, K. and Kavuri, S.N. (2003) ‘A review of process fault detection and diagnosis Part I: Quantitative model-based methods’, *Computers and Chemical Engineering*, 27(3), pp. 293–311.

Verhelst, J., Van Ham, G., Saelens, D. and Helsen, L. (2017) 'Economic impact of persistent sensor and actuator faults in concrete core activated office buildings', *Energy and Buildings*, 142, pp. 111–127.

Waddicor, D. A., Fuentes, E., Sisó, L., Salom, J., Favre, B., Jiménez, C. and Azar, M. (2016) 'Climate change and building ageing impact on building energy performance and mitigation measures application: A case study in Turin, northern Italy', *Building Environment*, 102, pp. 13-25.

Wang, S. and Xiao, F. (2004) 'AHU sensor fault diagnosis using principal component analysis method', *Energy and Buildings*, 36, pp. 147–160.

Wang, S. and Cui, J. (2005) 'Sensor-fault detection, diagnosis and estimation for centrifugal chiller systems using principal-component analysis method', *Applied Energy*, 82, pp. 197–213.

Wang, S. and Xiao, F. (2006) 'Sensor Fault Detection and Diagnosis of Air-Handling Units Using a Condition-Based Adaptive Statistical Method', *HVAC&R Research*, 12, no. 1.

Wang, L. and Hong, T. (2012) 'Modeling and simulation of HVAC faulty operations and performance degradation due to maintenance issues', *Proceedings of 1st Asia conference of International Building Performance Simulation Association*, Shanghai, China, 25-27 November. Available at: <https://escholarship.org/uc/item/1rn0h4mr> (Accessed: 6 August 2020)

Wang, P. and Gao, R.X. (2017) 'Automated Performance Tracking for Heat Exchangers in HVAC', *IEEE Transactions on Automation Science and Engineering*, 14, no. 2.

Wang, Z., Wang, L., Liang, K. and Tan, Y. (2018) 'Enhanced chiller fault detection using Bayesian network and principal component analysis', *Applied Thermal Engineering*, 141, pp. 898–905.

Wang, Z., Dong, Y., Liu, W. and Ma, Z. (2020) ‘A Novel Fault Diagnosis Approach for Chillers Based on 1-D Convolutional Neural Network and Gated Recurrent Unit’, *Sensors*, 20, no. 2458.

Wang, H., Feng, D. and Liu, K. (2021) ‘Fault detection and diagnosis for multiple faults of VAV terminals using self-adaptive model and layered random forest’, *Building Environment*, 193, no. 107667.

Wang, J., Guan, Y., Wu, L., Guan, X., Cai, W., Huang, J., Dong, W. and Zhang, B. (2021) ‘Changing Lengths of the Four Seasons by Global Warming’, *Geophysical Research Letters*, 48(6), no. e2020GL091753.

Walton, G.N. (1983) *Thermal Analysis Research Program Reference Manual*. Gaithersburg, United States: National Bureau of Standards.

Wen, J. and Li, S. (2011) *Tools for evaluating fault detection and diagnostic methods for air-handling units*. Atlanta, United States: American society of heating, refrigeration and air-conditioning engineers.

Wen, S., Zhang, W., Sun, Y., Li, Z., Huang, B., Bian, S. Zhao, L. and Wang, Y. (2023) ‘An enhanced principal component analysis method with Savitzky–Golay filter and clustering algorithm for sensor fault detection and diagnosis’, *Applied Energy*, 337, no. 120862.

Wetter, M., Zuo, W., Nouidui, T.S. and Pang, X. (2014) ‘Modelica buildings library’, *Journal of Building Performance Simulation*, 7, pp. 253-270.

Wold, S., Esbensen, K. and Geladi, P. (1987) ‘Principal component analysis’, *Chemometrics and Intelligent Laboratory Systems*, 2(1), pp. 37–52.

Winkler, J., Das, S., Earle, L., Burkett, L., Robertson, J., Roberts, D. and Booten, C. (2020) ‘Impact of installation faults in air conditioners and heat pumps in single- family homes on U.S. energy usage’, *Applied Energy*, 278, no. 115533.

Wu, S. and Sun J. (2011) ‘Cross-level fault detection and diagnosis of building HVAC systems’, *Building and Environment*, 46(8), pp. 1558-1566.

Xu, X., Xiao, F. and Wang, S. (2008) ‘Enhanced chiller sensor fault detection, diagnosis and estimation using wavelet analysis and principal component analysis methods’, *Applied Thermal Engineering*, 28(2-3), pp. 226-237.

Yan, R., Ma, Z., Zhao, Y. and Kokogiannakis, G. (2016) ‘A decision tree based data-driven diagnostic strategy for air handling units’, *Energy and Buildings*, 133, pp. 37–45.

Yan, K., Zhong, C., Ji, Z. and Huang, J. (2018) ‘Semi-supervised learning for early detection and diagnosis of various air handling unit faults’, *Energy and Buildings*, 181, pp. 75–83.

Yan, K., Huang, J., Shen, W. and Ji, Z. (2020) ‘Unsupervised learning for fault detection and diagnosis of air handling units’, *Energy and Buildings*, 210, no. 109689.

Yan, K., Chong, A. and Mo, Y. (2020) ‘Generative adversarial network for fault detection diagnosis of chillers’, *Building and Environment*, 172, no. 106698.

Yan, K. (2021) ‘Chiller fault detection and diagnosis with anomaly detective generative adversarial network’, *Building and Environment*, 201, no. 107982.

Yan, K., Su, J., Huang, J. and Mo, Y. (2022) ‘Chiller Fault Diagnosis Based on VAE-Enabled Generative Adversarial Networks’, *IEEE Transactions on Automation Science and Engineering*, 19(1), pp. 387 – 395.

Yan, K., Chen, X., Zhou, X., Yan, Z. and Ma, J. (2023a) ‘Physical Model Informed Fault Detection and Diagnosis of Air Handling Units Based on Transformer Generative Adversarial Network’, *IEEE Transactions on Industrial Informatics*, 19(2), pp. 2192 – 2199.

Yan, K., Lu, C., Ma, X., Ji, Z. and Huang, J. (2023b) ‘Intelligent fault diagnosis for air handling units based on improved generative adversarial network and deep reinforcement learning’, *Expert Systems with Applications* (in press).

Yang, X., Jin, X., Du, Z., Zhu, Y. and Guo, Y. (2013) 'A hybrid model-based fault detection strategy for air handling unit sensors', *Energy and Buildings*, 57, pp. 132–143.

Yoon, S., Yu, Y., Wang, J. and Wang, P. (2019) 'Impacts of HVACR temperature sensor offsets on building energy performance and occupant thermal comfort', *Building Simulation*, 12, pp. 259–271.

Yu, Z., Haghighat, F., Fung, B.C.M. and Zhou, L. (2012) 'A novel methodology for knowledge discovery through mining associations between building operational data', *Energy and Buildings*, 47, pp. 430–440.

Yu, Y., Woradechjumroen, D. and Yu, D. (2014) 'A review of fault detection and diagnosis methodologies on air-handling units', *Energy and Buildings*, 82, pp. 550-562.

Yu, F., Li, G., Chen, H., Guo, Y., Yuan, Y. and Coulton, B. (2018) 'A VRF charge fault diagnosis method based on expert modification C5.0 decision tree', *International Journal of Refrigeration*, 92, pp. 106-112.

Yuill, D.P., Cheung, H. and Braun, J.E. (2014) 'Validation of a Fault-Modeling Equipped Vapor Compression System Model Using a Fault Detection and Diagnostics Evaluation Tool', *Proceedings of 15th International Refrigeration and Air Conditioning Conference*, West Lafayette, United States, 14-17 July. Available at: <http://docs.lib.purdue.edu/iracc/1544> (Accessed: 2 August 2020).

Zhao, Y., Wang, S. and Xiao, F. (2013a) 'A statistical fault detection and diagnosis method for centrifugal chillers based on exponentially-weighted moving average control charts and support vector regression', *Applied Thermal Engineering*, 51, pp. 560-572.

Zhao, Y., Wang, S. and Xiao, F. (2013b) 'A system-level incipient fault-detection method for HVAC systems', *HVAC&R Research*, 19, pp. 593–601.

Zhao, Y., Li, T., Zhang, X. and Zhang, C. (2019) 'Artificial intelligence-based fault detection and diagnosis methods for building energy systems: Advantages, challenges and the future', *Renewable and Sustainable Energy Reviews*, 109, pp. 85-101.

Zhang, R. and Hong, T. (2017) 'Modeling of HVAC operational faults in building performance simulation', *Applied Energy*, 202, pp. 178-188.

Zhang, Y., Tennakoon, T., Chan, Y.H., Chan, K.C., Fu, S.C., Tso, C.Y., Yu, K.M., Huang, B. L., Yao, S.H., Qiu, H.H., Chao, C.Y.H. (2022a) 'Energy consumption modelling of a passive hybrid system for office buildings in different climates', *Energy*, 239, no. 121914.

Zhang, H., Li, C., Wei, Q. and Zhang, Y. (2022b) 'Fault detection and diagnosis of the air handling unit via combining the feature sparse representation based dynamic SFA and the LSTM network', *Energy and Buildings*, 269, no. 112241.

Zhang, Q., Tian, Z., Niu, J., Zhu, J. and Lu, Y. (2022c) 'A study on transfer learning in enhancing performance of building energy system fault diagnosis with extremely limited labeled data', *Building and Environment*, 225, no. 109641.

Zhang, J., Li, Z., Chen, H., Cheng, H., Xing, L., Wang, Y. and Zhang, L. (2022d) 'Integrated generative networks embedded with ensemble classifiers for fault detection and diagnosis under small and imbalanced data of building air condition system', *Energy and Buildings*, 268, no. 112207.

Zhang, F., Saeed, N. and Sadeghian, P. (2023) 'Deep learning in fault detection and diagnosis of building HVAC systems: A systematic review with meta analysis', *Energy and AI*, 12, no. 100235.

Zhang, J., Xu, Y., Chen, H. and Xing, L. (2023a) 'A novel building heat pump system semi-supervised fault detection and diagnosis method under small and imbalanced data', *Engineering Applications of Artificial Intelligence*, 123, no. 106316.

Zhang, H., Yang, W., Yi, W., Lim, J.B., An, Z. and Li, C. (2023b) ‘Imbalanced data based fault diagnosis of the chiller via integrating a new resampling technique with an improved ensemble extreme learning machine’, *Journal of Building Engineering*, 70, no. 106338.

Zheng, M., Li, T., Zhu, R., Tang, Y., Tang, M., Lin, L. and Ma, Z. (2020) ‘Conditional Wasserstein generative adversarial network-gradient penalty-based approach to alleviating imbalanced data classification’, *Information Sciences*, 512, pp. 1009–1023.

Zhong, C., Yan, K., Dai, Y., Jin, N. and Lou, B. (2019) ‘Energy Efficiency Solutions for Buildings: Automated Fault Diagnosis of Air Handling Units Using Generative Adversarial Networks’, *Energies*, 12, no. 527.

Zhong, F., Calautit, J.K. and Wu, Y. (2022) ‘Assessment of HVAC system operational fault impacts and multiple faults interactions under climate change’, *Energy*, 258, no. 124762.

Zhong, F., Calautit, J.K. and Wu, Y. (2023) ‘Fault data seasonal imbalance and insufficiency impacts on data-driven heating, ventilation and air-conditioning fault detection and diagnosis performances for energy-efficient building operations’, *Energy*, 282, no. 128180.

Zhou, Q., Wang, S. and Xiao, F. (2009) ‘A Novel Strategy for the Fault Detection and Diagnosis of Centrifugal Chiller Systems’, *HVAC&R Research*, 15, no. 1.

Zhou, Z., Chen, H., Li, G., Zhong, H., Zhang, M. and Wu, J. (2021) ‘Data-driven fault diagnosis for residential variable refrigerant flow system on imbalanced data environments’, *International Journal of Refrigeration*, 125, no. 34–43.

Zhou, Z., Li, G., Chen, H. and Zhong, H. (2021) ‘Fault diagnosis method for building VRF system based on convolutional neural network: Considering system defrosting process and sensor fault coupling’, *Building and Environment*, 195, no. 107775.

Zogg, D., Shafai, E. and Geering, H.P. (2006) ‘Fault diagnosis for heat pumps with parameter identification and clustering’, *Control Engineering Practice*, 14, pp. 1435–1444.

Zhou, Z., Chen, H., Xing, L., Li, G. and Gou, W. (2022) 'An experimental study of the behavior of a model variable refrigerant flow system with common faults', *Applied Thermal Engineering*, 202, no. 117852.

Zhu, X., Du, Z., Jin, X. and Chen, Z. (2019) 'Fault diagnosis based operation risk evaluation for air conditioning systems in data centers', *Building and Environment*, 163, no. 106319.