5-2018

Automated Fault Detection, Diagnostics, Impact Evaluation, and Service Decision-Making for Direct Expansion Air Conditioners

Andrew L. Hjortland
Purdue University

AUTOMATED FAULT DETECTION, DIAGNOSTICS, IMPACT EVALUATION, AND SERVICE DECISION-MAKING FOR DIRECT EXPANSION AIR CONDITIONERS

by

Andrew L. Hjortland

A Dissertation

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Doctor of Philosophy

School of Mechanical Engineering

West Lafayette, Indiana

May 2018

THE PURDUE UNIVERSITY GRADUATE SCHOOL

STATEMENT OF COMMITTEE APPROVAL

Dr. James E. Braun, Chair

Department of Mechanical Engineering

Dr. W. Travis Horton

Department of Architectural Engineering

Dr. Eckhard A. Groll

Department of Mechanical Engineering

Dr. Neera Jain

Department of Mechanical Engineering

Approved by:

Dr. Jay P. Gore

Head of the Graduate Program

For my thoughtful and loving wife, Nicole, for always motivating me

For my supportive parents, Angie and Randy, for laying the foundation

ACKNOWLEDGMENTS

I am very grateful to Professor James. E. Braun for his support, guidance, and attention to detail through my research at Ray W. Herrick Laboratories. Upon arriving at Purdue many years ago, his excellent instruction and regular research meetings helped me become a better researcher and motivated my interests in building science. I also appreciate his patience and encouragement to forge my own way in my research endeavors and numerous side projects. I also would like to thank my advisory committee members: Professor W. Travis Horton, Professor Eckhard A. Groll, and Professor Neera Jain for their valuable feedback and support.

I would also like to thank several research sponsors that have helped support my work. Notably, Lennox International, Inc provided both financial and technical support, especially from Jon Douglas. The support of Johnson Controls, Inc. is also appreciated, with special consideration to Kirk Drees for encouraging me to apply to Purdue University – literally starting the chain of events that has concluded with this dissertation.

Without fail, the great community at Ray W. Herrick Laboratories over the past several years has always been there to help. Special thanks must be given to the excellent technical support from the shop provided by Frank Lee, Orkan Kurtulus, and Bob Brown. I must also thank all my fellow graduate students who helped me along the way; especially Woohyun Kim, Howard Cheung, Christian Bach, Abhinav Krishna, Donghun Kim, Jie Cai, and Akash Patil.

In closing, I must acknowledge all the support from my life outside of Purdue University. First and foremost, I am indebted to my parents, Randy and Angie Hjortland, for supporting my education and always pushing me to do bigger and better things. I must also thank my brothers Dan and Sam for keeping me sane. Finally, this work could not have been completed without my wife, Nicole, who shared with me all her love and support while we both navigated through our time at Purdue.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS iv
TABLE OF CONTENTS v
LIST OF TABLES vii
LIST OF FIGURES x
LIST OF SYMBOLS xx
ABSTRACT xxiii
1. INTRODUCTION 1
1.1 Background and Motivation 1
1.2 Research Objectives 6
1.3 Thesis Organization 7
2. AUTOMATED FAULT DETECTION AND DIAGNOSTICS FOR UNITARY AIR CONDITIONING EQUIPMENT 10
2.1 Background and Motivation 10
2.2 Literature Review of Previous Work on AFDD for DX Systems 11
2.3 Description of AFDD System for Multi-stage and Variable-Speed RTUs 14
2.3.1 Overview of AFDD System Architecture 15
2.3.2 Description of Virtual Sensors used in AFDD Approach 17
2.4 Experimental Evaluation of AFDD System Applied to Multistage RTU 23
2.4.1 Description of Multistage RTU and Experimental Setup 23
2.4.2 Measured Detection and Diagnosis Performance of AFDD System 26
2.5 Experimental Evaluation of AFDD System Applied to Variable-Speed RTU 32
2.5.1 Description of Variable-Speed RTU and Experimental Setup 33
2.5.2 Extending AFDD System for Variable-Speed RTU 40
2.6 Description of AFDD System Hardware Design and Prototype 44
2.7 Summary of AFDD System and Important Results 49
3. SIMPLIFIED FAULT IMPACT MODELS FOR UNITARY AIR CONDITIONING EQUIPMENT 51
3.1 Background and Motivation 51

3.2 RTU Fault Impact Estimation Literature Review..... 53
3.3 Simplified Models for Normal Equipment Performance..... 56
3.3.1 Semi-Empirical Models for Fixed-Speed RTU Performance..... 57
3.3.2 Semi-Empirical Models for Variable-Speed RTU Performance..... 60
3.4 Estimating Actual Equipment Performance Using Virtual Sensors..... 75
3.5 Estimating Overall Performance Impacts of Faults..... 77
3.5.1 Estimating Total Cooling Capacity Impact of Faults..... 77
3.5.2 Estimating Cycle Efficiency Impacts of Faults..... 81
3.5.3 Estimating Sensible Heat Ratio Impacts of Faults..... 85
3.5.4 Estimating Run-time Impacts of Faults..... 89
3.5.5 Estimating Energy Consumption Impacts of Faults..... 93
3.6 Estimating Impacts of Multiple Faults using Semi-Empirical Models..... 97
3.7 Summary of Fault Impact Estimation Results..... 108
4. MODELING DX EQUIPMENT FAULT IMPACTS USING ARTIFICIAL NEURAL NETWORKS..... 110
4.1 Background and Motivation..... 110
4.2 Fault Impact Meta-Modeling Approach..... 112
4.3 Fault Impact Meta-model Training and Results..... 115
4.3.1 Overall Prediction Accuracy of Meta-model..... 115
4.3.2 Refrigerant Charge Fault Impact Prediction Accuracy..... 119
4.3.3 Condenser Fouling Fault Impact Prediction Accuracy..... 120
4.3.4 Evaporator Fouling Fault Impact Prediction Accuracy..... 121
4.3.5 Compressor Valve Leakage Fault Impact Prediction Accuracy..... 123
4.4 Review of Important Results..... 124
5. SIMULATING AIR CONDITIONING EQUIPMENT FAULTS IN SMALL COMMERCIAL BUILDINGS..... 127
5.1 Background and Motivation..... 127
5.2 Building Load Model Description..... 127
5.2.1 Climate and Weather Model Description..... 131
5.2.2 Utility, Equipment, and Service Cost Models..... 135
5.3 Fault Evolution Model Descriptions..... 137

6. COMPARING MAINTENANCE STRATEGIES FOR SMALL COMMERCIAL BUILDING AIR CONDITIONING EQUIPMENT 140
6.1 Background and Motivation 140
6.2 Literature Review of DX Equipment Service Decision Support 141
6.3 Optimal Service Policies 142
6.3.1 Optimal Service Decision Problem Formulation 142
6.3.2 Demonstration Results for Optimal Maintenance Schedules 147
6.4 Periodic Service Policies 164
6.4.1 Description and Implementation of Periodic Service Policies 164
6.4.2 Demonstration Results of Periodic Service Policies 165
6.5 Emergency Service Policies 177
6.5.1 Description and Implementation of Emergency Service Policies 177
6.5.2 Demonstration Results of Emergency Service Policies 178
6.6 Condition-based Service Policies 182
6.7 Operating Cost-based Service Policies 184
6.7.1 Rossi and Braun's Near Optimal Service Policy 185
6.7.2 Modification for Multiple Simultaneous Faults 187
6.8 Summary and Conclusions 203
7. CONCLUSION AND RECOMMENDATIONS 204
REFERENCES 210
APPENDIX A. LOAD-BASED TESTING METHOD FOR AIR CONDITIONING EQUIPMENT 218
APPENDIX B. COMPARISONS OF FAULT IMPACT NEURAL NETWORK META-MODEL AND DETAILED MODEL OUTPUTS 248
VITA 272
PUBLICATIONS 273

LIST OF TABLES

Table 2.1. Component descriptions of multi-stage RTU used for development and evaluation of AFDD implementation..... 24

Table 2.2. Rated cooling performance of multi-stage RTU used for development and evaluation of AFDD implementation..... 25

Table 2.3. Test conditions for RTU with finned-tube condenser and fixed orifice expansion device for low stage cooling operation in psychrometric test chambers..... 26

Table 2.4. Component descriptions of variable-speed RTU used for development and evaluation of AFDD implementation..... 35

Table 2.5. Rated cooling performance of variable-speed RTU used for development and evaluation of AFDD implementation..... 35

Table 2.6. Steady-state test scenarios for variable-speed RTU AFDD performance assessment..... 37

Table 2.7. Refrigerant-side sensor measurement uncertainty for sensors installed on variable-speed RTU..... 39

Table 2.8. Description of required refrigerant-side temperature sensors used for RTU AFDD methods..... 46

Table 3.1. Experimental test conditions for variable-speed RTU normal performance model development and evaluation. Experimental variables were indoor dry bulb temperature, indoor relative humidity, outdoor dry bulb temperature, indoor fan speed, and discharge air setpoint..... 64

Table 5.1. Building HVAC design conditions for selected locations used in fault impact simulation program [61]..... 131

Table 5.2. Building sensible and latent load parameters used in simulation framework for different locations..... 135

Table 6.1. Summary of service task costs used in equipment fault and maintenance simulations. When multiple tasks are performed, a 20% discount to the total cost is applied..... 148

Table 6.2. Summary of fault evolution rate parameters used in first multiple fault simulation and service scheduling optimization..... 161

Table 6.3. Summary of fault evolution rate parameters used in second multiple fault simulation and service scheduling optimization..... 162

Table 6.4. Summary of additional lifetime operating costs relative to optimal lifetime costs determined using dynamic programming for similar buildings in different locations. Inner-quartile ranges and means of 200 randomly sampled fault scenarios are compared..... 202

LIST OF FIGURES

Figure 1.1. General approach for automated fault detection and diagnostics described by Rossi and Braun [1]..... 3

Figure 2.1. Overall architecture of VOLTTRON™ enabled RTU AFDD system as an additional module of current RTU control units..... 16

Figure 2.3. VRC sensor prediction accuracy for RTU with finned tube condenser and fixed orifice expansion device applied to both stages of operation under different ambient conditions..... 24

Figure 2.4. Prediction accuracy of VRC sensor applied to charge level vaults showing charge fault probability for each test case. The data show that the FDD system can identify faults with high confidence when actual charge is ±20% of the normal charge level..... 27

Figure 2.5. Total capacity impact of refrigerant charge faults at 27.78 °C (82 °F) outdoor ambient temperature. The points are colored based on the fault probability determined for each test case, indicating faults that have larger impacts are identified with greater probability. IDF indicates the indoor fan torque setting; ODF indicates the outdoor fan speed setting..... 28

Figure 2.6. Total capacity impact of refrigerant charge faults at 35.00 °C (95 °F) outdoor ambient temperature. The points are colored based on the fault probability determined for each test case, indicating faults that have larger impacts are identified with greater probability. IDF indicates the indoor fan torque setting; ODF indicates the outdoor fan speed setting..... 29

Figure 2.7. Total capacity impact of refrigerant charge faults at 41.22 °C (108 °F) outdoor ambient temperature. The points are colored based on the fault probability determined for each test case, indicating faults that have larger impacts are identified with greater probability. IDF indicates the indoor fan torque setting; ODF indicates the outdoor fan speed setting..... 30

Figure 2.8. The FDD system can identify test cases with low evaporator air flow using the statistical fault detection and diagnostics method. Test cases without reductions in evaporator air flow were not identified with high fault probabilities. IDF indicates the indoor fan torque setting; ODF indicates the outdoor fan speed setting..... 31

Figure 2.9. The FDD system can identify test cases with low condenser air flow using the statistical fault detection and diagnostics method for test cases under 27.78 °C (82 °F) outdoor ambient temperature. Test cases without reductions in condenser air flow were not identified with high fault probabilities. IDF indicates the indoor fan torque setting; ODF indicates the outdoor fan speed setting..... 32

Figure 2.10. 5-ton variable-speed RTU that will be used to develop and evaluate AFDD implementation..... 34

Figure 2.11. Refrigerant-side temperature, pressure, and mass flow sensor layout for variable-speed RTU..... 39

Figure 2.12. Air-side temperature, pressure, and humidity sensor layout for variable-speed RTU..... 40

Figure 2.13. Variable-speed RTU virtual refrigerant mass flow rate sensor accuracy over range of operating conditions..... 41

Figure 2.14. Variable-speed RTU virtual compressor power sensor accuracy over range of operating conditions..... 42

Figure 2.15. Variable-speed RTU virtual condenser air flow rate sensor accuracy over range of operating conditions..... 43

Figure 2.16. Variable-speed RTU virtual evaporator air flow rate sensor accuracy over range of operating conditions..... 44

Figure 2.17. Locations of temperature (T) and relative humidity (ϕ) sensors used by RTU AFDD system..... 47

Figure 2.18. RTU AFDD system through-hole prototype implementation using pre-fabricated prototyping electronics boards..... 49

Figure 3.1. Prediction accuracy of semi-empirical model for total cooling capacity of a variable-speed RTU operating without faults over a range of ambient conditions and loads..... 66

Figure 3.2. Prediction accuracy of semi-empirical model for total power consumption of a variable-speed RTU operating without faults over a range of ambient conditions and loads..... 67
Figure 3.3. Prediction accuracy of semi-empirical model for coefficient of performance of a variable-speed RTU operating without faults over a range of ambient conditions and loads..... 68
Figure 3.4. Prediction accuracy of semi-empirical model for sensible cooling capacity of a variable-speed RTU operating without faults over a range of ambient conditions and loads..... 69
Figure 3.5. Prediction accuracy of semi-empirical model for sensible heat ratio of a variable-speed RTU operating without faults over a range of ambient conditions and loads..... 70
Figure 3.6. With return air wet bulb (63 °F) and compressor speed (100%) fixed, the predicted total capacity and predicted total power consumption are shown for varying evaporator air flow rates and outdoor air dry bulb temperatures..... 71
Figure 3.7. With return air wet bulb (63 °F) and compressor speed (100%) fixed, the predicted *COP* and predicted *SHR* are shown for varying evaporator air flow rates and outdoor air dry bulb temperatures..... 72
Figure 3.8. With outdoor air temperature (95 °F) and evaporator air flow rate (100%) fixed, the predicted total capacity and predicted total power consumption are shown for varying entering evaporator air wet bulbs and compressor speeds..... 73
Figure 3.9. With outdoor air temperature (95 °F) and evaporator air flow rate (100%) fixed, the predicted *COP* and predicted *SHR* are shown for varying entering evaporator air wet bulbs and compressor speeds..... 74
Figure 3.10. Comparison of simulated improper refrigerant charge level fault impacts for total cooling capacity on systems with different expansion devices..... 79
Figure 3.11. Comparison of simulated condenser fouling fault impacts on total cooling capacity for systems with different expansion devices..... 80
Figure 3.12. Comparison of simulated evaporator fouling fault impacts on total cooling capacity for systems with different expansion devices..... 81
Figure 3.13. Comparison of simulated improper refrigerant charge level fault impacts on *COP* for systems with different expansion devices..... 83

Figure 3.14. Comparison of simulated condenser fouling fault impacts on *COP* for systems with different expansion devices..... 84
Figure 3.15. Comparison of simulated evaporator fouling fault impacts on *COP* for systems with different expansion devices..... 85
Figure 3.16. Comparison of simulated improper refrigerant charge level fault impacts on *SHR* for systems with different expansion devices..... 87
Figure 3.17. Comparison of simulated condenser fouling fault impacts on *SHR* for systems with different expansion devices..... 88
Figure 3.18. Comparison of simulated evaporator fouling fault impacts on *SHR* for systems with different expansion devices..... 89
Figure 3.19. Comparison of simulated improper refrigerant charge level fault impacts on run-time requirement to meet equivalent loads for systems with different expansion devices..... 91
Figure 3.20. Comparison of simulated condenser fouling fault impacts on run-time requirement to meet equivalent loads for systems with different expansion devices..... 92
Figure 3.21. Comparison of simulated evaporator fouling fault impacts on run-time requirement to meet equivalent loads for systems with different expansion devices..... 93
Figure 3.22. Comparison of simulated improper refrigerant charge level fault impacts on energy consumption to meet equivalent loads for systems with different expansion devices..... 95
Figure 3.23. Comparison of simulated condenser fouling fault impacts on energy consumption to meet equivalent loads for systems with different expansion devices..... 96
Figure 3.24. Comparison of simulated evaporator fouling fault impacts on energy consumption to meet equivalent loads for systems with different expansion devices..... 97
Figure 3.25. Accuracy of models used to predict normal total cooling capacity and COP against neural network meta-model generated outputs..... 102

Figure 3.26. Accuracy of simplified charge fault impact models for predicting total cooling capacity and COP of systems with charge faults..... 103
Figure 3.27. Accuracy of simplified condenser fouling fault impact models for predicting total cooling capacity and COP of systems with condenser fouling faults..... 104
Figure 3.28. Accuracy of simplified evaporator fouling (reductions in evaporator air flow) fault impact models for predicting total cooling capacity and COP of systems with evaporator fouling faults..... 105
Figure 3.29. Accuracy of simplified fault impact models for predicting total cooling capacity and COP of systems with combinations of refrigerant charge and condenser fouling faults..... 106
Figure 3.30. Accuracy of simplified fault impact models for predicting total cooling capacity and COP of systems with combinations of refrigerant charge and evaporator fouling faults..... 107
Figure 3.31. Accuracy of simplified fault impact models for predicting total cooling capacity and COP of systems with combinations of refrigerant charge, condenser fouling, and evaporator fouling faults..... 108
Figure 4.1. Overview of supervisory learning process used to develop fault impact meta-model using detailed fault impact model inputs and outputs..... 113
Figure 3.2. Prediction accuracy of semi-empirical model for total power consumption of a variable-speed RTU operating without faults over a range of ambient conditions and loads..... 114
Figure 4.3. Overall comparison of suction pressure and enthalpy predictions from fault impact meta-model and detailed model for system with fixed orifice expansion valve..... 116
Figure 4.4. Overall comparison of discharge pressure and enthalpy predictions from fault impact meta-model and detailed model for system with fixed orifice expansion valve. 117
Figure 4.5. Overall comparison of liquid line (condenser outlet) pressure and enthalpy predictions from fault impact meta-model and detailed model for system with fixed orifice expansion valve..... 118

Figure 4.6. Overall comparison of evaporator inlet pressure and enthalpy predictions from fault impact meta-model and detailed model for system with fixed orifice expansion valve..... 119
Figure 4.7. Modeled total cooling capacity, cycle *COP*, and *SHR* for RTU with fixed orifice expansion valve at different levels of refrigerant charge..... 120
Figure 4.8. Comparison of modeled total cooling capacity, cycle COP, and SHR for RTU with fixed orifice expansion valve over a range of condenser fouling levels..... 121
Figure 4.9. Comparison of modeled total cooling capacity, cycle COP, and SHR for RTU with fixed orifice expansion valve over a range of evaporator fouling levels..... 122
Figure 4.10. Comparison of modeled total cooling capacity, cycle COP, and SHR for RTU with fixed orifice expansion valve over a range of compressor valve leakage levels..... 124
Figure 5.1. Hourly ambient temperatures and humidity for Atlanta, GA TMY3 data set for 72 hours before and 72 hours after peak dry bulb temperature..... 132
Figure 5.2. Hourly ambient temperatures and humidity for Chicago, IL TMY3 data set for 72 hours before and 72 hours after peak dry bulb temperature..... 133
Figure 5.3. Hourly ambient temperatures and humidity for Miami, FL TMY3 data set for 72 hours before and 72 hours after peak dry bulb temperature..... 134
Figure 6.1. Optimal service decision policy for refrigerant charge faults in Miami, FL for system that is 20% oversized to maximum annual load ($f_{os} = 0.2$) and refrigerant leakage rate equal to 5% per year..... 149
Figure 6.2 Comparison of optimal decision boundaries for refrigerant charge faults in Miami, FL for system that is 20% oversized to maximum annual load ($f_{os} = 0.2$) and fault leakage rate equal to 5% and 15% per year..... 152
Figure 6.3. The optimal service schedule for Atlanta, GA when the system is assumed be initially fully charged. When service is performed, the charge level is returned to 100% (20% oversized to maximum annual load; fault leakage rate equal to 5% per year)..... 153

Figure 6.4. An example optimal service schedule for Chicago, GA when the system is assumed is initially fully charged. When service is performed, the charge level is returned to 100% (20% oversized to maximum annual load; fault leakage rate equal to 5% per year)..... 154

Figure 6.5. An optimal service schedule for Miami, FL when the system starts with a clean coil (no reduction in condenser airflow). When service is performed, the condenser airflow level is returned to 100%, however a service cost is incurred (20% oversized to maximum annual load; condenser fouling rate equal to 5% per 5000 hours of run-time)..... 154

Figure 6.6. An optimal service schedule for Atlanta, GA when the system starts with a clean condenser coil (no reduction in condenser airflow). When service is performed, the condenser airflow level is returned to 100%, however a service cost is incurred (20% oversized to maximum annual load; condenser fouling rate equal to 5% per 5000 hours of run-time)..... 157

Figure 6.7. An optimal evaporator cleaning service schedule for Miami, FL when the system starts with a clean evaporator coil (no reduction in evaporator airflow). When service is performed, the evaporator airflow level is returned to 100%, however a service cost is incurred (20% oversized to maximum annual load; evaporator fouling rate equal to 5% per 5000 hours of run-time)..... 158

Figure 6.8. An optimal evaporator cleaning service schedule for Chicago, IL when the system starts with a clean evaporator coil (no reduction in evaporator airflow). When service is performed, the evaporator airflow level is returned to 100%, however a service cost is incurred..... 160

Figure 6.9. Optimal service schedule for system located in Miami, FL with multiple faults (described in Table 6.2). The optimal schedule tends to group multiple service tasks at each service interval..... 162

Figure 6.10. Optimal service schedule for system located in Miami, FL with multiple faults (described in Table 6.3). The optimal schedule tends to group multiple service tasks at each service interval..... 163

Figure 6.11. Comparison of lifetime refrigerant charge levels for system located in Miami, FL with a leakage rate of 5% percent per year using biennial and annual periodic service schedules..... 166
Figure 6.12. Comparison of average annual operating costs for system located in Miami, FL with different refrigerant charge leakage rates using biennial and annual periodic service schedules..... 167
Figure 6.13. Additional lifetime operating costs comparison between annual and biannual service schedules for building located in Miami, FL with different refrigerant leakage rates..... 168
Figure 6.14. Comparison of lifetime refrigerant charge levels for system located in Chicago, IL with a leakage rate of 5% per year using biennial and annual periodic service schedules..... 169
Figure 6.15. Lifetime average annual operating cost comparison between annual and biannual service schedules for building located in Chicago, IL with different refrigerant charge leakage rates..... 170
Figure 6.16. Additional lifetime operating costs comparison between annual and biannual service schedules for building located in Chicago, IL with different refrigerant leakage rates..... 171
Figure 6.17. Comparison of lifetime condenser fouling levels for system located in Miami, FL with a condenser fouling rate of rate of 5% percent per 5000 hours of condenser fan runtime using biennial and annual periodic service schedules..... 172
Figure 6.18. Comparison of average annual operating costs for system located in Miami, FL with different condenser fouling rates using biennial and annual periodic service schedules..... 173
Figure 6.19. Additional lifetime operating costs comparison between annual and biannual service schedules for building located in Miami, FL with different condenser fouling rates..... 174
Figure 6.20. Comparison of lifetime condenser airflow levels for system located in Chicago, IL with a condenser fouling rate of 5% 5000 hours of condenser fan runtime using biennial and annual periodic service schedules..... 175

Figure 6.21. Lifetime average annual operating cost comparison between annual and biannual service schedules for building located in Chicago, IL with different condenser fouling rates..... 176
Figure 6.22. Additional lifetime operating costs comparison between annual and biannual service schedules for building located in Chicago, IL with different condenser fouling rates..... 177
Figure 6.23. Trended refrigerant charge for an air conditioner with refrigerant leaking at a rate of 5% per year located in Miami, FL over life of equipment when an emergency service policy is used. Service is performed with insufficient capacity causes comfort violations..... 179
Figure 6.24. Comparison of average annual service costs for system in Miami, FL with different refrigerant charge leakage rates using optimal and emergency service policies..... 181
Figure 6.25. Trended refrigerant charge for an air conditioner with refrigerant leaking at a rate of 5% per year located in Chicago, IL over life of equipment when an emergency service policy is used. Service is performed with insufficient capacity causes comfort violations..... 181
Figure 6.26. Comparison of average annual service costs for system in Chicago, IL with different refrigerant charge leakage rates using optimal and emergency service policies..... 182
Figure 6.27. Distribution of randomly sampled fault rates used to evaluate the performance of different service decision strategies for multiple fault scenarios..... 192
Figure 6.28. Additional lifetime operating costs relative to optimal lifetime costs when the simplified service decision methodology developed by Rossi and Braun is used to service multiple faults in three different locations..... 194
Figure 6.29. Additional lifetime operating costs relative to optimal lifetime costs when the modified service decision methodology that considers multiple service tasks is used to service multiple faults in three different locations..... 195
Figure 6.30. Additional lifetime operating costs relative to optimal lifetime costs when annual service schedules are used to service multiple faults in three different locations..... 196

Figure 6.31. Additional lifetime operating costs relative to optimal lifetime costs when biennial service schedules are used to service multiple faults in three different locations..... 197
Figure 6.32. Additional lifetime operating costs relative to optimal lifetime costs when emergency service schedules are used to service multiple faults in three different locations..... 198
Figure 6.33. Additional lifetime operating costs relative to optimal lifetime costs when service is performed when capacity is degraded by 10% considering multiple faults in three different locations..... 199
Figure 6.34. Additional lifetime operating costs relative to optimal lifetime costs when service is performed when COP is degraded by 10% considering multiple faults in three different locations..... 200
Figure 6.35. Additional lifetime operating costs relative to optimal lifetime costs when service is performed when energy consumption is increased by 10% considering multiple faults in three different locations..... 201

LIST OF SYMBOLS

A	area, or set of service actions
a_i	service action
$a, b, c...$	empirical parameters
B	wet bulb temperature
BF	bypass factor
C_e	cost per unit electricity
C_s	cost per service action
COP	coefficient of performance
$c_{p,a}$	isobaric specific heat capacity
EC	equipment costs
F	dynamic programming cost function
FIR	fault impact ratio
FXO	fixed orifice expansion valve
f	correction factor or oversizing factor
g	correction factor
h	specific enthalpy
J	optimization variable
m	mass
$\dot{m}$	mass flow rate
NTU	number of transfer units
OC	operating cost
P	pressure
$\dot{Q}$	heat transfer rate
SC	service costs
SHR	sensible heat ratio
T	temperature, or dynamic programming state transition function

ΔT	temperature difference
TXV	thermostatic expansion valve
t	time
Δt	time difference
UA	overall heat transfer conductance
UC	utility costs
$\dot{V}$	volumetric flow rate
x	thermodynamic equilibrium quality, or fault level, or dynamic programming state
W	electrical energy consumption
$\dot{W}$	electrical power
ΔW	energy difference
α	heat loss fraction
β	discounting factor
ε	heat transfer effectiveness
ρ	density
ϕ	relative humidity
τ	time constant
Ω	speed/frequency
ω	humidity ratio, or service decision

Subscript Abbreviations

ADP	apparatus dew point
c, foul	condenser fouling
ca	condenser air
cai	condenser air inlet
chrg	refrigerant charge
cao	condenser air outlet
comp	compressor
cool	cooling capacity

cri	condenser refrigerant inlet
cro	condenser refrigerant outlet
dis	compressor discharge
e, foul	evaporator fouling
ea	evaporator air
eai	evaporator air inlet
eao	evaporator air outlet
elec	electricity
eri	evaporator refrigerant inlet
ero	evaporator refrigerant outlet
ID	indoor
idf	indoor fan
llo	liquid-line outlet
OD	outdoor
oa	outdoor air
odf	outdoor fan
os	oversize
ra	return air
run	run-time
sa	supply air
r	refrigerant
ref	refrigerant
s	sensible
suc	compressor suction
sh	superheat
sc	subcooling

ABSTRACT

Author: Hjortland, Andrew, L. PhD
Institution: Purdue University
Degree Received: May 2018
Title: Automated Fault Detection, Diagnostics, Impact Evaluation, and Service Decision-Making for Direct Expansion Air Conditioners
Major Professor: James E. Braun

This work describes approaches for automatically detecting, diagnosis, and evaluating the impacts of common faults in unitary rooftop air conditioning equipment. A semi-empirical component-based modeling approach using virtual sensors has been implemented using low-cost microcontrollers and tested on fixed-speed and variable-speed equipment using laboratory psychrometric test chambers. A previously developed virtual refrigerant charge sensor was applied to a fixed-speed rooftop unit with combinations of condenser types and expansion valve types and resulted in average prediction errors less than 10%. In addition, a methodology was developed that can be used to tune the empirical parameters of the model using data collected without psychrometric chambers, greatly reducing the experimental effort and costs required for the model. Virtual sensors previously developed for fixed-speed systems were also implemented for a variable-speed rooftop unit without significant loss of accuracy.

Much of this work has been devoted to estimating the performance impacts of faults that grow over time, like heat exchanger fouling or refrigerant charge leakage. To estimate these impacts, semi-empirical models for predicting the normal performance of fixed-speed and variable-speed systems have been developed and evaluated using experimentally collected data. In addition, the virtual sensor approaches for estimating the actual performance of systems using low-cost sensor measurements were evaluated. Together, normal performance models and virtual sensor estimations were used to estimate the overall impacts of several faults on system

performance. A methodology for quantifying the performance impacts of simultaneously occurring faults has been developed and tested using a detailed system model and experimental results. While relatively simple, simulated and experimentally collected results showed the fault impact models were accurate within 10% of the actual fault impacts. The fault impact evaluation models could be embedded in an AFDD system and used to determine when performance degradation faults should be serviced from an operating cost perspective.

In addition, different service and maintenance strategies are compared in this work using a simulation environment that was developed. A data-driven artificial neural network model of a rooftop unit with faults has been derived for this purpose using a detailed fault impact model for direct expansion cooling equipment. This model was coupled with a building model to simulate operating cost impacts of performance degradations and service over the life of cooling equipment. An optimization problem was formulated with the goal to minimize lifetime energy and service costs and was solved using dynamic programming. Using the optimal solution as a baseline, suboptimal service decision-making strategies were implemented and simulated using the building model. It was found that condition-based maintenance strategies using the outputs of automated fault detection and diagnostics tools can significantly reduce lifetime operating costs over periodic service policies.

1. INTRODUCTION

1.1 Background and Motivation

With advancements in low-cost hardware and cloud-based computing and communications infrastructure, machines that are used every day to operate our world and complete daily tasks are becoming more and more sophisticated and efficient. Machines in today's world are becoming increasingly connected with their surroundings. For example, new thermostats can "learn" the behavior of homeowners so that energy can be saved when they are away. Self-driving cars can prevent traffic accidents and other problems by optimizing their control. Large online retailers can deliver almost any item to customers within a day – no matter where the customer lives. The world and its machines are becoming increasingly optimized to meet the needs of its people.

Research and engineering for the optimization of machines and systems for critical applications has existed for many years. Controls and operations research related to complex chemical manufacturing, space exploration, military and defense applications, and power generation facilities has advanced considerably. In these applications, the complexity of the systems and decisions involved is often too great for the typical human to understand. Additionally, for some applications the speed at which decisions must be made is only suitable for computers. When this is the case, supervisory control systems must be designed to ensure safe and reliable operations.

Supervisory control systems have also been developed for several non-critical applications. Some of these applications are related to building systems and their heating, ventilation, and air conditioning (HVAC) systems. The benefits of improving the control of HVAC systems in

buildings come from improvements in reliability and comfort. Additionally, improvements in system control can lead to increases in efficiency and reductions in operating costs.

One supervisory-level process monitoring system intended to reduce operating costs in HVAC applications is automated fault detection and diagnostics (AFDD). AFDD systems are used to identify deviations in normal or expected performance that are due to faults or other unexpected failures that arise throughout the life of the system. The advantage of using AFDD systems for this purpose is that faults causing less efficient operation or reductions in capacity can be found more quickly than with a HVAC technician. Much like automotive diagnostics in today's cars, AFDD for HVAC systems can make it easier to identify problems and prevent more costly problems through preventative or proactive maintenance practices.

The general approach for HVAC system AFDD described by Rossi and Braun has been summarized in Figure 1.1 [1]. In the first stage of a fault detection and diagnostics tool, measurements recorded from the system of interest are made by a data acquisition system. These measurements will be used to determine whether the system is operating normally and to inform service decisions if a fault exists. Traditionally, the data collected from HVAC systems are temperature, pressure, and relative humidity measurements that can be used to fix the thermodynamic state points around the system. Once these state points are known, additional properties such as enthalpy or density can be determined and used to assess the condition of the system.

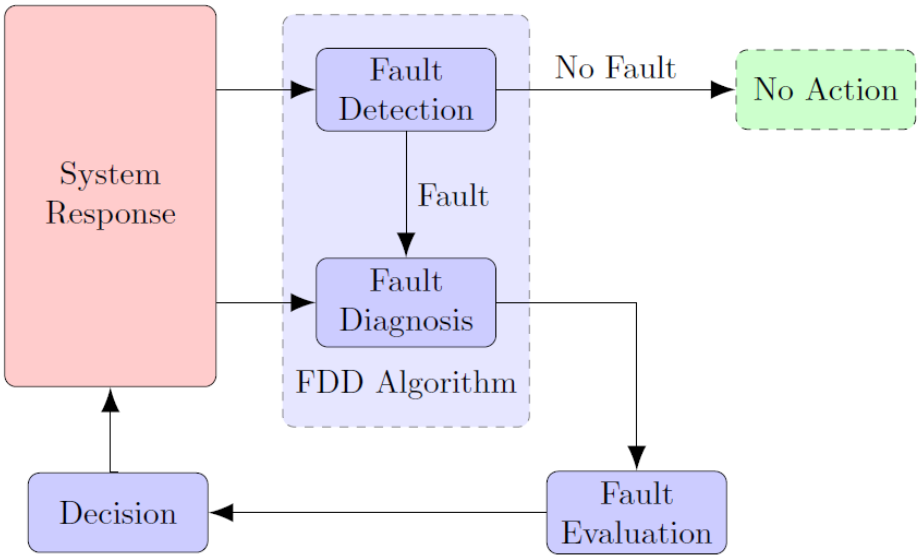

Figure 1.1. General approach for automated fault detection and diagnostics described by Rossi and Braun [1].

Sensor and hardware requirements for data acquisition have consistently been a significant upfront cost that has slowed widespread introduction of AFDD systems on lower cost commercial and residential direct-expansion (DX) HVAC equipment. Traditionally, these systems have been manufactured and sold with none or very few sensors installed for control purposes. The chief exceptions to this are the air-side temperature and relative humidity sensors installed with outdoor-air economizers for commercial applications. Because of the limited number of sensors available on existing systems, sensors that are required for AFDD systems should not lead to significant cost increases.

To reduce upfront instrumentation costs, past research has focused on designing low-cost AFDD systems and on the development of alternative methods for estimating physical quantities of interest. These methods, called *virtual sensors*, use a combination of low-cost sensors with physical or mathematical models of the system to estimate quantities that are impossible or expensive to measure directly [2-6]. In the current work, virtual sensors that have been developed in the past have been implemented using low-cost hardware and software and applied

to a two-stage fixed-speed commercial rooftop air conditioner (RTU) to demonstrate and assess the virtual sensor and AFDD algorithms. Additionally, these same virtual sensors have been applied to a variable-speed RTU with small modifications.

Using measured quantities and virtual sensors, automated fault detection for RTUs can be performed. In this step, a determination is made whether the system is operating normally or if a fault has developed. Many methods have been developed for this task including system-level model-based statistical classifiers, component-level virtual sensor methods. A thorough review of these methods and others have been discussed in the literature review of Chapter 2. What is common to all these methods is that a comparison between the performance that is expected and the observation performance at the current operating condition is made. When a sufficient deviation exists, a fault is identified.

In some systems and approaches, fault detection is the last step – i.e. the system only classifies if the machine is normal or not. In other, more developed systems, the step that occurs after fault detection is fault diagnosis. Unlike fault detection, the goal of fault diagnosis is to narrow down or isolate the cause of the fault. This can be done using the information available, including virtual sensors. Common examples of fault diagnosis in DX equipment is distinguishing between systems that are low on refrigerant from those that have dirty condenser coils. Fault diagnosis is in many ways an extension of fault detection – rather than classify if the system is not normal, diagnosis intends to classify if a system has a specific fault. In this work, the effectiveness of the AFDD system implemented will be analyzed using experimental data collected from systems with faults introduced artificially.

The focus of this research is related to the steps that come after a fault detection and diagnosis have been performed: fault impact evaluation. Because of the sensitivity of some

FDD methodologies, faults with relatively minor impacts on system performance can be detected and diagnosed. While it is not wrong that FDD tools can detect faults with minor impacts, a determination about the extent of the fault impact must be made before a decision can be made about whether to fix a fault. To this end, fault impact evaluation approaches are applied to the outputs of the FDD system to estimate how much a fault has impacted system performance. Impacts that are considered in this work include total cooling capacity, efficiency, sensible heat ratio, and energy consumption. While much attention and research focus has been applied to developing novel low-cost FDD approaches, less focus has been devoted to developing generalizable methods for estimating impacts of faults. In this work, a methodology that can be used to estimate the impacts of multiple faults occurring simultaneously has been developed and tested using detailed equipment simulation models and experimental data.

After faults have been detected, diagnosed, and evaluated, the information generated by each process can be used to improve service decisions. Not all faults may warrant service; faults may have limited performance impacts or may be very costly to fix. The economic tradeoffs between letting a fault persist and servicing it must be weighed before action is undertaken. Other considerations must also be made: will the system be able to maintain comfortable conditions if the fault persists, will the expected life of the system be shortened, or will additional faults exacerbate the impacts are all questions that could be considered before recommending service. Developing more optimal decision-making methodologies by coalescing data and other information collected during system operation is the last goal of this research.

AFDD systems are believed to have significant energy and operational cost benefits when they are applied to HVAC systems [7-9]. While many of these technologies have been researched and developed for small-to-midsize commercial buildings and residential applications,

few products are available from equipment manufacturers. While federal regulations have been used in the past to encourage or require these products, uncertainty about the actual benefit of AFDD to customers and building owners still exist [9-11]. To address some of these uncertainties, a computerized simulation framework has been developed in this work with the goal of assessing the cost effectiveness of different maintenance strategies. Simple strategies like performing annual maintenance are compared with more advanced conditioned-based maintenance strategies using the simulation framework. The comparisons made with this tool are used to develop deeper understanding of the benefits for HVAC AFDD systems and proactive maintenance.

1.2 Research Objectives

The goal of this thesis is to develop improved fault detection, diagnosis, evaluation, and service decision-making strategies for DX air conditioning equipment. The research objectives can be summarized into three main objectives, each covered in this thesis.

1. The first research objective is to evaluate implementation of a previously developed fault detection and diagnostics algorithm based on virtual sensors using low-cost sensor and data acquisition hardware. The system was tested on a two-stage fixed-speed RTU commonly used in small-to-midsize commercial buildings. The system was also tested using a similar RTU with a variable-speed compressor and control system. Variable-speed systems offer significant performance improvements, especially under part-load conditions. Evaluation of FDD systems applied to variable-speed systems has been limited.
2. The second objective of this research is to develop improved models for estimating the performance and economic impacts of different faults that could implemented within an

AFDD system. For this purpose, data collected from psychrometric chamber testing has been used to tune models for normal performance and measure degradation in cooling capacity, efficiency and sensible heat ratio with different combinations of faults imposed. A major thrust of this work is to identify methodologies and models that can be used to estimate the relative impacts of multiple faults that occur simultaneously. Isolating the impacts of different faults is important when making decisions about which faults to fix when multiple faults are present.

3. A simulation environment to evaluate different maintenance and service strategies has been implemented. Using simple building load models and fast steady-state equipment performance models, the energy consumption of systems with and without faults has be studied. An optimal maintenance decision-making strategy has been developed and implemented within the environment. Additional comparison between the optimal maintenance strategy and other sub-optimal strategies are presented.

1.3 Thesis Organization

This chapter has given a broad overview of automated fault detection and diagnostics (AFDD) systems as well as an outline of the research objectives of the completed work.

Chapter 2 describes the virtual sensors and AFDD methodologies developed for RTUs, including a literature review of past work by other authors. After describing the virtual sensors and AFDD methodology, the design and implementation of an example low-cost system for RTU AFDD is discussed. Additionally, the performance of the AFDD methodologies is evaluated using a two-stage fixed-speed RTU tested in psychrometric chambers with different faults imposed. The virtual sensors are also applied to a variable-speed RTU and the

performance of the methodology is evaluated using data collected from a system tested in psychrometric chambers.

Chapter 3 presents a literature review of previous work on fault impact estimation and presents a new approach for estimating the impacts of multiple simultaneous faults. The methodology developed requires estimations of the normal performance of a system without faults. To do this, a semi-empirical model is discussed and tested using experimental data collected using the variable-speed RTU. Using virtual sensors, processes that can be used to estimate the observed performance of an RTU are also described. Following this description, a fault impact equipment simulation model is used to develop simplified models for isolating the impacts of different faults when multiple faults are present.

Chapter 4 describes a methodology for approximating detailed equipment simulation models using artificial neural network models. These models are an integral part of simulation platform presented in Chapter 5. The process used to train the neural network meta-models are described and example results are presented. Additionally, the simulated impacts of different faults for a RTU are described and analyzed over different operating conditions.

Chapter 5 describes a computer model and framework for simulating the long-term performance impacts of DX equipment faults. The computer model implements the interaction between the air conditioning system with a simple single-node building load model. The weather data and operating cost calculations are described. Additionally, models that simulate how faults change with time and equipment run-time are described and implemented.

Chapter 6 introduces the optimal maintenance scheduling problem and uses the simulation framework to determine optimal maintenance strategies in different applications. Additionally, other suboptimal maintenance strategies are described and implemented within the simulation

framework for comparison. These strategies include periodic maintenance policies, conditioned-based maintenance policies, and policies based on comfort violations.

Finally, the work presented in this thesis is summarized in Chapter 7 and recommendations for further study are discussed.

2. AUTOMATED FAULT DETECTION AND DIAGNOSTICS FOR UNITARY AIR CONDITIONING EQUIPMENT

2.1 Background and Motivation

Studies have shown that packaged commercial rooftop air conditioning equipment (commonly called rooftop units, RTUs) tend to be poorly maintained and significant energy may be wasted due to unnoticed or unrepaired equipment faults. One often cited review of FDD technologies for building systems estimates between 10–30% additional annual energy in US commercial buildings may be caused by repairable faults [7, 8]. While uncertainty about the prevalence of different faults still exists, previous field studies have shown that RTUs often have undiagnosed faults and may benefit from automated FDD more than other building systems [12, 13]. Moreover, economic and statistical analyses by Li and Braun and Yuill and Braun estimate that FDD can provide positive economic value by reducing operating costs and improving in-field equipment repair [9, 11, 14]. The potential energy savings and efficiency gains that FDD may provide has garnered continued academic and private-sector research interest in equipment and building system automated FDD.

Manufacturers of RTUs have been slow to translate FDD research and technologies into new products and offerings. This may be attributed to a few issues that remain, to some extent, unsolved:

- FDD systems must be low-cost, easy to manufacture, and install,
- uncertainty with respect to economic benefits and potential savings still exists, especially regarding prevalence of different faults,
- and there has been a lack of integration and interoperability with other building technologies.

To address these issues, a complete automated FDD system has been designed that implements a previously developed automated FDD approach using virtual sensors [2-4, 6, 15, 16]. The system can continuously monitor the performance of an RTU and diagnose common faults during quasi-steady-state operation. An advantage of the selected methodology where many other FDD approaches may fail or have inconsistent results is the ability to diagnose multiple simultaneously occurring faults. Due to the sometimes-long service intervals for RTUs, multiple faults may develop over time and a well performing FDD system must be designed to handle these situations. In addition to multiple simultaneous fault diagnosis, an approach to estimate the total impact different faults have on equipment performance has been implemented. This approach uses a combination of virtual sensors to monitor actual performance and reference model outputs to estimate normal performance at similar operating conditions to estimate the severity of faults. By estimating the severity of faults on equipment performance, faults with inconsequential impacts with little benefits from servicing may be ignored.

2.2 Literature Review of Previous Work on AFDD for DX Systems

Rossi and Braun developed a statistical ruled-based FDD approach using a steady-state vapor compression cycle model for an RTU with fixed-speed compressor and fixed orifice expansion device [1]. The methodology compared the outputs generated by the cycle model to nine temperatures and one relative humidity measurement installed on the RTU to classify the performance as normal or faulty. For faulty classifications, the differences between the model outputs and measurements, termed residuals, were used to classify the specific type of fault. The method assumed that each fault type considered (improper refrigerant levels, condenser fouling, evaporator fouling, compressor valve leakage, and liquid line restrictions) results in a unique residual response. For single fault cases, this methodology was later demonstrated to

successfully diagnose different faults at several operating conditions using laboratory data by Breuker and Braun [17]. A sensitivity analysis of the statistical threshold for steady-state detection, fault detection, and fault diagnosis was also discussed.

Chen and Braun adapted the statistical rule-based FDD methodology originally proposed by Rossi and Braun to RTUs having thermostatic expansion valves [1, 18]. The researchers used steady-state laboratory test data to tune the thresholds of the methodology and assess the resulting performance of the system at diagnosing several common faults. The approach developed in this work reduced the number of sensors used to six temperature sensors and one relative humidity sensor. Rather than computing residuals, a methodology to calculate "sensitivity ratios" that were sensitive to individual faults was described. The advantage of this methodology was the improved FDD performance over a larger range of equipment operating conditions.

Siegal and Wray developed a methodology to detect refrigerant charging faults based on measured compressor superheat and outdoor air temperature for split-type air conditioning systems with fixed orifice expansion devices [19]. The methodology was tested using four different split-type air conditioning systems and compared with other commercially available refrigerant charge diagnostics methods. The methodology proposed in this work loses accuracy for different evaporator air inlet conditions and air flow rates.

Armstrong *et al*. developed a "non-intrusive load monitoring" method to diagnose common RTU faults based on transient electrical power signatures of the system [20]. The methodology requires three voltage measurements to be installed on the RTU. During start-up operation, the transient response of the voltage measurements was used to diagnose refrigerant slugging in the compressor, compressor valve leakage, and undercharged faults.

Kim and Payne developed a rule-based FDD methodology for a split-type air conditioner with fixed-speed compressor and thermostatic expansion valve [21]. The methodology identified several performance features that are sensitive to different faults. The directional changes between the expected features and the features observed when faults are present were used to develop a residual based fault classifier.

Li and Braun present a significant improvement to previously developed FDD methodologies for RTUs based on "virtual sensors" [4-6, 16, 22]. Virtual sensors are mathematical or physical models that use low-cost measurement to estimate quantities that are expensive or impossible to measure directly. The virtual sensors developed in the work are sensitive to specific faults, enabling the detection and diagnosis of multiple faults that occur simultaneously. The methodology used virtual sensors for detecting improper refrigerant charge levels, condenser fouling, evaporator fouling, compressor valve leakage, and liquid line restrictions.

Kim and Braun evaluated the performance of the virtual sensor methodology applied to several different units in laboratory settings and in the field [2, 3, 15, 23]. Additionally, the methodology was extended for a mini-split heat pump system with variable-speed compressor and electronic expansion valve. While many of the virtual sensors were improved, additional improvements to the virtual refrigerant charge sensor were suggested, especially for variable-speed heat pump systems with an accumulator. Additionally, the authors note that while the virtual sensors are decoupled, enabling simultaneous fault diagnosis, some of the virtual sensors lose accuracy for low and high levels of refrigerant charge. In these cases, suction superheat and liquid line subcooling tend to zero and make refrigerant property estimation error prone.

In addition to the previous works summarized, Katipamula & Brambley and Kim & Katipamula provide a literature review of several data-driven approaches for RTU and split-type

air conditioner AFDD that have been developed [7, 8, 24]. While the details of these methods differ, they all use essentially a black box model that is tuned using experimental data collected from systems subjected to different faults. Model types that have been studied include principle component analysis (PCA), artificial neural networks, and binary decision trees. These methodologies have the potential to enable very accurate fault detection and diagnosis capabilities, however they tend to yield unreliable results for conditions not contained in the original training data set. Additionally, these data driven methods can suffer from being overfit to the training data set which can result in poor performance and false alarms.

2.3 Description of AFDD System for Multi-stage and Variable-Speed RTUs

Previous work on automated fault detection and diagnostics (AFDD) for commercial and residential direct expansion (DX) air conditioners and heat pumps has predominantly been applied to fixed-speed or multistage equipment [1, 4, 6, 10, 15, 17-20, 25, 26]. Historically, fixed-speed and multistage air conditioning systems have covered many manufacturers' sales. Because of this, research and development has focused on AFDD for these systems since there is the largest potential impact at reducing energy in commercial and residential buildings. To a lesser extent, regulatory measures have also driven investments in AFDD for fixed and multistage equipment. The most notable of these regulations being the outdoor air economizer FDD requirement on commercial air conditioning equipment in California [27].

Even with these developments and advances towards autonomous FDD systems for air conditioning equipment, manufacturers' have been slow to introduce these technologies into new products. Manufacturers that have introduced FDD technologies have most often limited them to a single model as an optional add-on feature. To address these problems, a low-cost AFDD system for a multi-stage RTUs has been developed. One goal of this work was to demonstrate

development of an AFDD system for RTUs to make the engineering of such a system easier for equipment manufacturers. The ultimate goal of this work is to encourage equipment manufacturers to incorporate fault diagnostics technologies into new products by reducing the engineering effort.

2.3.1 Overview of AFDD System Architecture

Two hurdles that limit the adoption of FDD as a retrofit for existing RTUs and for integration in new RTUs are the lack of sensors and peripheral communications installed in equipment. It is difficult to detect problems with the data that is typically available. Additionally, it is difficult to alert building owners and facility managers about problems in a timely and effective manner. A dedicated FDD module has been developed that can be integrated within or in parallel with existing RTU controllers. The system addresses the need for sensors by adding low-cost measurement electronics and enables external communications by adding standard internet connectivity hardware. A schematic outline of the system architecture in relation to a typical RTU controller is shown in Figure 2.1.

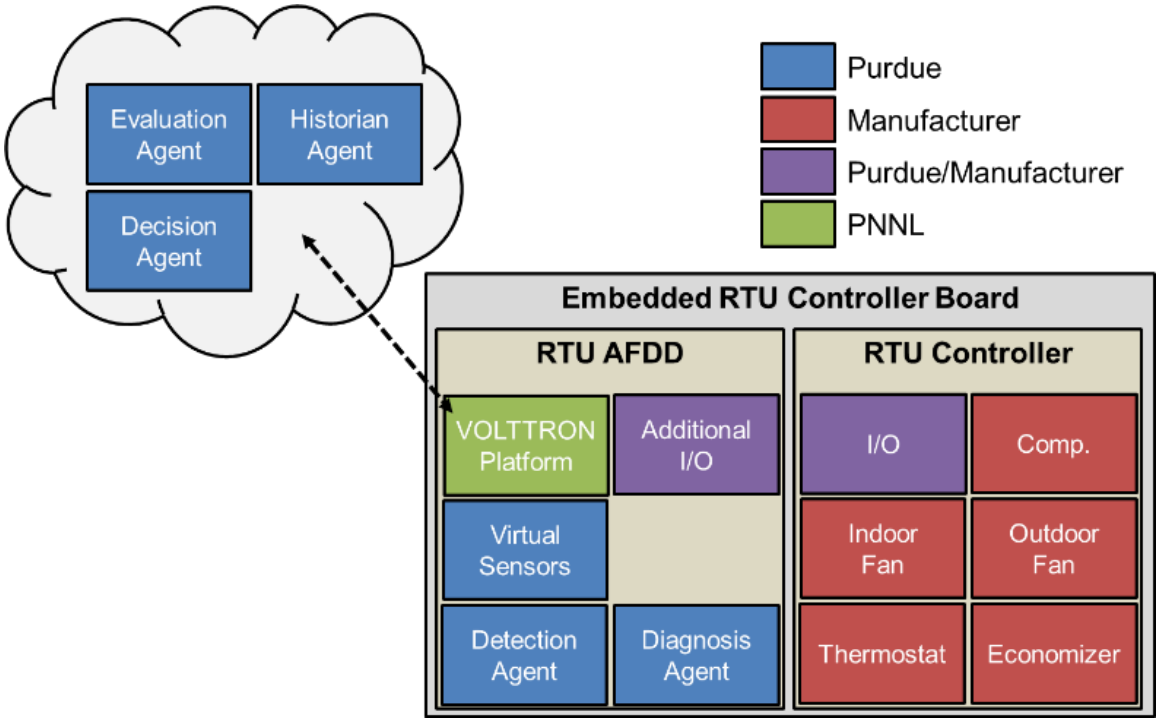

Figure 2.1. Overall architecture of VOLTTRON™ enabled RTU AFDD system as an additional module of current RTU control units.

It should be noted that integrating a dedicated FDD module into existing RTU controllers is not the only possible solution for detecting and diagnosing equipment faults. An alternative solution, commonly called in-field FDD, uses devices that can be temporarily installed on equipment to aid field technicians in determining problems. A significant advantage of these diagnostics tools is that the cost can be distributed across many different RTUs since they can be generally applied. However, designing a tool that can be generally applied to many different RTUs while still having sufficient fault detection sensitivity and false alarm rejection is challenging [10]. In addition, in-field FDD requires inspection of actual RTUs to detect and diagnose faults but previous studies have concluded that RTU inspection may be infrequent [12, 13]. Finally, the cost of sending service technicians into the field is significant in comparison with monitoring the health of equipment using electronics.

From its conception, the automated FDD (AFDD) system has been developed with the potential for existing as a passive, standalone system within the RTU controller. This means it was essential to incorporate all basic components of an AFDD product: a sensor measurement and data acquisition (DAQ) module, a fault detection module, and a fault diagnosis module. This involved both hardware design and extensive software development to implement the AFDD methods and algorithms. Rather than using propriety tools and code, the RTU AFDD system leverages open-source software exclusively, most importantly the VOLTTRON™ monitoring and control platform [28, 29]. The VOLTTRON™ platform, developed by the US Department of Energy, facilitates communication between software agents and other physical devices and resources with a primary focus on building systems and technologies. The role VOLTTRON™ plays in the RTU AFDD system is to act as a message broker between AFDD services so that development, operation, and management may be organized hierarchically. This modularity offers the flexibility to modify, add, or completely swap out individual subcomponents depending on system requirements and RTU configuration.

2.3.2 Description of Virtual Sensors used in AFDD Approach

Many automated FDD approaches have been previously proposed for RTUs [1, 3, 4, 6, 15, 18, 20, 30-33]. The system described in this work implements a virtual sensor-based methodology originally described by Li and Braun [4, 6, 16]. A virtual sensor uses low-cost sensor measurements with a mathematical model to estimate quantities that are difficult or expensive to measure directly. Several virtual sensors developed previously were implemented that can be used to detect and diagnose common faults, including:

- improper refrigerant charge levels,
- reduction in evaporator air flow caused by fouled filters or evaporator coils,

- reduction in condenser air flow cause by fouled condenser coils,
- loss of compressor volumetric efficiency caused by compressor valve or other internal leakage,
- and liquid-line restrictions.

One major advantage of the virtual sensor approach is the ability to diagnose faults in the presence of multiple simultaneous faults. This is important since many faults may cause system performance to degrade over time if regular maintenance is not performed. In addition to detecting and diagnosing common faults, the cooling capacity, electrical power input, and COP of the system can be estimated using virtual sensors when the system operates.

One differentiating quality of the method implemented in this work is that faults are diagnosed using virtual sensors that are sensitive only to certain faults. This approach is what enables the diagnosis of multiple faults that occur simultaneously, which other methodologies were unable to do [1].

A virtual refrigerant charge (VRC) sensor originally proposed by Li and Braun and improved by Kim and Braun was implemented in the AFDD system [2, 16]. The VRC sensor is based on a semi-empirical relationship between refrigerant charge level and the suction-line superheat ΔT_{sh}, the liquid-line subcooling ΔT_{sc}, and the quality of the refrigerant entering the evaporator x_{eri}. There are three empirical parameters k_{sh}, k_{sc}, and k_x of the VRC sensor that are must be tuned using regression using data collected from a specific system,

$$\begin{aligned} m_{ref,virtual} = \; & 1 + k_{sh} \cdot \left(\Delta T_{sh} - \Delta T_{sh,rated} \right) \\ & + k_{sc} \cdot \left(\Delta T_{sc} - \Delta T_{sc,rated} \right) \\ & + k_x \cdot \left(x_{eri} - x_{eri,rated} \right) \end{aligned} \tag{2.1}$$

where $\Delta T_{sh,rated}$, $\Delta T_{sc,rated}$, and $x_{eri,rated}$ are the superheat, subcooling, and inlet quality of the system at a rating point. The VRC sensor builds a correlation between the amount of refrigerant in an air conditioner and the change in superheat, subcooling, and evaporator inlet refrigerant quality. These quantities are related to the amount of subcooled liquid refrigerant in the condenser and superheated vapor refrigerant in the evaporator coil.

The refrigerant mass flow rate can be estimated using compressor inlet and outlet measurements in two ways: using standardized manufacturer maps for positive displacement compressors and energy balances. In the first methodology, standardized empirical relationships for the compressor can often be obtained from manufacturers which take the form

$$\begin{aligned} X = {} & c_0 + c_1 \cdot S + c_2 \cdot D \\ & + c_3 \cdot S^2 + c_4 \cdot S \cdot D + c_5 \cdot D^2 \\ & + c_6 \cdot S^3 + c_7 \cdot S^2 \cdot D + c_8 \cdot S \cdot D^2 + c_9 \cdot D^3 \end{aligned} \tag{2.2}$$

where X can represent the power input to the compressor, the mass flow rate, the current input to the compressor, or compressor efficiency, S is the suction dew point temperature of the refrigerant, D is the discharge dew point of the refrigerant, and $c_{0,\ldots,9}$ are empirical parameters determined using compressor calorimeter data [34]. Using Equation (2.2), the refrigerant mass flow rate can be estimated based on inlet and outlet dewpoint temperatures (which can be estimated using suction and discharge pressure measurements),

$$\begin{aligned} \dot{m}_{ref,map} = {} & c_0 + c_1 \cdot S + c_2 \cdot D \\ & + c_3 \cdot S^2 + c_4 \cdot S \cdot D + c_5 \cdot D^2 \\ & + c_6 \cdot S^3 + c_7 \cdot S^2 \cdot D + c_8 \cdot S \cdot D^2 + c_9 \cdot D^3. \end{aligned} \tag{2.3}$$

When compressor performance maps are generated using calorimeters, the suction superheat is typically held constant for all test conditions. This leads to systematic biases in mass flow rate predictions when the actual suction-line superheat is different the test superheat. To reduce this

bias, a correction correlation developed by Dabiri and Rice can be used to account for differences in suction density

$$\frac{\dot{m}_{ref,virtual,1}}{\dot{m}_{ref,map}} = 1 + F \cdot \left(\frac{\rho_{suc,actual}}{\rho_{suc,map}} - 1 \right) \tag{2.4}$$

where $\rho_{suc,map}$ is the suction-line refrigerant density corresponding the superheat maintained during the mapping process, $\rho_{suc,actual}$ is the actual suction-line superheat, and F is an empirical parameter to account for mass flow increase with increasing suction-line density [35].

An alternative methodology to estimate the mass flow rate of refrigerant produced by the compressor can be derived from a compressor energy balance,

$$\dot{m}_{ref,virtual,2} = \frac{(1-\alpha_{loss})\dot{W}_{comp,virtual}}{h_{dis} - h_{suc}} \tag{2.5}$$

where h_{suc} is the suction-line refrigerant enthalpy, h_{dis} is the discharge-line refrigerant enthalpy, α_{loss} is a parameter to account for heat loss to the environment, and $\dot{W}_{comp,virtual}$ is an estimate of the compressor power input. Using Equation (2.2), a virtual compressor power sensor can be derived based on suction-line and discharge-line refrigerant dew point temperatures,

$$\begin{aligned} \dot{W}_{ref,virtual} = {} & c_0 + c_1 \cdot S + c_2 \cdot D \\ & + c_3 \cdot S^2 + c_4 \cdot S \cdot D + c_5 \cdot D^2 \\ & + c_6 \cdot S^3 + c_7 \cdot S^2 \cdot D + c_8 \cdot S \cdot D^2 + c_9 \cdot D^3. \end{aligned} \tag{2.6}$$

Unlike the mass flow rate, the compressor power consumption is less sensitive to differences in superheat. Previous research has shown that superheat correction factors do not add significant increases in model prediction accuracy [3, 6].

Refrigerant mass flow rate can also be estimated using measurements at the inlet and outlet of the throttling device used in the air conditioner. Mass flow rate of liquid refrigerant entering short tube orifices can be modeled using the single-phase orifice equation

$$\dot{m}_{ref,FXO} = K \cdot A \sqrt{\rho \cdot \left(P_{in} - P_{out}\right)} \tag{2.7}$$

where A is the cross-section area of the orifice opening, ρ is the density of the entering refrigerant, P_{in} is the pressure of the entering refrigerant, P_{out} is the leaving refrigerant pressure, and K is an empirical valve parameter. For refrigerants entering as two-phase mixtures, Payne and O'Neil propose an empirical correlation for mass flow rate using non-dimensional analysis [36]. Kim and Braun have developed mass flow rate models for thermostatic expansion valves (TXV) and electronic expansion valves (EEV) [3]. The TXV mass flow rate model is derived from mass and force balances, and is expressed as

$$\dot{m}_{ref,TXV} = \left[a_0 + a_1 \cdot \left(\mathbf{P_{sat}}\left(T_{suc}\right) - P_{suc}\right)^2 + a_2 \cdot \left(\mathbf{P_{sat}}\left(T_{suc}\right) - P_{suc}\right) \right] \cdot \dot{m}_{max,TXV} \tag{2.8}$$

where T_{suc} is the suction-line temperature, P_{suc} is the suction-line pressure, $\mathbf{P_{sat}}$ is a function used to calculate the saturated pressure of a refrigeration using temperature, and $\dot{m}_{max,TXV}$ is the limiting flow rate through the valve when it is fully open which can be estimated using a correlation developed by Lenger et al.,

$$\dot{m}_{max,TXV} = b_0 \cdot \sqrt{2 \cdot \rho_{in} \cdot \left(P_{in} - P_{out} \cdot K\right)} \tag{2.9}$$

$$K = b_1 + b_2 \cdot \left(\frac{\Delta T_{sc} + 2}{T_{critical}} \right)^2 + b_3 \cdot \left(\frac{P_{critical} - P_{out}}{P_{critical}} \right) \tag{2.10}$$

where $T_{critical}$ and $P_{critical}$ are the critical temperature and pressure of the refrigerant [37].

A methodology for estimating the condenser air flow rate based on a condenser coil energy balance was proposed by Li and Braun [6]. Using this methodology, the condenser air mass flow rate virtual sensor is defined by

$$\dot{m}_{ca,virtual} = \frac{\dot{m}_{ref,virtual} \cdot \left(h_{dis} - h_{cro}\right)}{c_{p,ca} \cdot \left(T_{ca,out} - T_{ca,in}\right)} \tag{2.11}$$

where h_{dis} is the enthalpy of refrigerant entering the condenser coil (usually the compressor discharge), h_{cro} is the enthalpy of refrigerant leaving the condenser coil, $c_{p,ca}$ is the specific heat of condenser air, $T_{ca,out}$ is the temperature of air leaving the condenser coil, and $T_{ca,in}$ is the temperature of air entering the condenser coil. The mass air flow rate of condenser air can be used to detect and diagnose faults impacting the condenser heat rejection effectiveness or fan efficiency.

Li and Braun also propose a similar methodology for estimating the evaporator air flow rate [6]. The virtual evaporator air mass flow rate sensor is given by

$$\dot{m}_{ca,virtual} = \frac{\dot{m}_{ref,virtual} \cdot \left(h_{suc} - h_{eri}\right)}{\left(h_{ea,in} - h_{ea,out}\right)} \tag{2.12}$$

where h_{suc} is the enthalpy of refrigerant leaving the evaporator coil (usually the compressor suction), h_{eri} is the enthalpy of refrigerant entering the evaporator coil, $h_{ea,in}$ is the enthalpy of air entering the evaporator coil, and $h_{ea,out}$ is the enthalpy of air leaving the evaporator coil. The enthalpy of entering and leaving the evaporator coil can be estimated when air temperature and relative humidity measurements are available. The mass air flow rate of evaporator air can be used to detect and diagnose faults impacting the evaporator heat transfer effectiveness or fan efficiency.

The empirical and mathematical models for the different virtual sensors were implemented using several open-source numerical computing libraries in the Python programming language [38]. These models were incorporated into VOLTTRON™ compatible agents within a low-cost computing device. This software also included data acquisition functionality to interface with sensors required to monitor system operation.

2.4 Experimental Evaluation of AFDD System Applied to Multistage RTU

2.4.1 Description of Multistage RTU and Experimental Setup

To test the performance of the fault detection and diagnostics system, an RTU with 5-ton rated capacity was installed in psychrometric chambers at Ray W. Herrick Laboratories (shown in Figure 2.2). The system was installed with a two-stage scroll compressor, a finned-tube evaporator, a finned-tube condenser, and R410A as the working fluid. Detailed descriptions of the RTU components are presented in Table 2.1. The system that was tested had moderately high rated efficiencies when compared to other fixed-speed and multistage equipment currently available by manufacturers. The actual nameplate rated efficiencies of the system tested are described in Table 2.2.

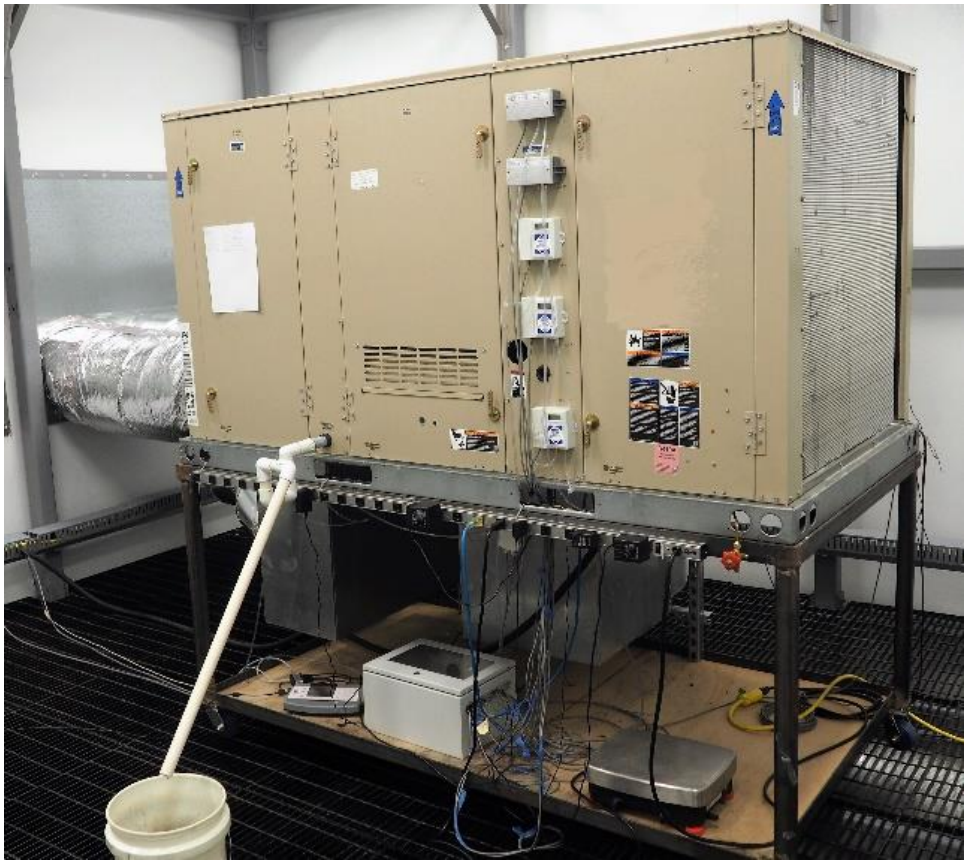

Figure 2.2. 5-ton RTU used to test automated fault detection and diagnostics system installed in psychrometric chambers.

Table 2.1. Component descriptions of multi-stage RTU used for development and evaluation of AFDD implementation.

Component	**Description**
Compressor	Multi-Stage Scroll Compressor
Indoor Blower	Direct-Drive, Multi-Stage Centrifugal Blower
Outdoor Fan	Direct-Drive, Variable-speed Propeller Fan
Expansion Valve	Balance Port TXV, Suction Superheat Control
Evaporator Coil	Round Tube, Plate Fin Heat Exchanger
Condenser Coil	Round Tube, Plate Fin Heat Exchanger
Refrigerant	R-410A, 7.57 kg (16 lbs 11 oz)

Table 2.2. Rated cooling performance of multi-stage RTU used for development and evaluation of AFDD implementation.

Rated Cooling Performance[1]	
Gross Cooling Capacity	18.05 kW (61.6 kBtu/h)
Net Cooling Capacity[2]	17.58 kW (60.0 kBtu/h)
AHRI Rated Air flow	0.8259 m^3/s (1750 CFM)
Total Unit Power	4.7 kW
EER	12.7 Btu/h / W
SEER	17.1 Btu/h / W

[1] AHRI Certified to AHRI Standard 210/240: 95 °F outdoor air temperature and 80 °F dry bulb / 67 °F wet bulb entering evaporator air, minimum external duct pressure.
[2] Net capacity includes evaporator blower motor heat deduction.

To test the performance of the AFDD system, the RTU was operated under a range of operating conditions, described in Table 2.3. This included both wet coil and dry coil tests under outdoor ambient conditions ranging from 20.6 °C (69.0 °F) to 42.2 °C (108.0 °F). To test the performance of the AFDD system at diagnosing refrigerant charge faults, the system charge level was adjusted to values ranging from 60% (40% undercharge) to 120% (20% overcharge). The ability to diagnose reductions in evaporator and condenser air flow (simulating fouling) was tested by adjusting the indoor blower and outdoor fan control inputs to prescribed levels. Two levels of air flow were tested for each fan: normal and reduced. Finally, combinations of all three faults were tested simultaneously at the different ambient conditions.

Table 2.3. Test conditions for RTU with finned-tube condenser and fixed orifice expansion device for low stage cooling operation in psychrometric test chambers.

Test Variable	**Test Values**	
Compressor Stage, -	LOW	HIGH
Indoor Dry Bulb, °C	26.7	26.7
Indoor Wet Bulb, °C	13.9, 20.6	13.9, 20.6
Outdoor Dry Bulb, °C	20.6, 27.8, 35.0, 41.2	20.6, 27.8, 35.0, 41.2
Charge Level[1], %	60, 70, 80, 90, 100, 110, 120	60, 70, 80, 90, 100, 110, 120
Indoor Fan Torque[2], %	30, 60	50, 90
Outdoor Fan Speed[3], %	40, 70	70, 100

[1] Charge is measured relative to the recommended charge according to the manufacturer's nameplate data.
[2] Indoor fan torque is set according to a nominal flow rate of 0.637 m^3/s (1350 CFM) for low stage operation; 0.944 m^3/s (2000 CFM) for high stage operation.
[3] Outdoor fan speeds are set using the manufacturer's default value for low and high stage operation.

2.4.2 Measured Detection and Diagnosis Performance of AFDD System

The accuracy of the VRC model for each stage of operation is shown in Figure 2.3 over the range of ambient conditions and fault scenarios tested. The virtual refrigerant charge sensor required empirical data to tune the model parameters. These parameters were tuned using an automated training algorithm without the use of psychrometric chambers [34]. In both stages of operation, the prediction accuracy was mostly within 10% of the measured charge levels. Moreover, the root-mean-square error (RMSE) was approximately 6.20% for both stages operation. It should also be noted that the relative accuracy remains approximately unchanged over the range of refrigerant charge levels tested and minimal bias exists as a function of outdoor ambient temperature.

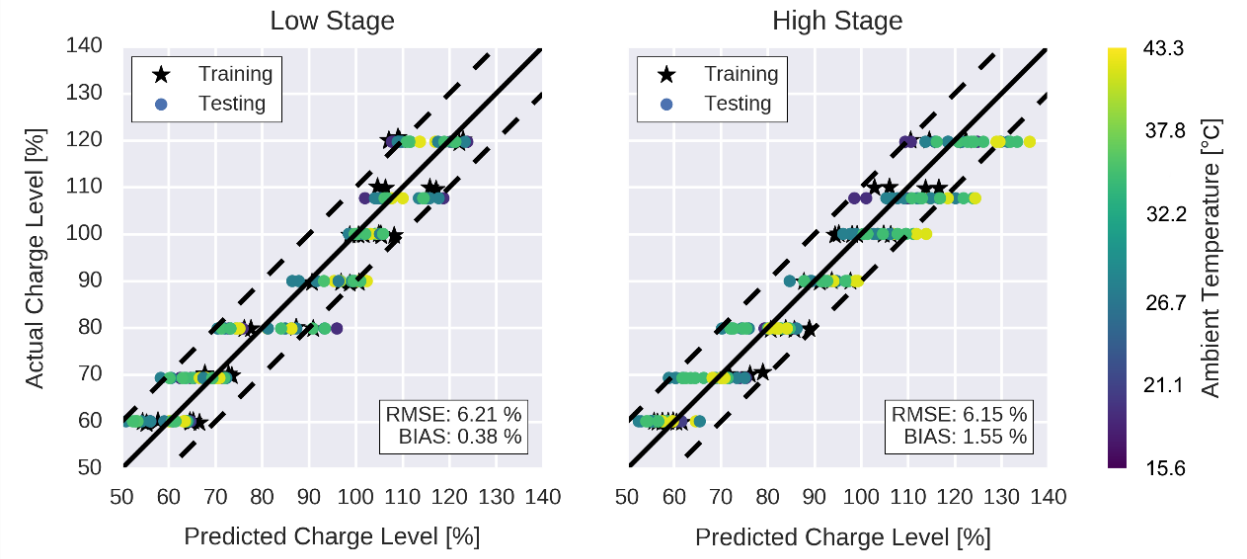

Figure 2.3. VRC sensor prediction accuracy for RTU with finned tube condenser and fixed orifice expansion device applied to both stages of operation under different ambient conditions.

Figure 2.4 shows the accuracy of the virtual charge sensor as well as the calculated fault probability for tests cases with normal air flow rate levels. Good agreement between the predicted charge level and the measured charge level was observed for these test cases. The charge fault probability was calculated using a statistical methodology originally described by Rossi and Braun [1]. Essentially, this assumes that two Gaussian distributions can be used to represent the expected and observed refrigerant charge levels. The degree that these two distributions overlap is analogous to the probability that the observed charge level is equal to the expected charge level. This probability is then calculated by integrating the area contained by both distributions. The results shown in Figure 2.4 show that the FDD system can identify refrigerant charge faults with probabilities greater than 95% when the actual charge level deviates ±15% from normal.

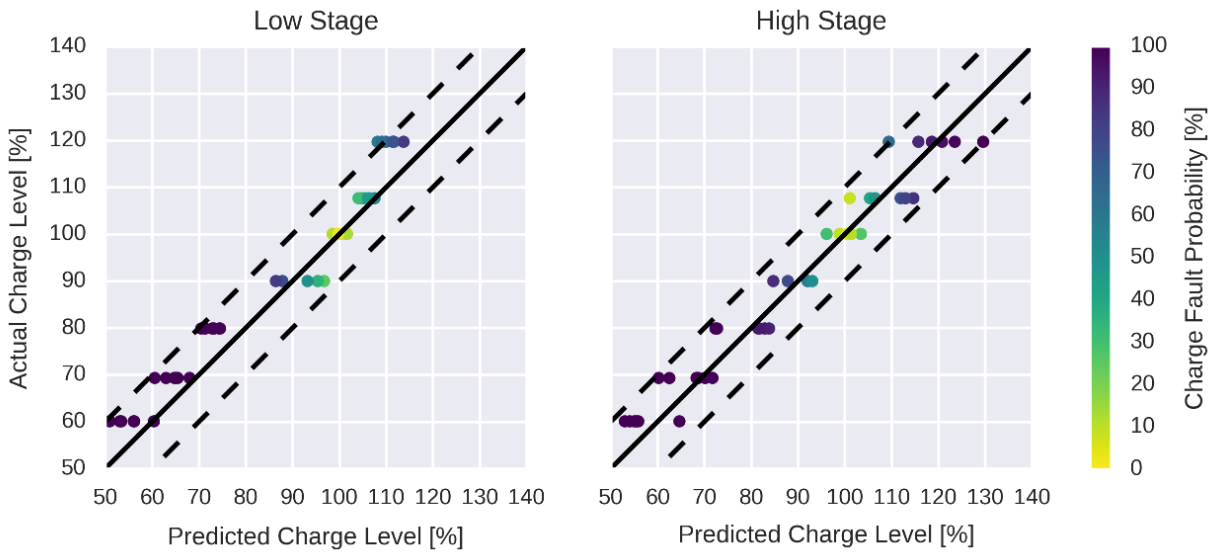

Figure 2.4. Prediction accuracy of VRC sensor applied to charge level vaults showing charge fault probability for each test case. The data show that the FDD system can identify faults with high confidence when actual charge is ±20% of the normal charge level.

It is important that the FDD system can accurately identify faults when there are multiple simultaneous faults affecting the system. Figure 2.5 shows the performance of the charge fault diagnostics for several combinations of improper refrigerant levels, indoor air flow levels, and outdoor air flow levels. In these results, tested under 27.78 °C (82.0 °F) outdoor temperature, refrigerant charge faults were correctly identified with probabilities greater than 95% in all cases where the charge level deviated by at least ±20%. It is also noteworthy that the VRC sensor tends to predict refrigerant charge faults with greater probability when the impact on system capacity of the faults increases. This indicates that when significant refrigerant charge faults affect the system, the FDD system can identify them.

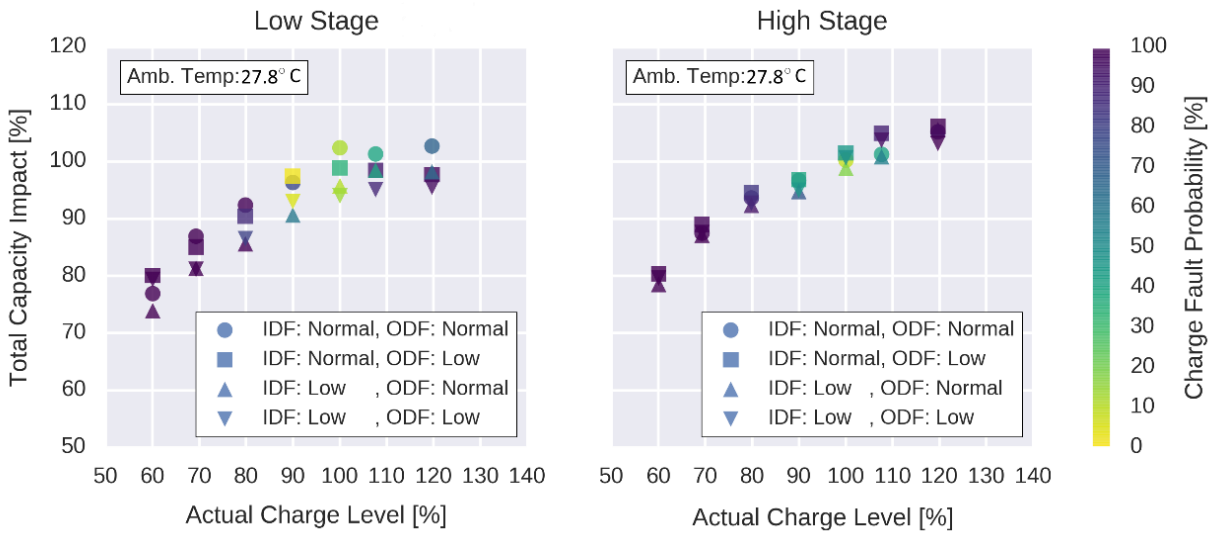

Figure 2.5. Total capacity impact of refrigerant charge faults at 27.78 °C (82 °F) outdoor ambient temperature. The points are colored based on the fault probability determined for each test case, indicating faults that have larger impacts are identified with greater probability. IDF indicates the indoor fan torque setting; ODF indicates the outdoor fan speed setting.

Similar results for test conditions at 35.00 °C (95.0 °F) and 41.22 °C (108.0 °F) outdoor temperature are shown in Figure 2.6 and Figure 2.7 respectively. In these results, refrigerant charge faults were correctly identified with high probabilities in test cases where the refrigerant charge deviated by ±10%. It should also be noted that refrigerant charge faults tended to have large impacts on the total capacity of the system at the higher outdoor temperatures.

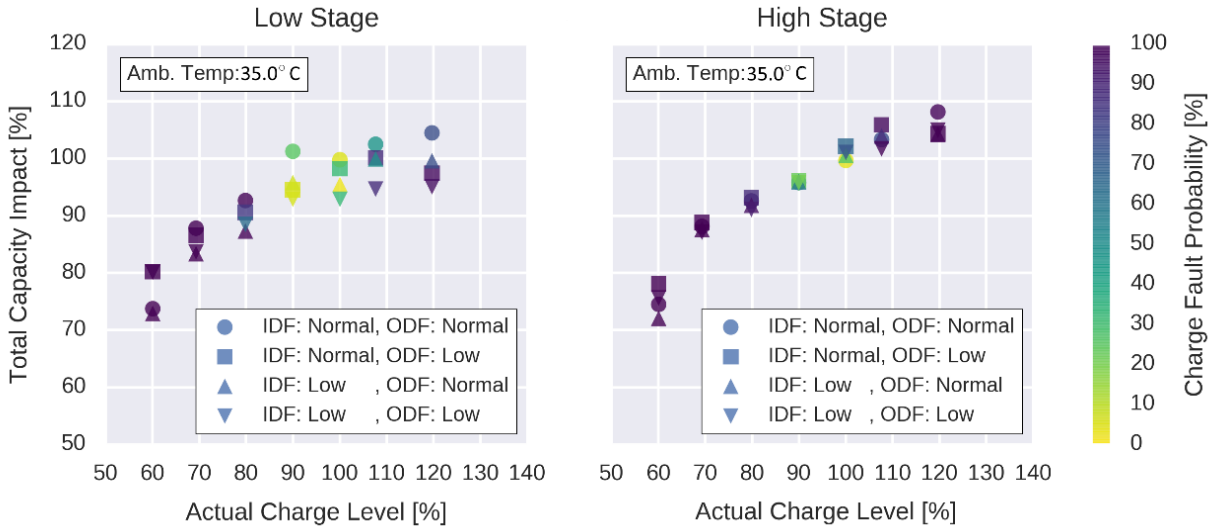

Figure 2.6. Total capacity impact of refrigerant charge faults at 35.00 °C (95 °F) outdoor ambient temperature. The points are colored based on the fault probability determined for each test case, indicating faults that have larger impacts are identified with greater probability. IDF indicates the indoor fan torque setting; ODF indicates the outdoor fan speed setting.

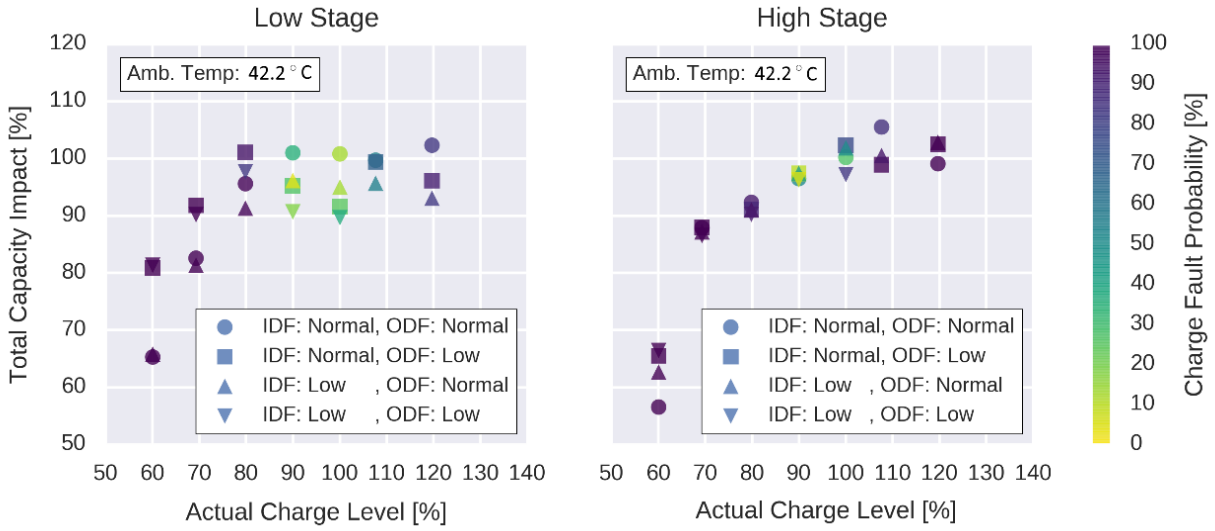

Figure 2.7. Total capacity impact of refrigerant charge faults at 41.22 °C (108 °F) outdoor ambient temperature. The points are colored based on the fault probability determined for each test case, indicating faults that have larger impacts are identified with greater probability. IDF indicates the indoor fan torque setting; ODF indicates the outdoor fan speed setting.

The probability of an evaporator air flow fault has been calculated for each test case to test the effectiveness of the FDD system at identifying evaporator fouling faults. This probability was calculated using the previously described method using the virtual evaporator air flow rate sensor. The results in Figure 2.8 compare test cases at 27.78 °C (82.0 °F) ambient conditions where evaporator air flow was normal and reduced to simulate a fouling fault. Fault cases that had reductions in evaporator air flow resulted in high fault probabilities, which can be used to identify faults. The results illustrate that evaporator fouling can be identified even in the presence of other faults, like low refrigerant charge.

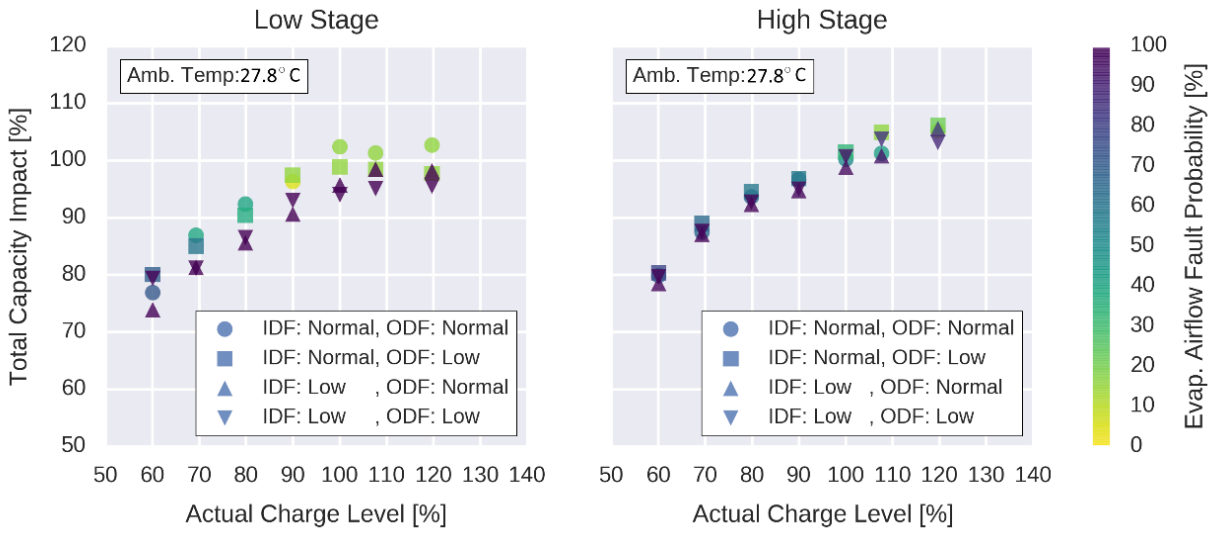

Figure 2.8. The FDD system can identify test cases with low evaporator air flow using the statistical fault detection and diagnostics method. Test cases without reductions in evaporator air flow were not identified with high fault probabilities. IDF indicates the indoor fan torque setting; ODF indicates the outdoor fan speed setting.

The probability of a condenser air flow fault has been calculated for each test case to test the effectiveness of the FDD system at identifying condenser fouling faults. This probability was calculated using the previously described method using the virtual condenser air flow rate sensor. The results in Figure 2.9 compare test cases at 27.78 °C (82.0 °F) ambient conditions where

condenser air flow was normal and reduced to simulate a fouling fault. Fault cases that had reductions in condenser air flow resulted in high fault probabilities, which can be used to identify faults. The results illustrate that condenser fouling can be identified even in the presence of other faults, like low refrigerant charge.

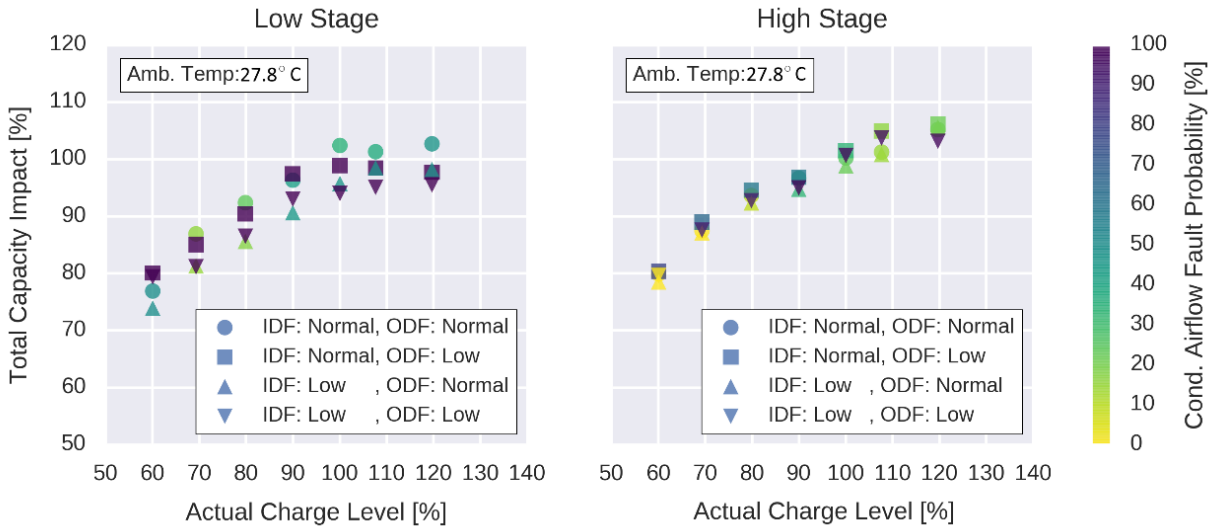

Figure 2.9. The FDD system can identify test cases with low condenser air flow using the statistical fault detection and diagnostics method for test cases under 27.78 °C (82 °F) outdoor ambient temperature. Test cases without reductions in condenser air flow were not identified with high fault probabilities. IDF indicates the indoor fan torque setting; ODF indicates the outdoor fan speed setting.

2.5 Experimental Evaluation of AFDD System Applied to Variable-Speed RTU

While there have been developments in AFDD systems for variable-speed direct-expansion equipment, this has been to a much lesser extent when compared to fixed-speed systems. Moreover, much of this work has focused on ductless heat pumps (sometimes called mini-split systems) and use mostly data-driven methods which have been shown to be susceptible to errors when models are overfit or when extrapolation is required. The performance advantages of systems with variable-speed compressors are significant. By using a variable-speed compressor, air conditioners and heat pumps can be controlled to match the time-varying, dynamic heating

and cooling loads of a building. This has the potential to increase heating and cooling efficiencies of equipment, especially at part-load conditions by reducing the cycling inefficiencies associated with ON/OFF or staged equipment.

These efficiency improvements, along with increasing energy efficiency regulatory measures, have driven manufacturers to design and sell more systems with variable-speed compressors having advanced controls. Thus, more and more variable-speed systems have been introduced into US commercial and residential buildings. Like fixed-speed systems, systems with variable-speed components still have the potential to develop faults that degrade performance over time. Though with the relative lack of AFDD system developments and demonstrations for variable-speed equipment as compared with fixed-speed system, investigation of low-cost approaches is needed. In this section, the virtual sensor AFDD approach has been extended so that it can be applied to a variable-speed RTU. Using experimental testing, the methodology is evaluated using data collected from a system at different operating conditions and fault levels.

2.5.1 Description of Variable-Speed RTU and Experimental Setup

A new 5-ton variable-speed RTU was installed in a pair of psychrometric chambers at Ray W. Herrick Laboratories for the purposes of testing a novel supervisory level controller designed to minimize energy consumption of variable-speed RTUs. This system, pictured in Figure 2.10, was used to perform a series of tests to evaluate the performance of this controller in comparison to several baseline control methodologies. In addition to testing control algorithms, the impacts of common faults that affect RTUs were measured using the psychrometric chambers. These faults included improper refrigerant charge level, condenser fouling, and evaporator fouling.

Figure 2.10. 5-ton variable-speed RTU that will be used to develop and evaluate AFDD implementation.

The major components of the variable-speed RTU are described in Table 2.4. The compressor used in the system is an inverter driven variable capacity scroll compressor. The compressor input frequency was controlled to maintain a constant supply air temperature set point by the main RTU controller. The evaporator blower was controlled to produce different air flow rates in response to the zone temperature using a constant torque fan motor. It is also important to note that the system utilized a thermostatic expansion valve (TXV) to control suction superheat entering the compressor. In comparison to the previous multistage RTU tested, the only significant difference between the systems was the compressor used in the system. Additionally, the control of the compressor and fan speeds by the RTU controller was also significantly different.

Table 2.4. Component descriptions of variable-speed RTU used for development and evaluation of AFDD implementation.

Component	Description
Compressor	Inverter Driven, Variable Capacity Scroll Compressor
Indoor Blower	Direct-Drive, Multi-Stage Centrifugal Blower
Outdoor Fan	Direct-Drive, Variable-speed Propeller Fan
Expansion Valve	Balance Port TXV, Suction Superheat Control
Evaporator Coil	Round Tube, Plate Fin Heat Exchanger
Condenser Coil	Round Tube, Plate Fin Heat Exchanger
Refrigerant	R-410A, 7.57 kg (16 lbs 11 oz)

The rated performance of the variable-speed RTU that was be used to evaluate the variable-speed RTU AFDD implementation is shown in Table 2.5. In comparison to the multi-stage RTU tested previously, a significant improvement in the seasonal energy efficiency ratio (SEER) can be observed, which is due to improved performance at part-load conditions. This illustrates the energy saving potential provided by the variable-speed compressor that avoids degradation in performance due to cycling at part loads.

Table 2.5. Rated cooling performance of variable-speed RTU used for development and evaluation of AFDD implementation.

Rated Cooling Performance[1]	
Gross Cooling Capacity	17.44 kW (59.5 kBtu/h)
Net Cooling Capacity[2]	17.00 kW (58.0 kBtu/h)
AHRI Rated Air flow	0.8595 m^3/s (1800 CFM)
Total Unit Power	4.5 kW
EER	12.0 Btu/h / W
SEER	20.0 Btu/h / W

[1] AHRI Certified to AHRI Standard 210/240: 95 °F outdoor air temperature and 80 °F dry bulb / 67 °F wet bulb entering evaporator air, minimum external duct pressure.
[2] Net capacity includes evaporator blower motor heat deduction.

To test the performance of the AFDD implementation applied to the variable-speed RTU, the psychrometric chamber test facilities at Ray W. Herrick Laboratories was used to maintain test conditions simulating typical equipment operating conditions. One interesting aspect of the control system of the variable-speed RTU is that the system can be operated in full variable-speed mode (as is intended by the manufacturer) but also using a fixed-speed and multistage

control approach. Not only did this allow for a direct comparison of different control approaches, but it will also be possible to compare the performance of the AFDD system when applied to different control approaches. Based on a literature review, no previous study on the performance of AFDD systems with respect to the capacity control mechanism of RTUs has been studied.

A combination of steady-state testing and dynamic load-based testing scenarios were used for validation of the AFDD system within psychrometric chamber test facilities. A description of the load-based testing approach has been included in Appendix A. The steady-state test scenarios served two purposes: for development and tuning of empirical models used in some of the virtual sensors and for evaluation of the FDD systems when systems are under different constant loads. The steady-state tests that were performed are based on the equipment performance rating standard relevant for the equipment (AHRI Standard 210/240) and are described in Table 2.6 [39]. Scenarios with both wet coil and dry coil conditions were included in the steady-state testing program to evaluate whether FDD performance degrades with different evaporator inlet conditions. Besides analyzing the performance of the AFDD system, reference models for the normal performance of the system were developed using the data generated from these tests. These results are presented in Chapter 3.

Table 2.6. Steady-state test scenarios for variable-speed RTU AFDD performance assessment.

Test Variable	Test Values
Part-Load Ratio, -	0.60, 0.75, 0.90, 1.00
Indoor Dry Bulb, °C	26.7 (80.0 °F)
Indoor Relative Humidity, %	25, 45, 55
Outdoor Dry Bulb, °C	25.0, 30.0, 35.0 (77.0, 86.0, 95.0 °F)
Charge Level[1], %	60, 70, 80, 100, 120, 130
Indoor Fouling, -	None, Light, Heavy
Outdoor Fouling[3], %	None, Light, Heavy

[1] Charge is measured relative to the recommended charge according to the manufacturer's nameplate data.
[2] Indoor fouling was imposed by restricting air flow through an orifice damper to increase pressure head on indoor fan.
[3] Outdoor fouling was imposed by uniformly adding flow restriction through condenser coil using different layers of porous sheets.

To evaluate the performance of the AFDD system experimentally, the performance of the RTU while subject to different faults was observed. For the purposes of this testing, faults were injected into the system to measure their impacts on the system and to assess how well the AFDD system is able to identify and classify them. Three fault types were considered for this study: improper refrigerant charge levels, condenser fouling (simulated by uniformly adding flow restriction through the condenser coil using different layers of porous sheets), and evaporator fouling (simulated by restricting air flow through an orifice damper to increase pressure head on the indoor fan). It should be noted that combinations of different faults will also be studied (but not exhaustively).

The amount of refrigerant contained in the RTU was measured prior to each test scenario and was held constant throughout the duration of the test scenario. The normal charge level was determined using the manufacturer's nameplate data, while adjusting for additional refrigerant piping due to the mass flow meter. To reduce uncertainty about the amount of refrigerant contained in the system and to reduce the time required to adjust the refrigerant charge level, a

previously developed automated refrigerant charging apparatus was used to add and remove different amounts of charge between tests [40].

Condenser fouling was simulated using porous sheets that were placed uniformly across the condenser coil face area. Different number of layers were added to the condenser inlet to simulate different levels of condenser fouling. During each test, the number of layers was not changed.

The supply air static pressure head was controlled using a variable area orifice damper to simulate evaporator coil fouling. For normal fault conditions, the orifice damper was opened to its largest cross-sectional flow area. For subsequent tests, the cross-sectional area was reduced, producing greater static pressure at the outlet of the RTU. This resulted in lower air flow rates and higher fan power requirements.

Instrumentation was added to the variable-speed RTU to evaluate the performance of the AFDD system and to collect data for model tuning. The refrigerant-side sensors and their locations relative to the major components of the RTU are shown in Figure 2.11 and the associated sensor accuracies are shown in Table 2.7. The sensors installed on the refrigerant-side of the system enabled the calculation of refrigerant-side cooling capacity, COP, and system superheat and subcooling. This was important for both the AFDD implementation evaluation and for assessing performance impacts of the faults.

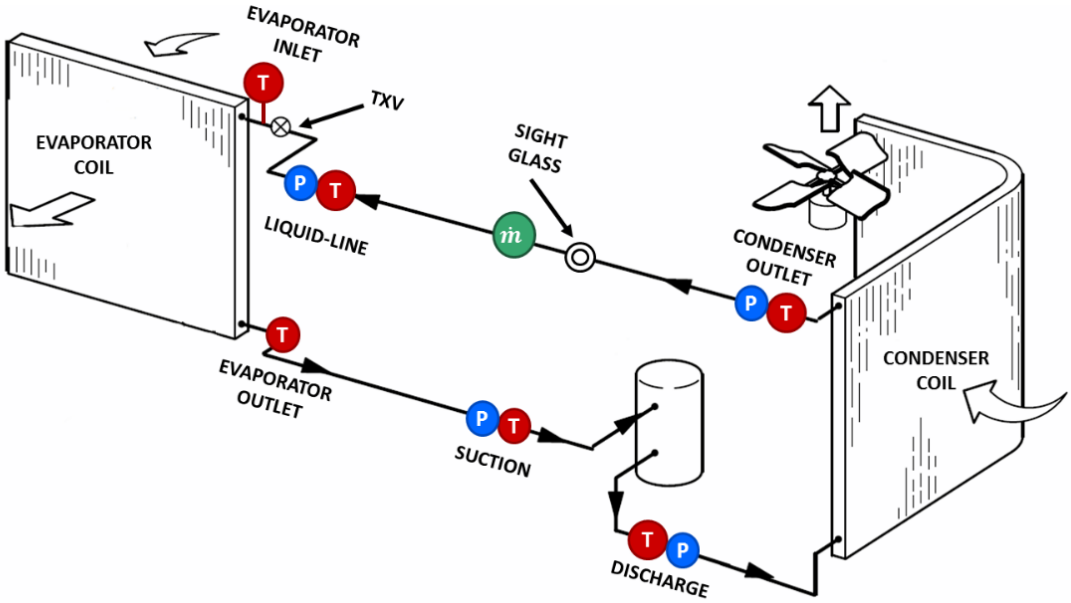

Figure 2.11. Refrigerant-side temperature, pressure, and mass flow sensor layout for variable-speed RTU.

Table 2.7. Refrigerant-side sensor measurement uncertainty for sensors installed on variable-speed RTU.

Location	Temperature °C (°F)	Pressure kPa (PSI)	Mass Flow g/s (lbm/min)
Compressor Suction	0.5 (0.9)	5.2 (0.75)	N/A
Compressor Discharge	0.5 (0.9)	8.6 (1.25)	N/A
Condenser Outlet	0.5 (0.9)	8.6 (1.25)	0.76 (0.10)
Liquid-Line	0.5 (0.9)	8.6 (1.25)	N/A
Evaporator Inlet	0.5 (0.9)	N/A	N/A
Evaporator Outlet	0.5 (0.9)	N/A	N/A

In addition to the refrigerant-side sensors installed, air-side temperature, pressure, and humidity sensors were installed in the ductwork used to transfer air to and from the RTU. These sensors were installed in accordance with ASHRAE Standard 41 to ensure reliability and reproducibility of the experimental data [41-43]. In addition to the sensors previously mentioned, it was intended to use a calibrated air flow nozzle to measure the supply air flow rate delivered by the unit to calculate air-side total and sensible capacity. However, significant errors existed with the air flow rates measured for reasons that could not be explained during testing. As a replacement, a virtual air flow rate sensor output developed by the manufacturer was used to

estimate the supply air flow rate. To verify the virtual sensor accuracy, refrigerant-side to air-side energy balances were used. Air-side measurements were required to evaluate the impacts faults have on sensible heat ratio (SHR) since it is not possible to directly measure this using only refrigerant-side measurements.

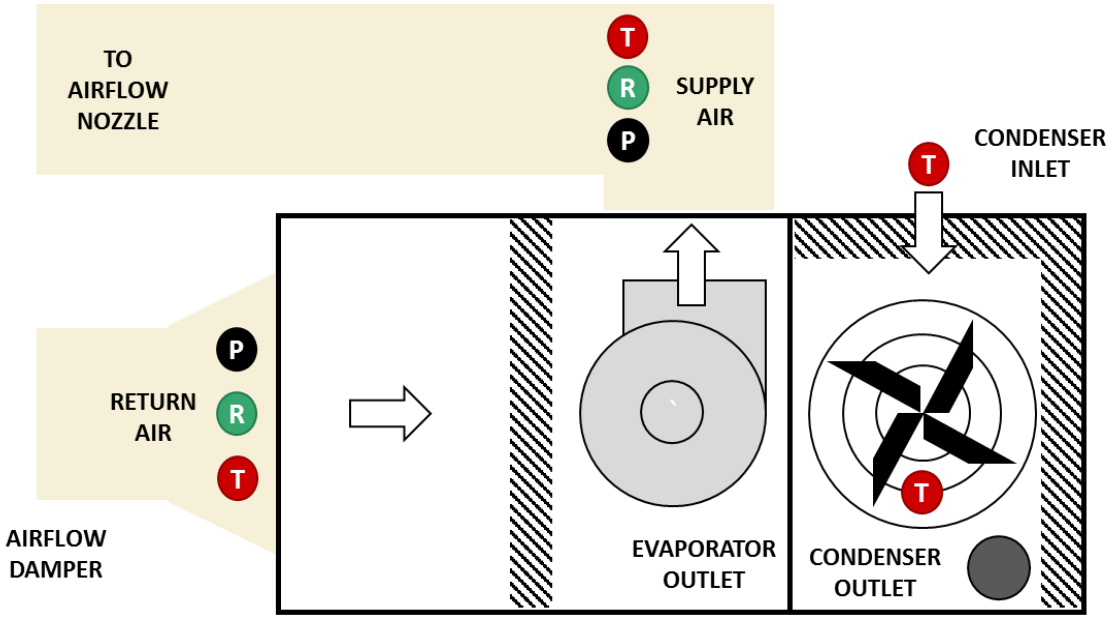

Figure 2.12. Air-side temperature, pressure, and humidity sensor layout for variable-speed RTU.

2.5.2 Extending AFDD System for Variable-Speed RTU

The virtual sensors developed for the multistage RTU described previously in this chapter could largely be applied to the variable-speed RTU without modifications. For a refrigerant mass flow rate sensor based on the compressor map, a modification was required to account for the performance of the compressor at different operating frequencies. To account for this, different sets of empirical coefficients were generated for the compressor for different operating frequencies. Thus, for 22, 30, 40, 50, 60 Hz operating frequencies, the refrigerant mass flow rate and compressor input power could be calculated using Equations (2.3) and (2.6) respectively.

Linear interpolation was used to predict the mass flow rate and input power when the compressor operated at intermediate frequencies.

Using interpolation of the compressor mass flow rate maps, the predicted and measured refrigerant mass flow rates are compared at different operating conditions in Figure 2.13. Over the range of operating conditions tested, the refrigerant mass flow rate predictions had errors within approximately 5% of the measured values. It should be noted that the range of operating frequencies tested ranged from approximately 35% to 100%. Over this range, the flow rates were reduced approximately 50% compared to full-load operation. It should also be noted that the density correction in Equation (2.4) was applied to account for variations in suction-line superheat. Without adjusting for the suction-line superheat, the refrigerant mass flow rate predictions tended to underestimate the measured flow rates by approximately 5%. This was due to the compressor maps being generated using a constant 11.1 °C (20.0 °F), while during testing the superheat was maintained around 5.5 °C (10.0 °F).

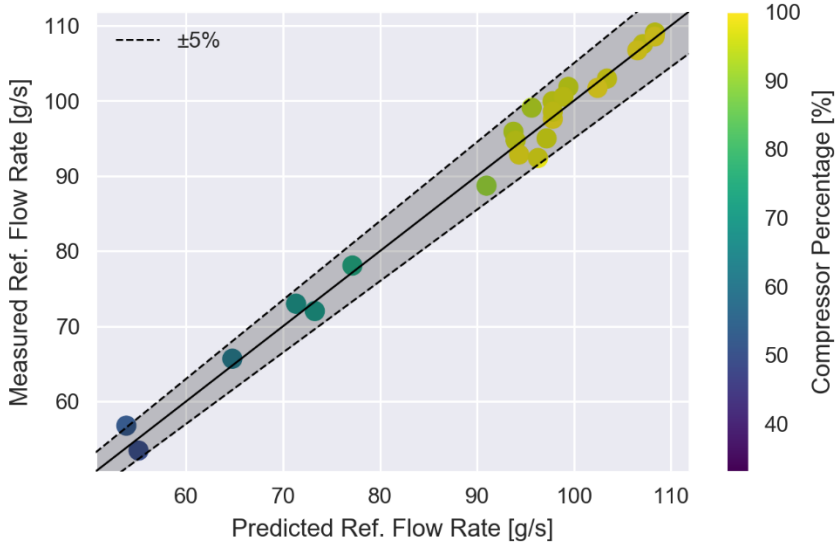

Figure 2.13. Variable-speed RTU virtual refrigerant mass flow rate sensor accuracy over range of operating conditions.

A comparison between the predicted and measured compressor power consumption for the experimental test cases is shown in Figure 2.14. Using interpolation of the compressor input power maps, the predictions agreed to within 5% of the measured power consumption for all but one test case. For the data collected, the range of compressor powers tested ranged from 40% to 100% of the full-load power consumption. This indicates that interpolation between compressor maps for different input frequencies can be performed without significant loss of accuracy in the virtual sensor. This provides a straightforward methodology to adapt the virtual refrigerant mass flow rate and input power sensors previously developed for fixed-speed and multistage systems to variable-speed systems.

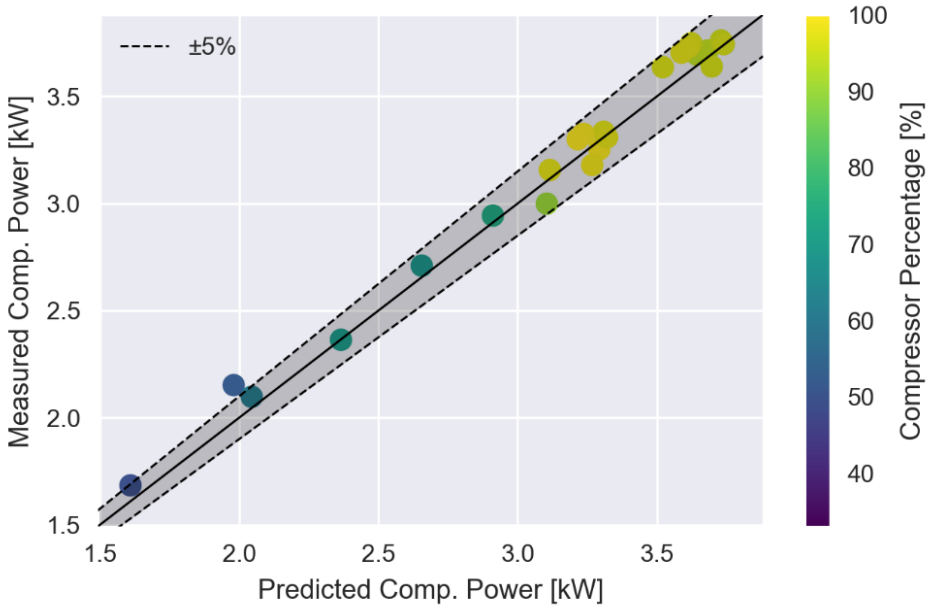

Figure 2.14. Variable-speed RTU virtual compressor power sensor accuracy over range of operating conditions.

Using the virtual mass flow rate sensor and measurements of the refrigerant and air entering and leaving the condenser coil, the virtual condenser airflow sensor was applied to the variable-speed RTU. The predicted and measured condenser airflow rates are compared in Figure 2.15.

Over the range of operating conditions and compressor speeds tested, the virtual condenser air flow sensor outputs agreed within 5% of the condenser airflow calculated using direct refrigerant mass flow rate and an energy balance on the condenser coil. In this system, the condenser fan speed was controlled based on compressor speed. Thus, it is important to evaluate the condenser airflow accuracy over different ambient conditions and compressor input frequencies. From the results shown in Figure 2.15, no bias is observed over the range of compressor input frequencies tested.

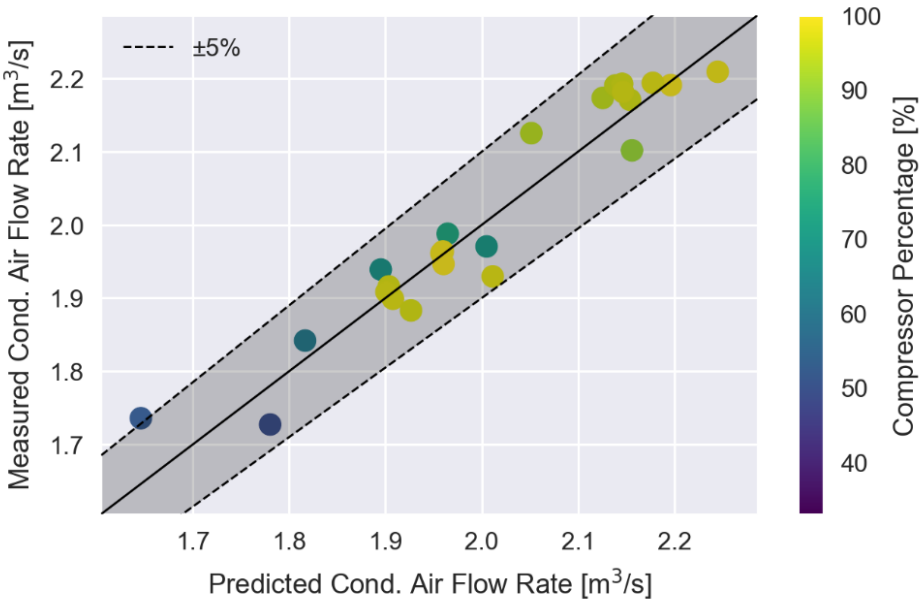

Figure 2.15. Variable-speed RTU virtual condenser air flow rate sensor accuracy over range of operating conditions.

The accuracy of the virtual evaporator airflow rate sensor applied to the variable-speed RTU in comparison to the directly measured airflow rate is shown in Figure 2.16. For all the test cases analyzed, the predicted and measured evaporator airflow rate errors were within 10%. In this system, the indoor fan is controlled to maintain zone air temperature setpoint. It can be inferred from these results that the virtual evaporator air flow rate outputs were relatively accurate for the

range of part-load fractions tested. No discernable correlation between the prediction error and the compressor operating percentage is observed.

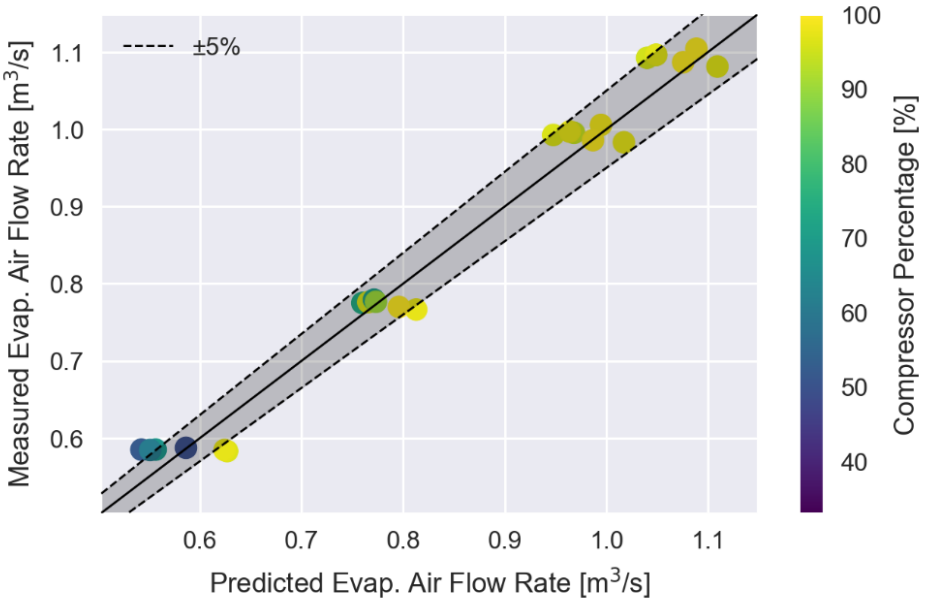

Figure 2.16. Variable-speed RTU virtual evaporator air flow rate sensor accuracy over range of operating conditions.

2.6 Description of AFDD System Hardware Design and Prototype

Due to the limited availability of sensors installed on existing RTUs in production and the relatively modest computing resources available, additional electronics hardware has been designed for implementing the AFDD algorithms. The system designed can be considered standalone – these electronics can be installed on a typical RTU and with the proper initial configuration, an effective AFDD system can be utilized by building operators without any other sensor requirements or hardware. While this system could seemingly be applied as a retrofit, it was primarily designed from the standpoint of being embedded by equipment manufacturers during the production process.

To implement the virtual-sensor-based AFDD algorithms, several refrigerant-side and air-side temperature measurements (shown in Table 2.8) are required. To measure these temperatures, a low-cost buffered analog-to-digital thermistor circuit was designed. The thermistors selected for the application can be easily surface-mounted to the RTU refrigerant circuit in the locations required. In comparison to other types of temperature sensors (thermocouples, RTDs, etc.) thermistors offer a good combination of accuracy, reliability, and cost. When using thermistors, the highly nonlinear relationship between temperature and internal resistance must be considered during the design process. While there are different ways to address this problem, each with their own tradeoffs, a more expensive (yet still relatively inexpensive) analog-to-digital converter (ADC) with a higher resolution was selected for this application.

Table 2.8. Description of required refrigerant-side temperature sensors used for RTU AFDD methods.

Symbol	Type	Description
T_{eri}	10K Thermistor[1]	Evaporator Refrigerant Inlet Temperature
T_{suc}	10K Thermistor	Compressor Refrigerant Suction Temperature
T_{dis}	10K Thermistor	Compressor Refrigerant Discharge Temperature
T_{crs}	10K Thermistor[2]	Condenser Refrigerant Saturation Temperature
T_{cro}	10K Thermistor	Condenser Refrigerant Outlet Temperature
T_{cai}	10K Thermistor	Condenser Air Inlet Temperature
T_{cao}	10K Thermistor	Condenser Air Outlet Temperature
T_{eai}	Temp/RH Chip	Evaporator Air Inlet Temperature
ϕ_{eai}	Temp/RH Chip	Evaporator Air Inlet Relative Humidity
T_{eao}	Temp/RH Chip	Evaporator Air Outlet Temperature
ϕ_{eao}	Temp/RH Chip	Evaporator Air Outlet Relative Humidity

[1] In some applications, a compressor suction pressure measurement is available. When this is the case, the T_{eri} sensor is not required since the evaporating temperature can be calculated using two-phase property relations.

[2] In some applications, a compressor discharge pressure measurement is available. When this is the case, the T_{crs} sensor is not required since the condensing temperature can be calculated using two-phase property relations.

It should also be noted that pressure measurements can be used to calculate the evaporator refrigerant inlet temperature and condenser refrigerant saturation temperature since the refrigerant at these points is a two-phase fluid. Systems that already have these pressure sensors installed for control purposes do not need to install additional temperature sensors which reduces the additional instrumentation costs for AFDD. It is also worth noting that pressure sensors may be required for systems with micro-channel condensers. On these systems, locating a consistent and reliable saturation temperature point over the expected equipment operating range is not a trivial task [40].

The RTU AFDD methodology requires a calculation of the enthalpy of the entering and leaving evaporator air to determine if the evaporator coil is fouled or air flow has been reduced. Measuring enthalpy directly is not possible, so the temperature and relative humidity at these

points is measured instead (described in Table 2.8). In the initial design, two solid-state sensors that measured both the dry-bulb temperature and relative humidity were used. Unlike the analog thermistor circuits described previously, these sensors provide a digital output using the I^2C communications protocol. Besides using these sensors for the virtual evaporator air flow sensor, the entering evaporator air dry bulb temperature and wet bulb temperature are used as inputs in reference models of normal performance also used by the AFDD algorithms.

An overall schematic view of the temperature, pressure, and relative humidity sensors used by the RTU AFDD system is shown in Figure 2.17. In comparison with the sensors typically found on RTUs, significantly more sensors are required. The additional sensor sensors enable the ability to diagnosis simultaneous faults and estimate impacts of faults on cooling capacity and COP.

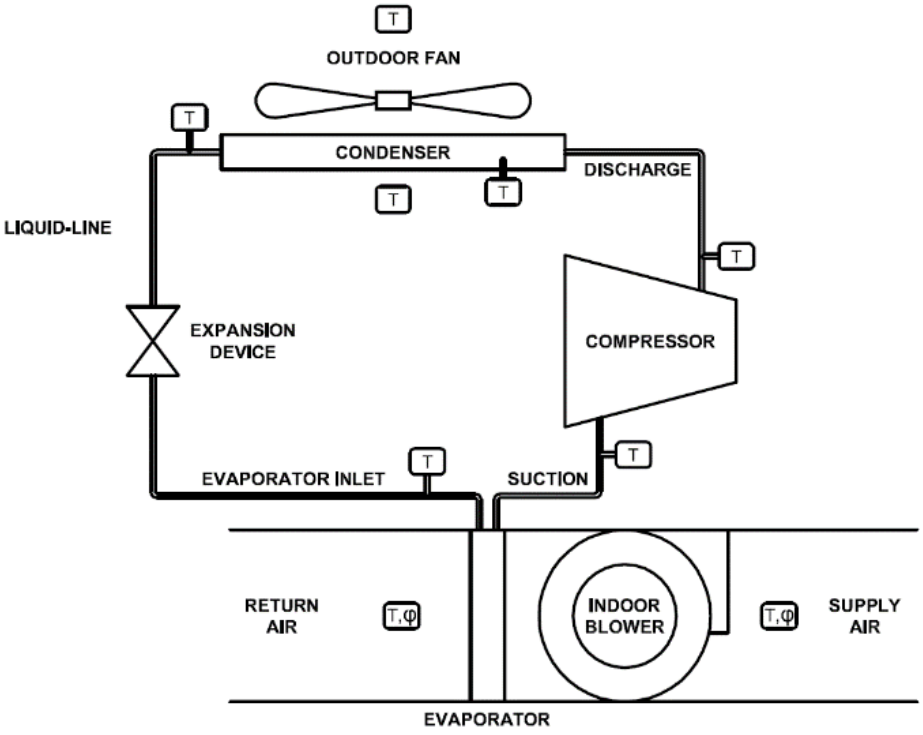

Figure 2.17. Locations of temperature (T) and relative humidity (ϕ) sensors used by RTU AFDD system.

Besides sensors and signal conditioning circuitry, the RTU AFDD data acquisition system requires a computational engine that can monitor the sensor outputs and perform the required mathematical transformations to the data and communicate the results with the VOLTTRON™ message bus. Typically, in HVAC applications, a low-cost microcontroller is used for data acquisition purposes. Most microcontrollers by themselves do not have the type of communications abilities that are required by the VOLTTRON™ communication protocol. While with enough time and effort, this could be implemented in some microcontrollers, an alternative type of device was selected to perform the required data analysis and communications.

The BeagleBone Black is a low-cost, open-source, community supported development platform with a TI Sitara™ ARM® Cortex A8 microprocessor that can run the Linux operating system. In other words, the BeagleBone Black is a computer with all essential components (microprocessor, RAM, hard drive, etc.) on a single circuit board. The first role of the BeagleBone Black in this application is to provide analog and digital interfaces between the data acquisition software and the required sensors installed on the RTU. Compared to other microcontrollers and microprocessors, application development using the BeagleBone Black is easier since many of these low-level hardware and software interfaces are provided out of the box. The second role of the BeagleBone Black is to support a run-time environment for the central VOLTTRON™ application as well as the embedded RTU AFDD software agents. The BeagleBone Black is not the only system capable of this; other development platforms are available with similar functionality. Development using the BeagleBone Black was selected since the TI Sitara™ microprocessor is widely available. Because of this, any work done with the prototype platform is almost directly translatable to any future (potentially lower cost) platform using a similar chipset.

The RTU AFDD electronics system designed was implemented using actual components, shown in Figure 2.18. The hardware selected for the prototype is typical and widely available. The system implemented is relatively low-cost when compared to similar data acquisition applications within the HVAC market – the final system could be built for approximately $120 USD. With consideration for economies of scale and optimized manufacturing process, significant cost reduction should be possible. While the through-hole prototyping system shown in Figure 2.18 can be readily manufactured, a printed circuit board design was also developed.

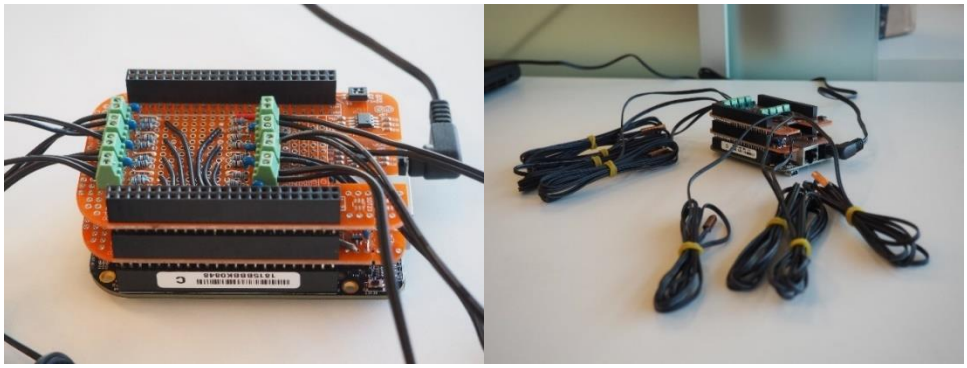

Figure 2.18. RTU AFDD system through-hole prototype implementation using pre-fabricated prototyping electronics boards.

2.7 Summary of AFDD System and Important Results

Previous research on automated fault detection and diagnosis (AFDD) for building systems has shown that identifying problems that cause degradation in system and equipment performance is technically feasible and has the potential to significantly reduce energy consumption. However, manufacturers have faced hurdles such as high instrumentation costs, lack of integration with other building system equipment, and uncertainty about savings that has prevented technologies from seeing market adoption. An AFDD system for packaged rooftop air conditioning equipment has been developed and implemented using a prototyping platform. AFDD algorithms developed by previous researchers have been implemented into the system

using open-source software. The low-cost electronics hardware that has been designed to implement the data acquisition, communications, and computational requirements of the system has also been implemented and tested. It is envisioned that this work can be adapted by future researchers as well as equipment manufacturers to reduce initial engineering and development effort.

Extensive psychrometric chamber testing has been conducted to assess the performance of the FDD system designed for rooftop units. Testing was performed over a wide range of outdoor and indoor ambient conditions while also injecting combinations of improper refrigerant charge, evaporator air flow reduction, and condenser air flow reduction faults. The virtual sensor outputs have been evaluated for a multistage RTU and variable-speed RTU. A statistical method for determining the probability of a fault being present in the system was able to identify improper charge levels, evaporator air flow reduction and condenser air flow reduction. Additionally, tests with multiple faults also showed good fault isolation, especially when impacts on total capacity and COP were significant. This indicates that the FDD system can identify problems that may cause significant additional energy consumption.

3. SIMPLIFIED FAULT IMPACT MODELS FOR UNITARY AIR CONDITIONING EQUIPMENT

3.1 Background and Motivation

Automated fault detection and diagnosis (AFDD) can be used to identify and isolate faults before significant impacts on comfort are noticed by building occupants or before degradations in performance cause equipment failure. AFDD systems can thus be used to monitor HVAC system operation and proactively schedule maintenance and service tasks when they are most economical. To make these decisions, the impacts of different types of faults must be characterized and modeled. This requires an understanding of how different faults impact the ability of air conditioning systems to maintain comfort, the potential to cause premature component failures and reduce equipment life, and the operating costs of the system.

The purpose of the following chapter is to develop models for characterizing the fault impacts of different DX equipment faults that could be implemented within an AFDD system so that more economical service decisions can be made. Before getting into the details of these models, a survey of previous work related to understanding and modeling of equipment fault impacts is presented. This review highlights major works related to experimental testing of equipment subjected to artificially imposed faults. In addition, it identifies the work of previous researchers that has developed models and other methodologies for estimating the impacts faults have on equipment performance and operating costs.

Following this literature review, models that can be used to predict the normal performance of DX systems are described. For fixed-speed equipment, a widely-used model based on the apparatus dew point / bypass factor method is described that is suitable for predicting total cooling capacity, system efficiency, SHR at off-design ambient conditions and air flow rates. A

previously developed modification of this model for variable-speed equipment is also presented along with a comparison to measured performance obtained using psychrometric chamber testing.

After results of the normal performance models are evaluated, the underlying methodology used to characterize the total impacts different faults have on equipment performance is described. The methodology that is described uses measurements of actual equipment performance (based on virtual sensors) and models of normal equipment performance to assess the impacts faults have on total cooling capacity, system COP, sensible heat ratio (SHR), required run-time, and total energy consumption. The methodology illustrates the importance of virtual sensors used in the AFDD system. Virtual sensors are not only used for detection and diagnosis, but also for determining the significance faults have on the operation of a DX system.

To this point, work on fault impact estimation has focused on the total, combined impact of faults on equipment performance – even when multiple faults affect a system simultaneously. At the close of this chapter, the development of fault impact models that can be used to predict the relative impacts of different faults that occur simultaneously is presented. In this work, these models will be referred to as *fault impact isolation* models and will have the implicit goal to estimate the benefit of servicing different faults. Since multiple faults may affect equipment simultaneously over time, evaluating which faults affect performance most significantly is important. For example, a RTU may be slightly undercharged and have a significantly fouled condenser. An automated FDD tool should have sufficient intelligence to ignore faults that do not result in significant impacts while recommending service for faults that do.

To develop these models, a more detailed physics-based fault impact model for DX equipment has been used. This model is then used to develop simplified models that can be used to estimate the relative impacts of individual faults. These simplified models can be used to

predict the total capacity and efficiency impacts of different faults relative to the performance of a fault-free system.

3.2 RTU Fault Impact Estimation Literature Review

The impacts of faults on the operation of direct expansion (DX) equipment have been studied in laboratory settings for at least two decades. One of the earliest experimental studies of these impacts on DX air conditioning equipment can be attributed to Breuker and Braun [44]. In that work the impacts of several faults, including refrigerant charge leakage (undercharged system), condenser fouling, evaporator fouling, compressor valve leakage, and liquid line restriction, were studied with respect to how different temperatures of the refrigeration cycle were affected. Breuker and Braun also studied the impacts of faults on equipment cooling capacity and *COP* [44]. The temperatures that were studied: evaporation temperature, compressor suction temperature, compressor discharge temperature, condensing temperature, liquid line temperature, and the temperatures of air entering and leaving the evaporator and condenser coils showed different behaviors for different types of faults. These responses were then used to assess the performance of an FDD tool applied to a 3-ton R22 RTU with a fixed orifice expansion valve [17].

Breuker and Braun also published findings from a survey of the most common faults that affect packaged RTUs in the field [44]. Of the faults that affect the refrigeration cycle performance, the most common faults were refrigerant leaks, condenser fouling, evaporator fouling, compressor problems, and liquid line restrictions. The most common problems overall were controls errors and electrical problems – which were also found to cause the most total service costs. Of the refrigerant cycle faults, problems with compressors tended to have the highest total costs. The transient impacts of different faults were also studied during this work,

however no difference between the normal start up transient and fault start up transients were found [17].

Harms *et al*. tested a 5-ton R22 RTU with a thermostatic expansion valve at different levels of charge to compare the accuracy of different void fraction models with respect to charge inventory modeling of the vapor compression cycle [45]. The result of the testing showed that the choice of void fraction model had negligible impact on the accuracy of the estimation of system performance when proper charge tuning was used.

Grace *et al*. studied the impacts of refrigerant charge level on the performance of a built-up refrigeration system using R404A as the working fluid and thermostatic expansion valve [30]. The performance of the system was found to be strongly dependent on the amount of refrigerant contained in the system. The experimental data were used to develop an on-line fault detection system.

Shen *et al*. studied several DX air conditioning systems, including a 5-ton R407C RTU with fixed orifice, a 3-ton R410A RTU with fixed orifice, a 3-ton R410A split-type air conditioner with fixed orifice, and a 3-ton R410A split-type air conditioner with thermostatic expansion valve [46]. Evaporator fouling, condenser fouling, and several levels of refrigerant charge were tested with the impacts on system performance measured. The experimental results were primarily used to develop an improved charge tuning method for vapor compression cycle models as well as improve the accuracy of an existing vapor compression model at off-design conditions.

Armstrong *et al*. studied the impacts of different faults on electrical measurements of a RTU [20]. The study found that different faults lead to unique fault signatures that present

themselves in the transient electrical characteristics of the system. The data were used to develop a non-intrusive fault FDD methodology based solely on electrical measurements.

Kim *et al*. studied the impacts of improper charge levels, condenser fouling, evaporator fouling, compressor valve leakage, liquid line restriction, and the presence of non-condensable gases in the refrigerant circuit on a 2.5-ton R410A split-type air conditioner with thermostat expansion valve [25]. The study was performed to determine temperatures that were significantly affected by different faults for the development of an FDD methodology.

Palmiter *et al*. measured the seasonal performance impacts of indoor air flow rate and refrigerant charge on the 3-ton R410A split-type heat pump with thermostatic expansion valve [47]. Some cycling tests were also performed during this testing and it was found that COP decreases with reductions in indoor air flow rate and increasing refrigerant charge level.

Li and Braun developed a methodology to assess the economic impact of operating DX systems with faults based on the combined effects of total cooling capacity, COP, and sensible heat ratio (SHR) impacts [14]. The model is based on actual measured performance of the system and typical performance of systems without faults. In the work, a case study is presented which evaluated the economic impacts of different faults that were artificially applied to an existing air conditioner under different operating conditions.

Kim and Braun studied the impacts of refrigerant charge on the performance of several split-type heat pumps and air conditioners [23]. Laboratory measurements were used to estimate the seasonal performance impacts of improper refrigerant charge levels and associated operating cost impacts.

Cheung and Braun used experimental data collected from several previous studies to develop a component-level fault impact inverse model for vapor compression air conditioning

equipment [48, 49]. The model can simulate the performance impacts of combinations of improper refrigerant charge level, condenser fouling, evaporator fouling, compressor valve leakage, liquid line restriction, and presence of non-condensable gas in the refrigerant circuit at different operating conditions. The resulting model was validated using experimental data collected from several RTUs and split-type air conditioners.

Yuill and Braun developed a methodology to evaluate the performance of FDD tools for HVAC equipment [10]. In this study, different criteria measuring the performance of FDD systems, e.g. fault detection rate, false alarm rate, misdiagnosis rate, etc. were used to evaluate several generic FDD tools that can be applied to systems in the field. Yuill and Braun used data collected from several RTUs and split-type air conditioners collected in laboratory psychrometric chambers under different fault conditions. Rather than assess the performance of FDD systems based on the severity of a different fault, e.g. the percent deviation in charge level, Yuill and Braun evaluated FDD performance based on how much an impact a fault has on cooling capacity or COP. The authors reason that this leads to a fairer assessment of different FDD tools since faults that have negligible or small impacts can be ignored by some FDD tools.

3.3 Simplified Models for Normal Equipment Performance

Prior to estimating the impacts that certain faults may have on system performance, the baseline performance of the system must be determined. To estimate the base-line performance, semi-empirical models of fixed-speed and variable-speed RTUs are implemented and tested. The models are based on measurements of the outdoor temperature and condition of the air entering the evaporator coil. For variable-speed equipment, an additional correlation accounting for changes in the compressor frequency is used to capture changes in equipment performance.

3.3.1 Semi-Empirical Models for Fixed-Speed RTU Performance

One commonly used semi-empirical model for direct expansion air conditioning equipment approximates the total cooling capacity, the total power consumption, and the sensible heat ratio by adjusting rated system performance to account for differences in ambient conditions and air flow rate. In this model, the total cooling capacity of a system is determined using

$$\dot{Q}_{cool,normal} = f_T \cdot f_{\dot{V}} \cdot \dot{Q}_{cool,rated} \tag{3.1}$$

where $\dot{Q}_{cool,rated}$ is the total cooling capacity at some rating condition, f_T is a correction factor accounting for differences in return air and outdoor air temperatures, and $f_{\dot{V}}$ is a correction factor accounting for differences in the volumetric air flow rate through the evaporator coil. The ambient correction factor, f_T, is often given a biquadratic polynomial

$$f_T = a_0 + a_1 \cdot T_{OD} + a_2 \cdot B_{ID} + a_3 \cdot T_{OD}^2 + a_4 \cdot T_{OD} \cdot B_{ID} + a_5 \cdot B_{ID} \tag{3.2}$$

where T_{OD} is the outdoor dry bulb temperature (entering the condenser coil), B_{ID} is the return air wet bulb temperature (entering the evaporator coil), and $a_{0\ldots5}$ are empirical parameters determined using regression. The evaporator volumetric air flow rate correction factor, $f_{\dot{V}}$, is often a linear function given by

$$f_{\dot{V}} = b_0 + b_1 \cdot \left(\frac{\dot{V}_{actual}}{\dot{V}_{rated}} \right) \tag{3.3}$$

where $\dot{V}_{actual}$ is the actual volumetric air flow rate, $\dot{V}_{rated}$ is the evaporator air flow rate of the system at the rating condition, and $b_{0,1}$ are empirical parameters determined using regression for a particular system.

To accurately predict the moisture content of the air leaving the evaporator coil, the model must properly predict the fraction of total cooling rate that results in sensible heat transfer and

latent heat transfer over the typical range of equipment operating conditions. It is widely held that the total cooling capacity delivered by a DX coil operating with moisture removal is a function of the ambient dry-bulb temperature and wet-bulb of the air entering the evaporator. For these cases, often termed wet coil conditions, Equation (3.1) estimates the total cooling capacity of the system with sufficient accuracy to estimate the leaving air enthalpy, but does provide separate estimates of leaving air temperature and humidity. In addition, estimates using Equation (3.1) need to be adjusted when no moisture is removed though the cooling coil process.

One commonly used approach for addressing these aspects of cooling coil performance is described by the apparatus dew-point/bypass factor (ADP/DP) methodology that has been implemented in several equipment performance and energy calculation programs [50, 51]. This approach is like the NTU-effectiveness calculation procedure used for sensible heat transfer calculations in air-to-refrigerant heat exchanger analysis where $C_{min}/C_{max} = 0$ [52]. In this analysis, the bypass factor is defined as

$$BF = 1 - \varepsilon = \exp\left(-NTU\right) \tag{3.4}$$

and

$$\frac{NTU}{NTU_{rated}} = \left(\frac{\dot{V}_{actual}}{\dot{V}_{rated}}\right)^{n} \tag{3.5}$$

where n is an empirical parameter related to the design of the heat exchanger.

Using the predicted total cooling capacity from Equation (3.1) and the evaporator air mass flow rate, the enthalpy of air leaving the cooling coil can be determined

$$h_{SA} = h_{ID} - \frac{\dot{Q}_{cool}}{\dot{m}_a} \tag{3.6}$$

where h_{ID} is the enthalpy of air entering the evaporator and $\dot{m}_a$ can be determined using the volumetric air flow rate and density calculated using the known dry-bulb and wet-bulb temperature of the entering air. The apparatus dew point condition can now be determined using the bypass factor,

$$h_{ADP} = h_{ID} - \frac{h_{ID} - h_{SA}}{1 - BF}. \tag{3.7}$$

The enthalpy of the air at the apparatus dew point condition can be used to calculate the humidity ratio (the corresponding saturation conditions). With these known, the sensible heat ratio of the cooling process can be determined

$$SHR = \frac{\mathbf{h}\left(T_{ID}, \omega_{ADP}\right) - h_{ADP}}{h_{ID} - h_{ADP}} \tag{3.8}$$

where **h** is a function that returns humid air enthalpy using psychrometric property calculations.

The approach that has been presented to this point can be used to calculate the total cooling capacity delivered by DX cooling equipment, as well as the sensible/latent performance of the coil at off-design conditions. For wet coil conditions, this procedure is sufficient by itself in predicting the *SHR* maintained by the cooing coil. For dry coil conditions where there is no moisture removal, a widely used iterative procedure can be used that adjusts a "virtual entering wet-bulb" used in Equation (3.1) to modify the predicted outputs so that $SHR = 1$.

In a similar manner to modeling the total cooling capacity of the system, the total power consumption of the system can be estimated using

$$\dot{W}_{total,normal} = g_T \cdot g_{\dot{V}} \cdot \dot{W}_{total,rated} \tag{3.9}$$

where $\dot{W}_{total,rated}$ is the total power consumption of the system at the rating condition, g_T is a correction factor accounting for differences in ambient operating conditions, and $g_{\dot{V}}$ is a

correction factor accounting for the effect of evaporator volumetric air flow rate. The ambient correction factor is given by

$$g_T = c_0 + c_1 \cdot T_{OD} + c_2 \cdot B_{ID} + c_3 \cdot T_{OD}^2 + c_4 \cdot T_{OD} \cdot B_{ID} + c_5 \cdot B_{ID}^2 \tag{3.10}$$

where T_{OD} is the outdoor dry bulb temperature, B_{ID} is the return air wet bulb temperature, and $c_{0..5}$ are empirical parameters determined using regression. The air flow rate correction factor is given by

$$g_{\dot{V}} = d_0 + d_1 \cdot \left(\frac{\dot{V}_{actual}}{\dot{V}_{rated}} \right) \tag{3.11}$$

where $\dot{V}_{actual}$ is the actual volumetric air flow rate, $\dot{V}_{rated}$ is the evaporator air flow rate of the system at the rating condition, and $d_{0,1}$ are empirical parameters determined using regression for a particular system. The coefficient of performance of the system at different operating conditions can be calculated using Equations (3.1) and (3.9),

$$COP_{normal} = \frac{\dot{Q}_{cool,normal}}{\dot{W}_{total,normal}}. \tag{3.12}$$

Using the methodology to determine normal equipment performance at different operating conditions outlined in Equations (3.1) through (3.12), the empirical parameters used in the different models were tuned using experimental data collected from a fixed-speed RTU with different types of condenser coils and expansion devices.

3.3.2 Semi-Empirical Models for Variable-Speed RTU Performance

Equations (3.1) through (3.12) will not estimate the performance of variable-speed DX systems accurately when operating at reduced load with reduced compressor speeds, since the model assumes compressor speed is fixed. Nyika *et al*. developed a methodology for modeling

the performance of variable-speed heat pumps in both cooling and heating modes of operation at different operating conditions [53]. Additionally, Cheung and Braun developed a methodology for mapping the performance of variable-speed heat pumps in heating operation with defrost [54]. A more detailed, component-based gray-box model of a ductless multi-split heat pump system was developed by Cheung and Braun [55].

One adaptation to the model described in Section 3.3.1 has been proposed by Cai and Braun to account for different compressor speeds in the total cooling capacity prediction using the part-load ratio of the system [56]. Part-load ratio is not straightforward to measure in an online application – it requires a model for the full load total capacity as well as a measurement of the actual capacity of the system. This requirement can be circumvented by adapting the model developed by Cai and Braun to use the compressor speed,

$$\dot{Q}_{cool,normal} = j_T \cdot j_{\dot{V}} \cdot j_\Omega \cdot \dot{Q}_{cool,rated} \tag{3.13}$$

where $\dot{Q}_{cool,rated}$ is the total cooling capacity at some rating condition, j_T is a correction factor accounting for differences in return air and outdoor air temperatures, $j_{\dot{V}}$ is a correction factor accounting for differences in the volumetric air flow rate through the evaporator coil, and j_Ω is a correction factor for differences in compressor speed . The ambient correction factor, j_T, is given a biquadratic polynomial

$$j_T = e_0 + e_1 \cdot T_{OD} + e_2 \cdot B_{ID} + e_3 \cdot T_{OD}^2 + e_4 \cdot T_{OD} \cdot B_{ID} + e_5 \cdot B_{ID} \tag{3.14}$$

where T_{OD} is the outdoor dry bulb temperature (entering the condenser coil), B_{ID} is the return air wet bulb temperature (entering the evaporator coil), and $e_{0\ldots5}$ are empirical parameters determined using regression. The evaporator volumetric air flow rate correction factor, $j_{\dot{V}}$, is given by

$$f_{\dot{V}} = m_0 + m_1 \cdot \left(\frac{\dot{V}_{actual}}{\dot{V}_{rated}} \right) + m_2 \cdot \left(\frac{\dot{V}_{actual}}{\dot{V}_{rated}} \right)^2 \tag{3.15}$$

where $\dot{V}_{actual}$ is the actual volumetric air flow rate, $\dot{V}_{rated}$ is the evaporator air flow rate of the system at the rating condition, and $m_{0\ldots2}$ are empirical parameters determined using regression for a particular system. The compressor speed correction factor is given by

$$j_{\Omega} = n_0 + n_1 \cdot \Omega + n_2 \cdot \Omega^2 + n_3 \cdot \Omega^3 \tag{3.16}$$

where Ω is the compressor speed and $n_{0\ldots3}$ are empirical parameters determined using regression.

In a similar manner to modeling the total cooling capacity of the system, the total power consumption of the system can be estimated using

$$\dot{W}_{total,normal} = k_T \cdot k_{\dot{V}} \cdot k_{\Omega} \cdot \dot{W}_{total,rated} \tag{3.17}$$

where $\dot{W}_{total,rated}$ is the total power consumption of the system at the rating condition, k_T is a correction factor accounting for differences in ambient operating conditions, $k_{\dot{V}}$ is a correction factor accounting for the effect of evaporator volumetric air flow rate, and k_{Ω} is a correction factor for differences in compressor speed. The ambient correction factor is given by

$$k_T = l_0 + l_1 \cdot T_{OD} + l_2 \cdot B_{ID} + l_3 \cdot T_{OD}^2 + l_4 \cdot T_{OD} \cdot B_{ID} + l_5 \cdot B_{ID}^2 \tag{3.18}$$

where T_{OD} is the outdoor dry bulb temperature, B_{ID} is the return air wet bulb temperature, and $l_{0\ldots5}$ are empirical parameters determined using regression. The air flow rate correction factor is given by

$$k_{\dot{V}} = q_0 + q_1 \cdot \left(\frac{\dot{V}_{actual}}{\dot{V}_{rated}} \right) + q_2 \cdot \left(\frac{\dot{V}_{actual}}{\dot{V}_{rated}} \right)^2 \tag{3.19}$$

where $\dot{V}_{actual}$ is the actual volumetric air flow rate, $\dot{V}_{rated}$ is the evaporator air flow rate of the system at the rating condition, and $q_{0,1}$ are empirical parameters determined using regression for a particular system. The compressor speed correction factor is given by

$$k_{\Omega} = r_0 + r_1 \cdot \Omega + r_2 \cdot \Omega^2 + r_3 \cdot \Omega^3 \tag{3.20}$$

where Ω is the compressor speed and $r_{0\ldots3}$ are empirical parameters determined using regression. To estimate the sensible heat ratio of the air conditioner, the bypass factor method described previously in Equations (3.2) through (3.8) can be used.

The methodology outlined in Equations (3.13) through (3.20) to determine normal equipment performance at different operating conditions, was implemented using empirical parameters tuned using experimental data collected from a variable-speed RTU. A variable-speed RTU with 5-ton rated capacity was installed in a pair of psychrometric chambers and instrumented with laboratory grade sensors using recommended standards for unitary equipment testing [41-43, 57-60].

Steady-state testing at constant indoor and outdoor temperatures and humidity was conducted over a range of conditions representative of typical equipment operation. These conditions, including indoor room temperature, indoor room relative humidity, and outdoor room temperature are shown in Table 3.1. The indoor fan speed and discharge air temperature setpoints were also maintained during each test. The operating points of the RTU are also shown in Table 3.1. The psychrometric chambers and RTU were operated until steady-state was reached before transitioning to subsequent tests. Steady-state was determined using a steady-state filter which measured temporal changes in total cooling capacity, sensible cooling capacity, and power consumption of the RTU. When negligible changes in any of the variables occurred

for a period of two minutes, the current test was concluded and the subsequent tests at the next operating conditions was started.

Table 3.1. Experimental test conditions for variable-speed RTU normal performance model development and evaluation. Experimental variables were indoor dry bulb temperature, indoor relative humidity, outdoor dry bulb temperature, indoor fan speed, and discharge air setpoint.

T_{ID} [°F]	ϕ_{ID} [%]	T_{OD} [°F]	γ_{IDF} [%]	$T_{SA,SP}$ [°F]	Dry/Wet Coil
80	25	95	100	55	Dry
80	25	95	80	55	Dry
80	25	95	50	55	Dry
80	25	95	30	55	Dry
80	45	95	100	55	Wet
80	45	95	80	55	Wet
80	45	95	50	55	Wet
80	45	95	30	55	Wet
80	55	95	100	55	Wet
80	55	95	80	55	Wet
80	55	95	50	55	Wet
80	55	95	30	55	Wet
80	25	86	100	55	Dry
80	25	86	80	55	Dry
80	25	86	50	55	Dry
80	25	86	30	55	Dry
80	45	86	100	55	Wet
80	45	86	80	55	Wet
80	45	86	50	55	Wet
80	45	86	30	55	Wet
80	55	86	100	55	Wet
80	55	86	80	55	Wet
80	55	86	50	55	Wet
80	55	86	30	55	Wet

The empirical parameters of Equations (3.13) and (3.17) were fit using ordinary least squares regression with the experimental data collected. Because the semi-empirical models defined by Equations (3.13) and (3.17) use an iterative algorithm to adjust the evaporator air entering wet bulb to constrain the model output to $SHR \leq 1$, only wet-coil data was used to determine the empirical coefficients. Additionally, it was not possible to directly measure the variable-speed compressor speed using available instrumentation. During testing, it was noticed that the

compressor speed and outdoor fan speed was coincidently controlled – i.e. if the compressor speed was 50%, the condenser fan was 50%. Since it was possible to measure the outdoor fan power, the compressor speed was estimated based on fan affinity relationships

$$\Omega_{comp} = \Omega_{odf} = \Omega_{odf,rated}\left(\frac{\dot{W}_{odf}}{\dot{W}_{odf,rated}}\right)^3 \tag{3.21}$$

where Ω_{comp} is the compressor loading percent, Ω_{odf} is the outdoor fan speed, $\Omega_{odf,rated}$ is the outdoor fan speed at the rating condition, $\dot{W}_{odf}$ is the outdoor fan power consumption, and $\dot{W}_{odf,rated}$ is the power consumption of the outdoor fan at the rating condition. To determine the rated values, the system was operated at a condition resulting in full load – which resulted in the fan operating at 100% speed with the maximum power.

The predicted total cooling capacity of the variable-speed RTU using Equation (3.13) is shown with the measured total cooling capacity in Figure 3.1. The semi-empirical model for total cooling capacity agreed to within 5% of the measured total cooling capacity for all test conditions. Additionally, no clear correlation between the prediction error and the sensible heat ratio was observed. This provides some evidence that using only wet coil data to tune the empirical parameters of the model did not result in poor performance when used to predict dry-coil operation. It should also be noted that the part load ratio of the equipment ranged from 100% down to approximately 50%. This indicates the model can predict total cooling capacity without significant errors over the range of continuous system operation.

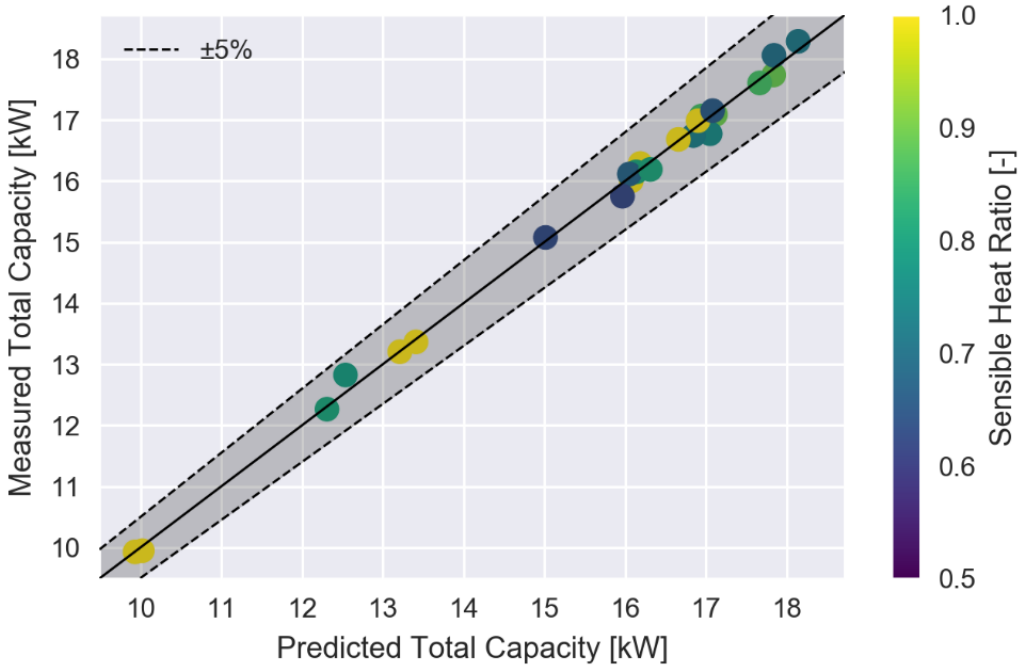

Figure 3.1. Prediction accuracy of semi-empirical model for total cooling capacity of a variable-speed RTU operating without faults over a range of ambient conditions and loads.

The total system power consumption prediction of the semi-empirical model given by Equation (3.17) is compared with experimentally measured power consumption in Figure 3.2. The results show that the semi-empirical model can be used to predict power consumption of the variable-speed RTU with small deviations over the range of operating conditions tested. No dependence between the prediction error and the sensible heat ratio was observed for the experimental test data collected. This includes test cases that had no moisture removal – dry coil cases. For dry coil cases, the iterative algorithm based on bypass factor analysis is used to adjust the entering evaporator wet bulb that restricts $SHR = 1$. This is noteworthy since dry coil cases were not included in the data set used to fit the empirical parameters using least squares regression.

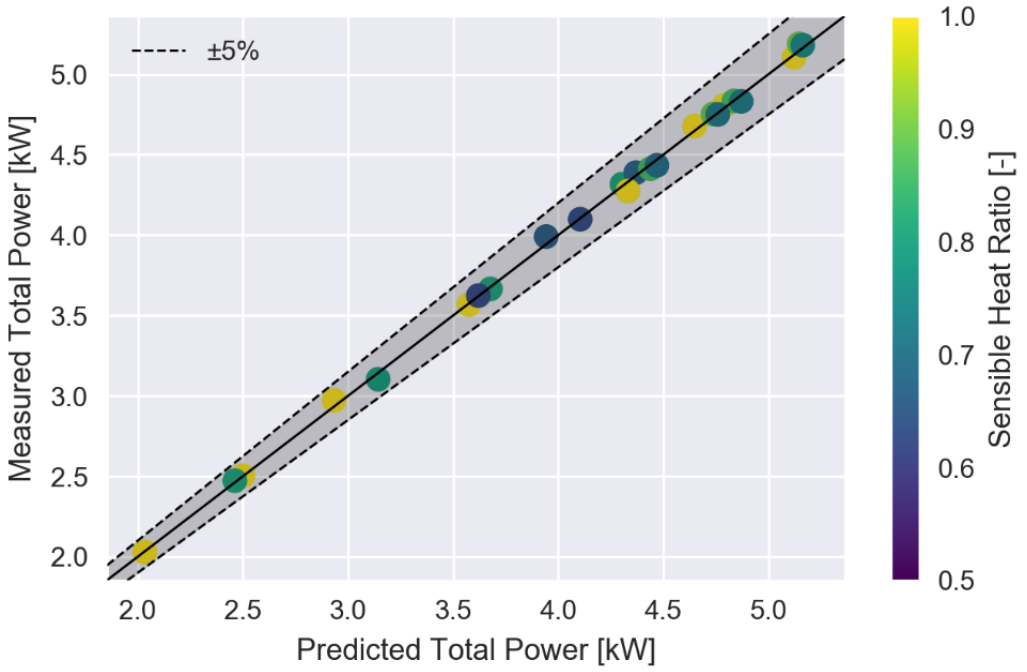

Figure 3.2. Prediction accuracy of semi-empirical model for total power consumption of a variable-speed RTU operating without faults over a range of ambient conditions and loads.

The coefficient of performance of each experimental test case was predicted using predicted capacity and power. These predictions, along with the *COP* derived from measured total cooling capacity and total power consumption, are compared in Figure 3.3. It is observed from the results that the prediction errors for all the test cases fall within 5% of the measured *COP*. From an FDD perspective, smaller modeling errors lead to more sensitive detection rates. These results show that deviations in *COP* greater than 5% could be detected with reasonable expected probability provided the actual *COP* can be accurately measured or estimated.

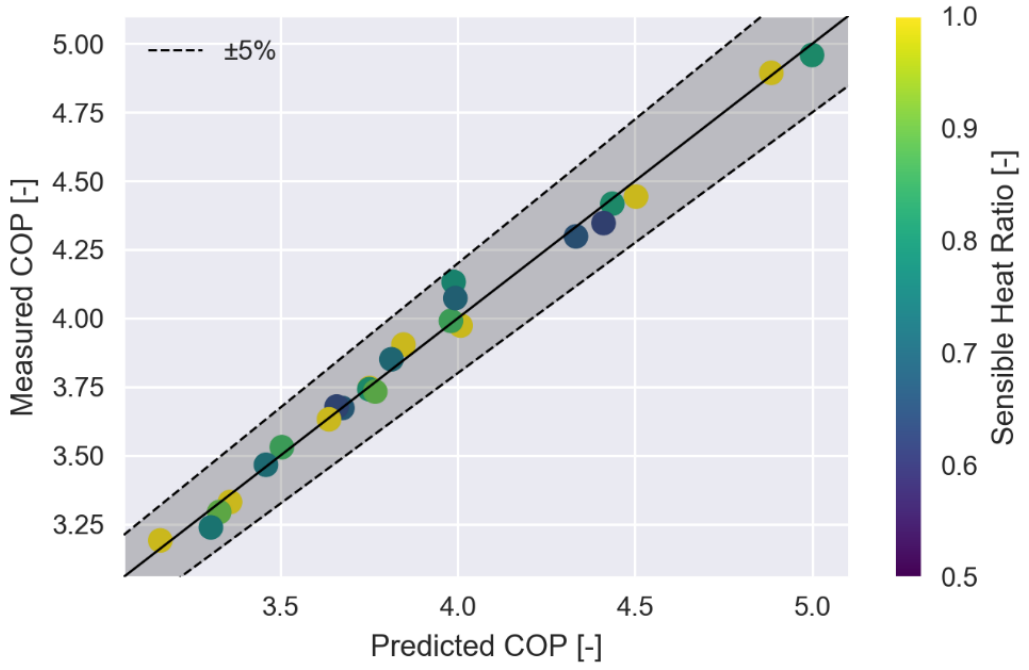

Figure 3.3. Prediction accuracy of semi-empirical model for coefficient of performance of a variable-speed RTU operating without faults over a range of ambient conditions and loads.

Using the total capacity prediction, the sensible capacity was determined using Equations (3.6) and (3.8). This predicted sensible capacity is compared with the measured sensible capacity in Figure 3.4. These results show that the predicted sensible capacity of the variable-speed RTU falls within 5% of the measured sensible capacity for all test cases.

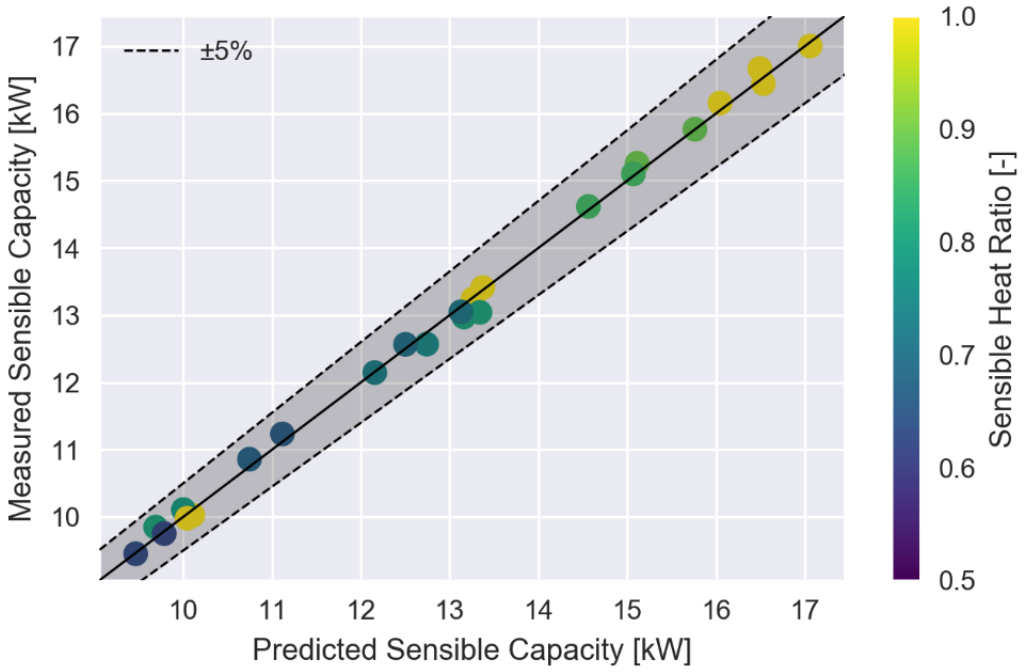

Figure 3.4. Prediction accuracy of semi-empirical model for sensible cooling capacity of a variable-speed RTU operating without faults over a range of ambient conditions and loads.

The predicted sensible heat ratio is compared with the measured equipment sensible heat ratio in Figure 3.5 for all test cases. The prediction errors for all experimental tests cases are within 5%. It should be noted that a relatively large range of equipment sensible heat ratios were tested – from approximately 60% to 100%. The results show that the iterative algorithm based on bypass factor can be applied to the variable-speed RTU with good results.

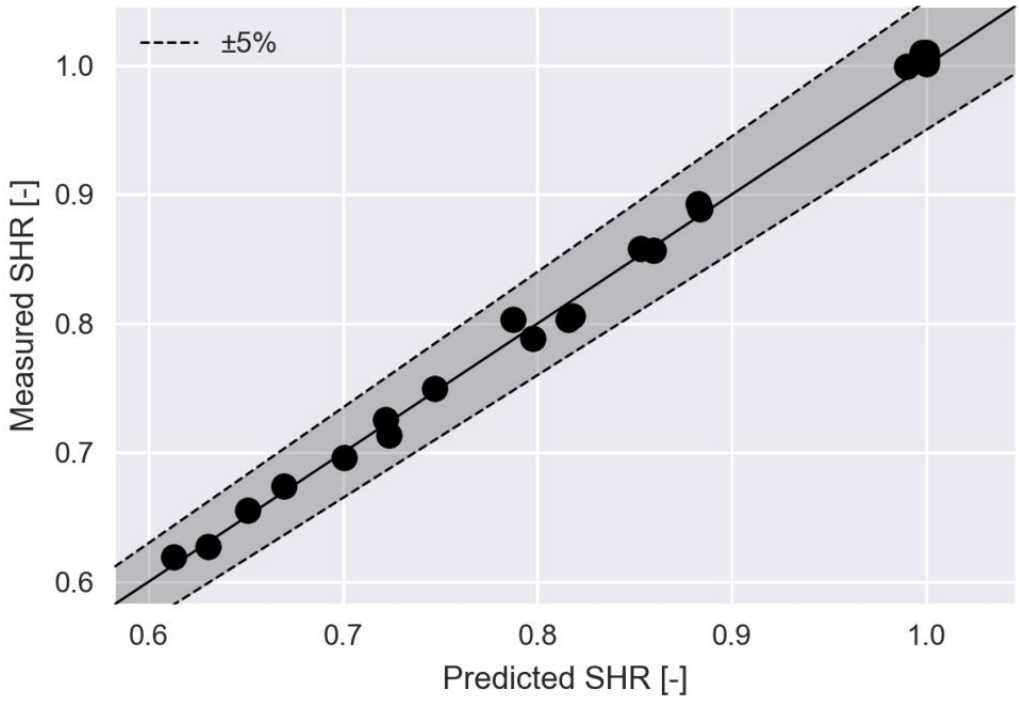

Figure 3.5. Prediction accuracy of semi-empirical model for sensible heat ratio of a variable-speed RTU operating without faults over a range of ambient conditions and loads.

Semi-empirical models tuned using linear regression can suffer from overfitting problems when careful model evaluation is not performed. To evaluate the total capacity and total power consumption predictions, the model was applied to a generated data set designed to test the interpolation and extrapolation performance. Holding the entering evaporator air wet bulb and compressor speed constant at 64 °F and 100% respectively, the predicted total capacity and predicted total power consumption were calculated over range of outdoor air temperatures and evaporator air flow rates, shown in Figure 3.6. The results generated by the model are smooth without any unrealistic inflection points or oscillatory behavior. From a physical perspective, the total cooling capacity predictions show expected trends. As outdoor air temperature and evaporator air flow rate increase, cooling capacity decreases and power consumption increases.

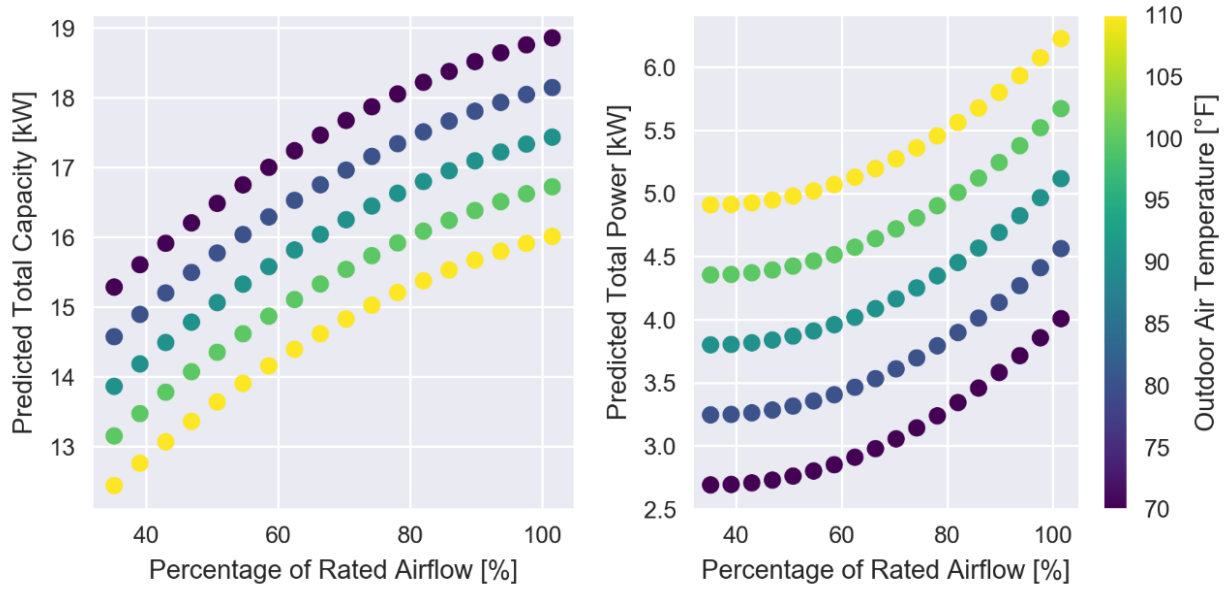

Figure 3.6. With return air wet bulb (63 °F) and compressor speed (100%) fixed, the predicted total capacity and predicted total power consumption are shown for varying evaporator air flow rates and outdoor air dry bulb temperatures.

For the same conditions, the predicted *COP* and *SHR* were generated by the semi-empirical normal performance model, shown in Figure 3.7. These results show that *COP* can be optimized by adjusted indoor air flow rate which also depends on the ambient condition. The *SHR* response is relatively linear with both evaporator air flow rate and outdoor air temperature. As air flow rate increases, *SHR* is shown to increase. This is an expected result since the evaporator saturation temperature tends to increase at higher air flow rates.

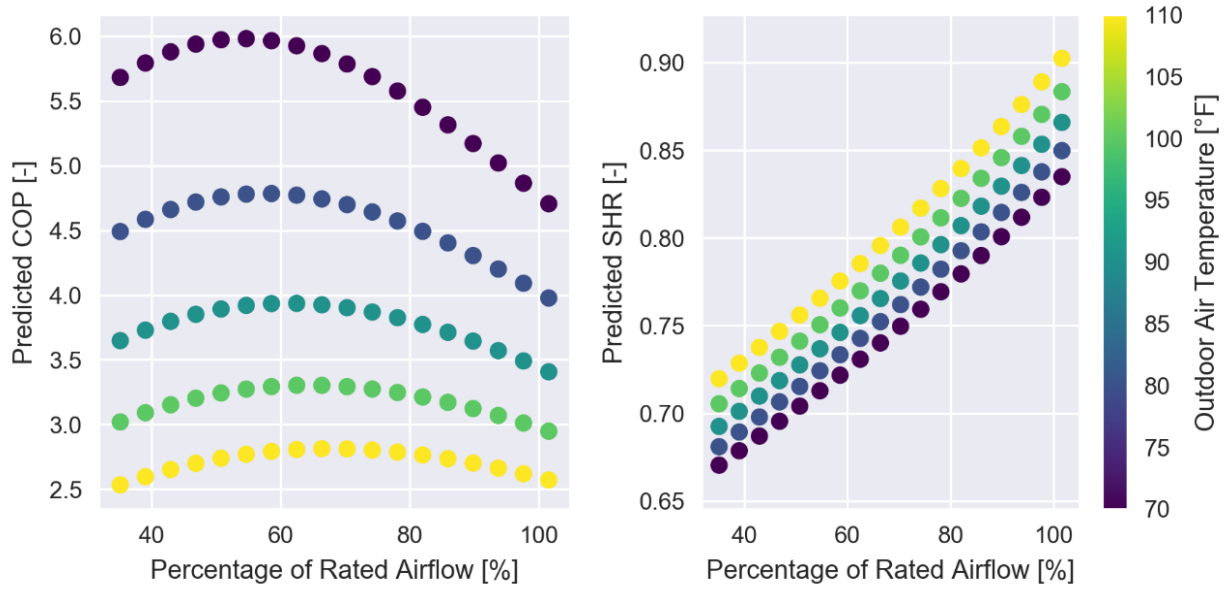

Figure 3.7. With return air wet bulb (63 °F) and compressor speed (100%) fixed, the predicted *COP* and predicted *SHR* are shown for varying evaporator air flow rates and outdoor air dry bulb temperatures.

In a similar manner, the semi-empirical model was used to predict the total cooling capacity and total power consumption over a range of compressor speeds and evaporator air entering wet bulbs for constant outdoor air temperature (95 °F) and air flow rate (100% of rated air flow). The responses shown in Figure 3.11 show reasonable results – capacity and power consumption increase with compressor speed and entering wet bulb. It should be noted that total power consumption has relatively weak dependence on entering wet bulb over the range of test cases.

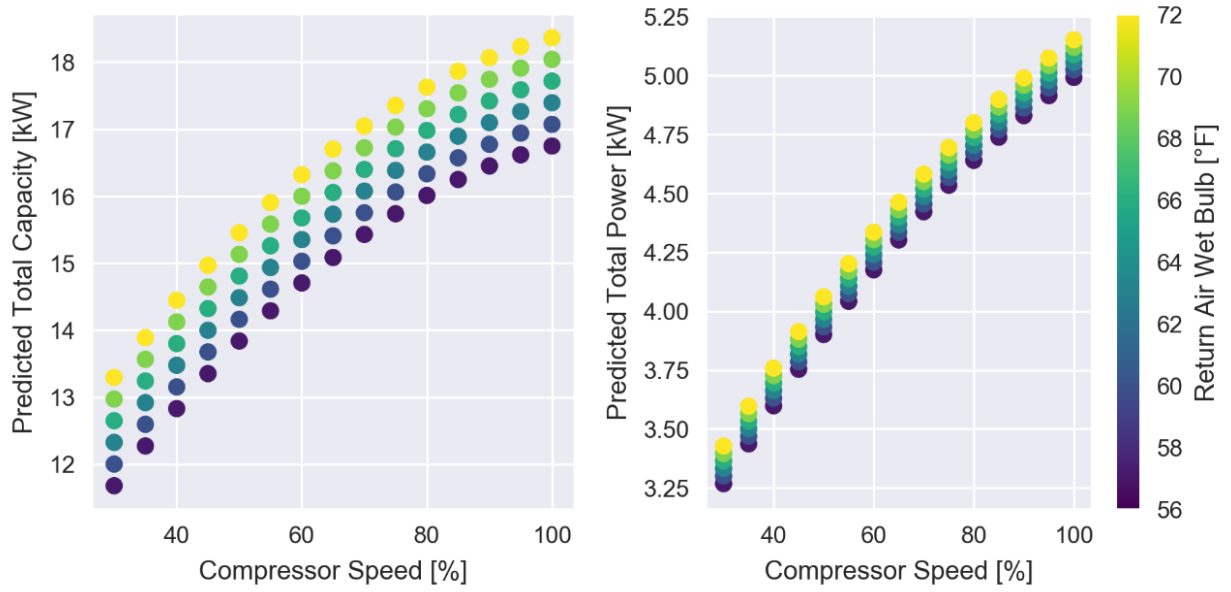

Figure 3.8. With outdoor air temperature (95 °F) and evaporator air flow rate (100%) fixed, the predicted total capacity and predicted total power consumption are shown for varying entering evaporator air wet bulbs and compressor speeds.

The semi-empirical model outputs for the *COP* and *SHR* of the system have been plotted in Figure 3.9 for varying compressor speeds and evaporator entering air wet bulbs. The results show that COP decreases with increasing compressor speed. The *SHR* responses show the model limits the maximum *SHR* to 1. The *SHR* output does show some unexpected behavior for increasing compressor speeds: the *SHR* is not monotonically decreasing. As compressor speed decreases, the discharge air temperature should increase and generally result in higher *SHR*. This intuition is not perfectly captured by the model.

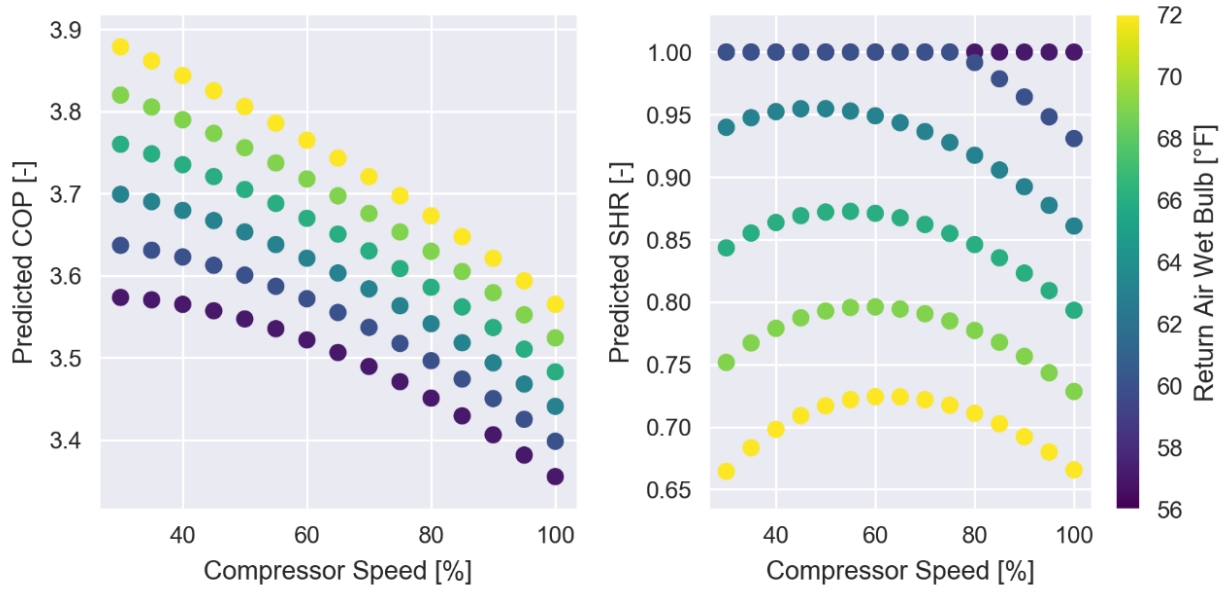

Figure 3.9. With outdoor air temperature (95 °F) and evaporator air flow rate (100%) fixed, the predicted *COP* and predicted *SHR* are shown for varying entering evaporator air wet bulbs and compressor speeds.

A semi-empirical model has been applied to a variable-speed RTU and evaluated using experimental data over a range of typical equipment operating conditions. The model requires measurements of the entering evaporator air wet bulb, outdoor air dry bulb temperature, evaporator air flow rate, and compressor speed. The model was able to predict the total cooling capacity, the total power consumption, the *COP*, the sensible cooling capacity, and the *SHR* with 5% of the experimentally measured values for all test cases. Because the compressor speed could not be measured directly, it was approximated using measurements of the condenser fan speed based on the fan affinity laws. This was possible since the condenser fan speed and the compressor speed were controlled at the same percentage by the system controller. It can be concluded from the results that the semi-empirical model used can be used to estimate normal equipment performance by a fault detection and diagnostics algorithm. Additionally, the model is useful for estimating the impact of faults on different equipment performance.

3.4 Estimating Actual Equipment Performance Using Virtual Sensors

The actual cooling capacity can be estimated using virtual cooling capacity sensors that have been previously developed by Li and Braun [6]. The virtual cooling capacity sensor is based on a refrigerant-side energy balance of the evaporator coil where the mass flow rate of refrigerant and the inlet and outlet refrigerant enthalpy are known. Equation (3.22) expresses the virtual cooling capacity sensor, $\dot{Q}_{cool,actual}$, in terms of a virtual refrigerant mass flow rate estimate, $\dot{m}_{r,virtual}$,

$$\dot{Q}_{cool,actual} = \dot{m}_{r,virtual} \cdot \left(h_{ero} - h_{eri} \right) \tag{3.22}$$

where h_{ero} is the enthalpy of the refrigerant leaving the evaporator and h_{eri} is the enthalpy of refrigerant entering the evaporator. The enthalpy of the entering and leaving refrigerant can be inferred using thermodynamic property relations and refrigerant-side temperature or pressure measurements that are present in the AFDD system instrumentation. Typically, the expansion process is assumed to be isenthalpic and refrigerant enthalpy entering the evaporator is equal to the enthalpy of refrigerant in the liquid line.

The coefficient of performance (*COP)* of a system can also be estimated using virtual sensor outputs. Using the virtual cooling capacity sensor given by Equation (3.22), the virtual coefficient of performance, COP_{actual}, can be estimated using

$$COP_{actual} = \frac{\dot{Q}_{cool,actual}}{\dot{W}_{elec,virtual}} \tag{3.23}$$

where $\dot{W}_{elec,virtual}$ is the output of a virtual power sensor tuned for the system based on compressor and fan performance data.

Another important performance criteria that can be affected by faults is the sensible heat ratio, *SHR*. When the evaporator entering and leaving air temperature and humidity are measured, the actual sensible heat ratio can be estimated using

$$SHR_{actual} = \frac{c_{p,a} \cdot \left(T_{ra} - T_{sa}\right)}{h_{ra} - h_{sa}} \tag{3.24}$$

where T_{ra} and h_{ra} is the return air temperature and enthalpy, T_{sa} and h_{sa} is the supply air temperature and enthalpy, and $c_{p,a}$ is the isobaric specific heat of air. The enthalpy of humid air can be estimated using moist air property relations [61]. In applications that do not have relative humidity measurements, it is also possible to estimate the actual *SHR* using

$$SHR_{actual} = \frac{\rho_a \cdot c_{p,a} \cdot \dot{V}_{a,virtual} \cdot \left(T_{ra} - T_{sa}\right) + \dot{W}_{idf,virtual}}{\dot{Q}_{cool,actual}} \tag{3.25}$$

where ρ_a is the density of air, $\dot{V}_{a,virtual}$ is evaporator air flow rate determined using a virtual sensor, and $\dot{W}_{idf,virtual}$ is an estimate of the indoor fan power. In either Equation (3.24) or (3.25), the result should be equivalent provided the sensor readings are not significantly biased. The importance of the indoor fan power in the numerator of Equation (3.25) must be stressed when the supply air sensors are located after the indoor fan since refrigerant-side capacity measurements are unable to account for the net capacity delivered by the system. It could also be argued that Equation (3.24) is preferred since measurements of flow are not needed. With that said, relative humidity sensors have been shown to lose calibration over time and fail to meet stated accuracy [62].

3.5 Estimating Overall Performance Impacts of Faults

In order estimate the impact of a fault on system performance it is convenient to compare the actual performance of a system with the performance of a normal system at the same operating conditions. One method that can be used for this comparison is to simply form the ratio between actual equipment performance to normal equipment performance at the same operating condition. These ratios, which may be called *fault impact ratios* (*FIR*), take the form

$$FIR = \frac{\text{Actual Performance}}{\text{Normal Performance}} \tag{3.26}$$

where the numerator is some direct or indirect measurement of equipment operation and the denominator is the output of a model used to estimate normal system behavior. In other words, a *FIR* is an indicator of how much performance has changed relative to the performance that would be expected if the system were in perfect working condition. Several indicators of performance can be directly represented by these fault impact ratios, including cooling capacity, cycle efficiency, sensible capacity.

3.5.1 Estimating Total Cooling Capacity Impact of Faults

The impact different faults have on total cooling capacity of a system is important since it ultimately affects the rate of cooling that is delivered to a conditioned space. The magnitude or significance of the cooling capacity impact of faults may be characterized using a fault impact ratio, FIR_{cool} ,

$$FIR_{cool} = \frac{\dot{Q}_{cool,actual}}{\dot{Q}_{cool,normal}} \tag{3.27}$$

where $\dot{Q}_{cool,actual}$ is the actual cooling capacity delivered by the system and $\dot{Q}_{cool,normal}$ is the normal cooling capacity of the system without faults at the same operating condition. When

faults are present that degrade cooling capacity, such as low refrigerant charge or a fouled evaporator coil, FIR_{cool} becomes less than 1. In some cases, a fault may have a positive effect on system cooling capacity, e.g. an overcharged system with a fixed orifice. When this occurs, FIR_{cool} becomes greater than 1. In the trivial case when the system is normal or has faults that are inconsequential with respect to cooling capacity, $FIR_{cool} = 1$.

To illustrate how different types of faults affect the cooling performance of DX equipment, a component-based model of DX units that predicts fault impacts was used to simulate fault impacts under different ambient conditions [48, 49]. Two fixed-speed RTUs were simulated, each having identical compressors, evaporator coils, and condenser coils. The only difference between systems was the type of expansion valve used: one system was modeled with a fixed orifice (FXO), the other was modeled with a thermostatic expansion valve (TXV). Three faults were considered to illustrate the impacts on performance: improper refrigerant charge, condenser fouling, and evaporator fouling.

The impacts on total cooling capacity caused by improper refrigerant charge level faults are shown in Figure 3.10. From these plots, operating DX equipment at improper refrigerant charge levels can lead to significant reductions in total cooling capacity, even with relatively moderate losses in charge. The simulated results also show that some performance benefits are gained at high charge levels. While systems operating at higher refrigerant charge levels may operate with greater total cooling capacity – degradations in expected equipment life may occur due to liquid slugging at the suction-side of the compressor due to lower evaporator outlet quality. Additionally, increases in system power consumption can lead to decreases in COP for overcharged cases.

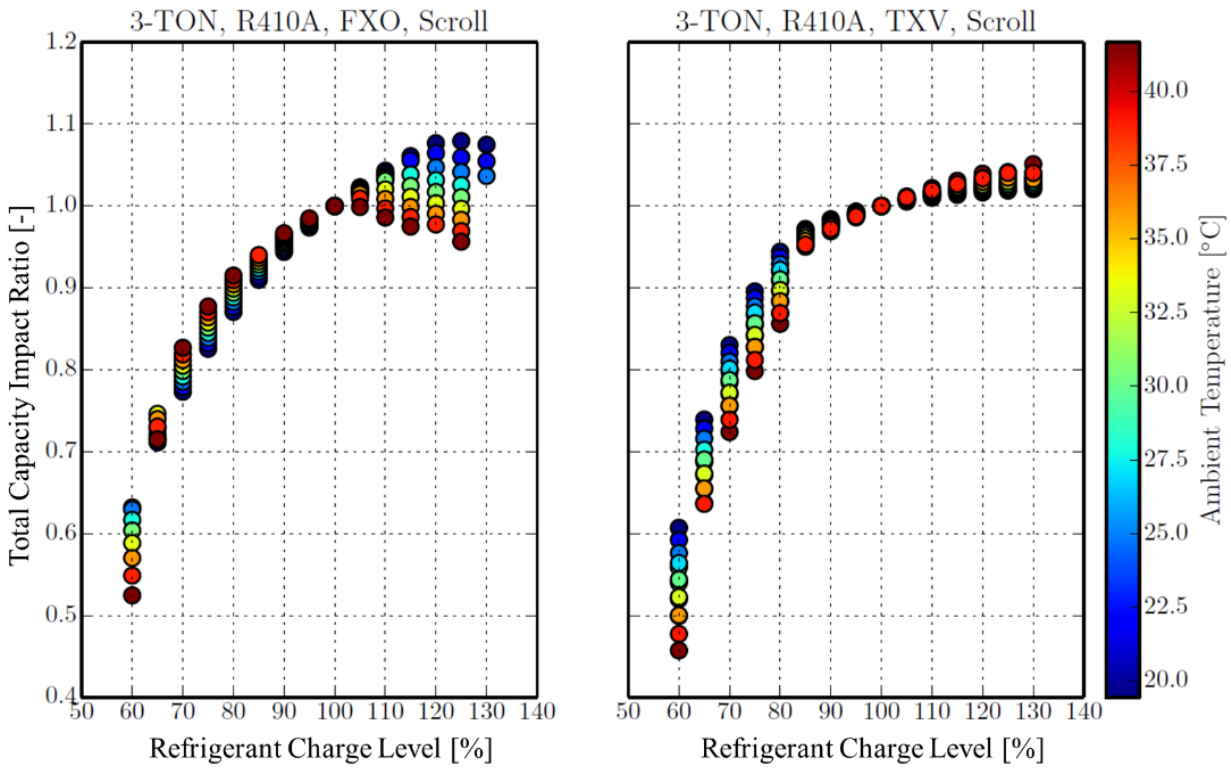

Figure 3.10. Comparison of simulated improper refrigerant charge level fault impacts for total cooling capacity on systems with different expansion devices.

A comparison of the impacts of total cooling capacity on the two systems caused by condenser fouling of different magnitudes is shown in Figure 3.11. The fault impact simulation shows that the system using a TXV is impacted significantly less than the system with the FXO. In both cases, relatively large reductions are required before degradations in cooling capacity exceed 10%. This is much different than the results that were shown for improper refrigerant charge level faults.

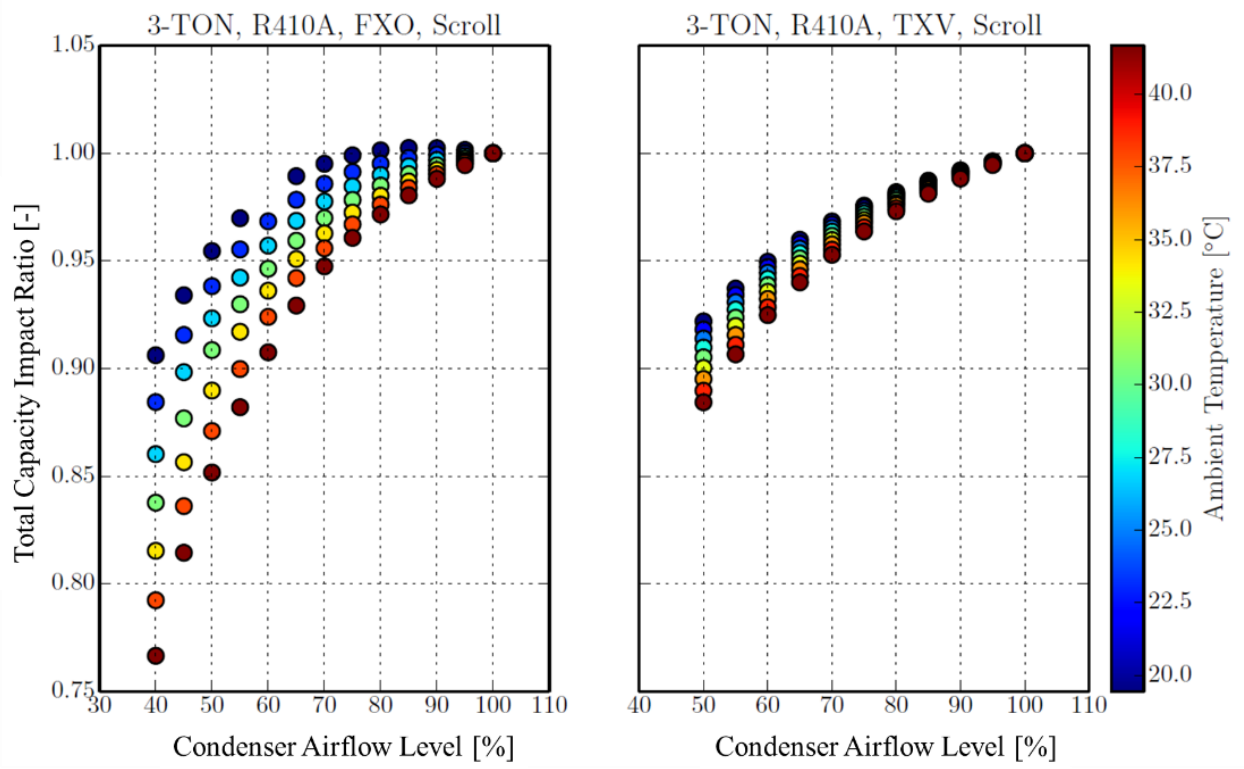

Figure 3.11. Comparison of simulated condenser fouling fault impacts on total cooling capacity for systems with different expansion devices.

The impacts of evaporator fouling on total cooling capacity at different ambient temperatures are shown in Figure 3.12. These comparisons show that evaporator fouling faults affect systems with FXOs more than systems with TXVs. However, in comparison to condenser fouling faults, reductions in evaporator air flow lead to capacity impacts that are much less dependent on ambient temperature.

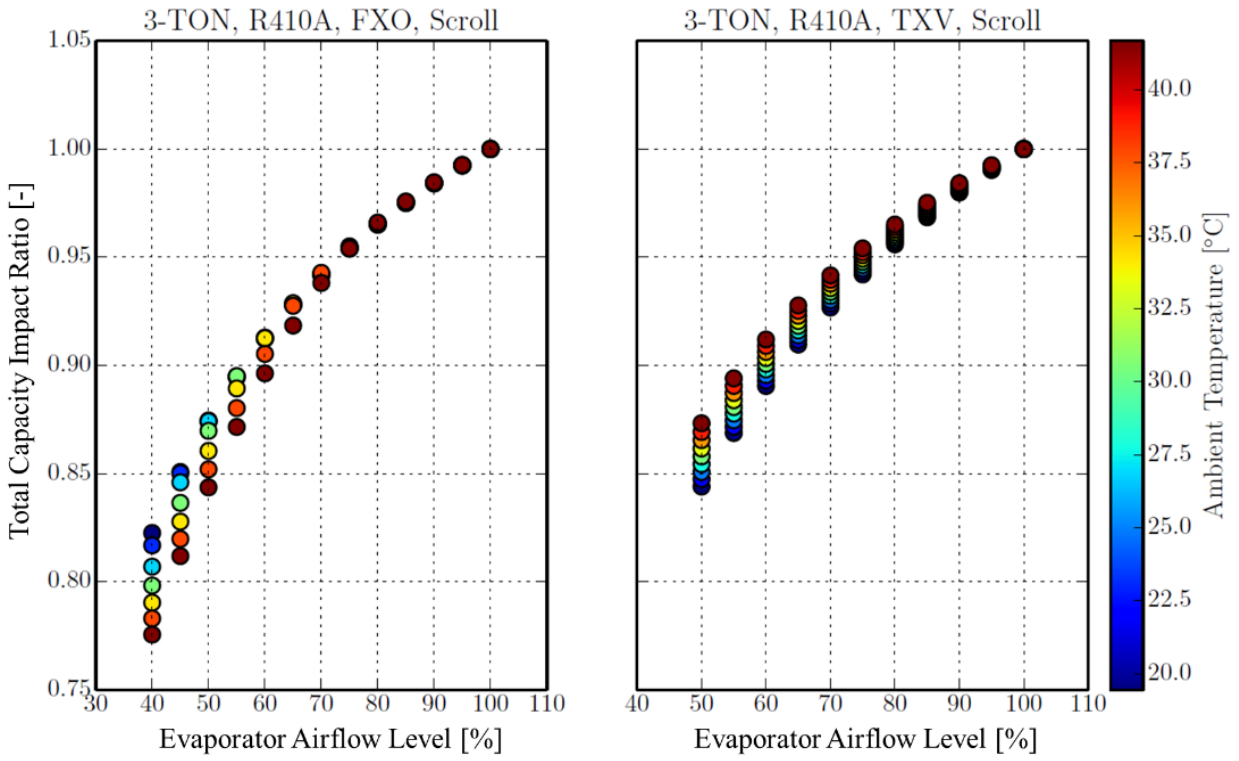

Figure 3.12. Comparison of simulated evaporator fouling fault impacts on total cooling capacity for systems with different expansion devices.

3.5.2 Estimating Cycle Efficiency Impacts of Faults

While the cooling capacity of a system determines the extent to which it can cool a building, the efficiency of the system determines the energy cost to provide this cooling. Faults often impact the cooling efficiency of a DX system since they may lead to relative decreases in cooling capacity and relative increases in electrical input power. The relative magnitude of this impact on the cooling cycle efficiency may be characterized using the efficiency impact ratio, FIR_{COP},

$$FIR_{COP} = \frac{COP_{actual}}{COP_{normal}} \tag{3.28}$$

where COP_{actual} is the actual coefficient of performance of the system and COP_{normal} is the normal coefficient of performance without faults. The COP of a DX system is defined as the ratio of cooling capacity delivered the total electrical input requirement to operate the cycle,

$$COP = \frac{\dot{Q}_{cool}}{\dot{W}_{total}} \tag{3.29}$$

Thus, to evaluate the *COP* of the system, estimates of the cooling capacity and power consumption of the system are required.

Using Equation (3.28), the *COP* impacts of refrigerant charge faults were compared between systems with FXOs and TXVs. The results shown in Figure 3.13 illustrate the importance of maintaining proper refrigerant charge levels – small changes in charge can lead to large decreases in *COP*. From the results, systems with FXOs that are overcharged may operate at higher *COP* than normal. This is due in part to the increase in evaporator temperature as well as the refrigerant entering the compressor at lower quality and specific volume which results in less compression power. However, liquid slugging may cause premature failure for some types of positive displacement compressors – a costly problem and should generally be avoided.

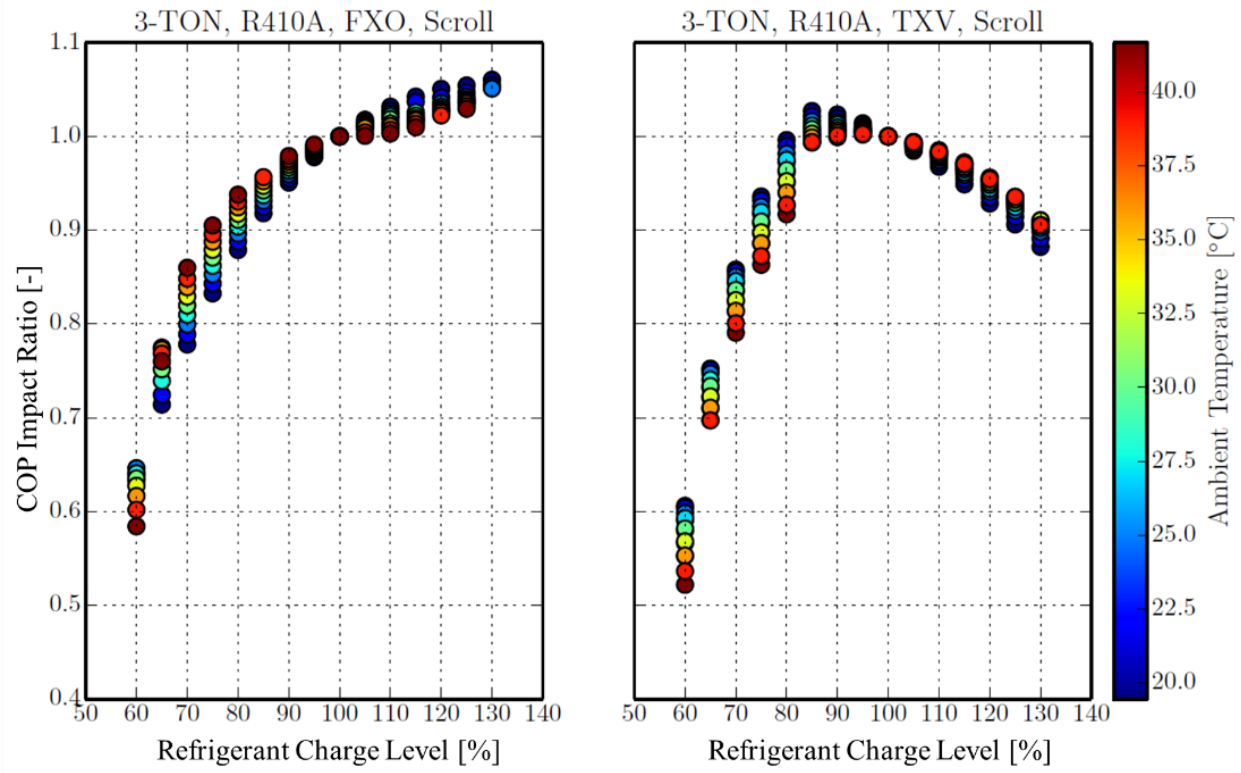

Figure 3.13. Comparison of simulated improper refrigerant charge level fault impacts on *COP* for systems with different expansion devices.

The impacts on system efficiency caused by condenser fouling have been simulated to compare systems with FXOs and TXVs. From the results shown in Figure 3.14, it is observed that systems with FXOs tend to be impacted more severely by condenser fouling than similar systems with TXVs. It should also be noted that compared with the impacts on total capacity, condenser fouling impacts system *COP* much more significantly. This is primarily due to the increases in high-side operating pressure which causes significant increases in compressor power consumption. It can also be observed that ambient temperature only moderately impacts the relative impact on system *COP* caused by condenser fouling.

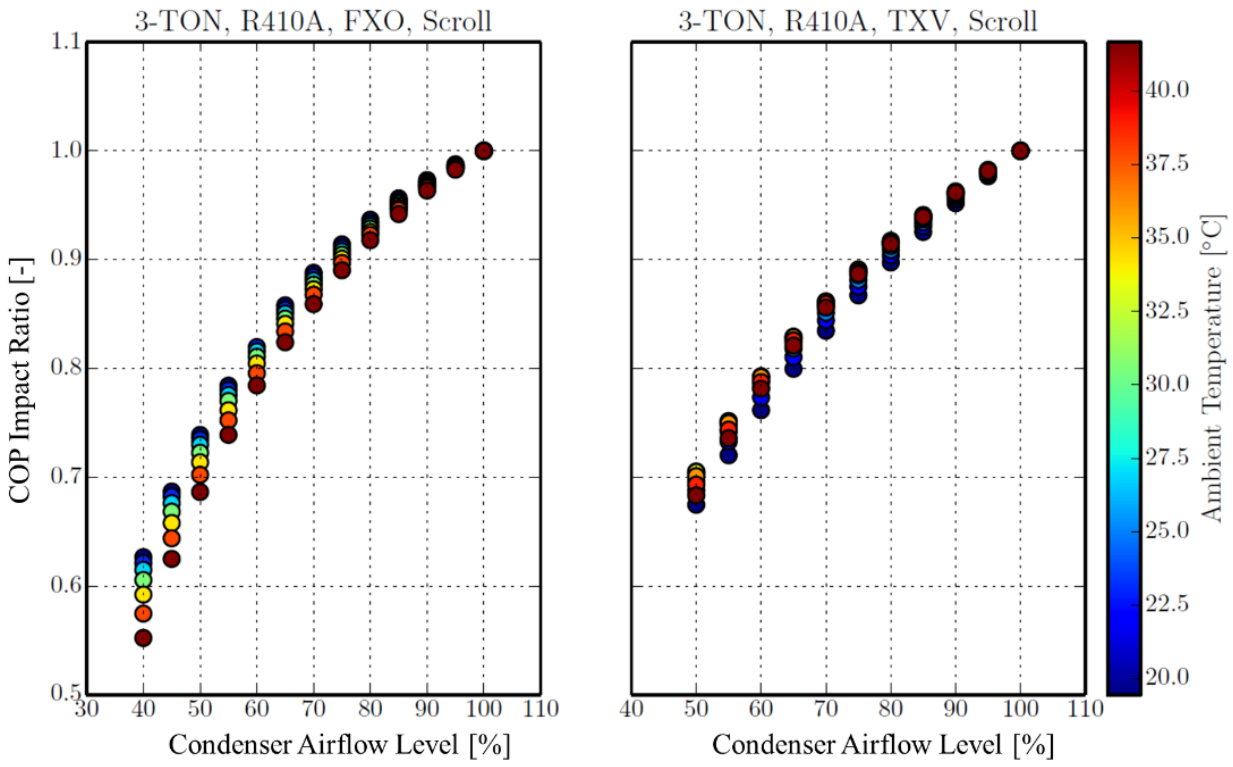

Figure 3.14. Comparison of simulated condenser fouling fault impacts on *COP* for systems with different expansion devices.

When evaporator fouling is present in DX equipment, the air flow rate across the cooling coil is reduced. These faults impact the efficiency of the cooling cycle since reductions in cooling capacity and slight increases in compressor power result. A comparison of simulated evaporator fouling impacts is shown in Figure 3.15. The relative impacts on *COP* for systems with FXOs and TXVs are approximately the same, however systems with TXVs are more sensitive to the ambient temperature. The reason for this is due to reduced refrigerant mass flow rates controlled by the TXV to maintain suction superheat.

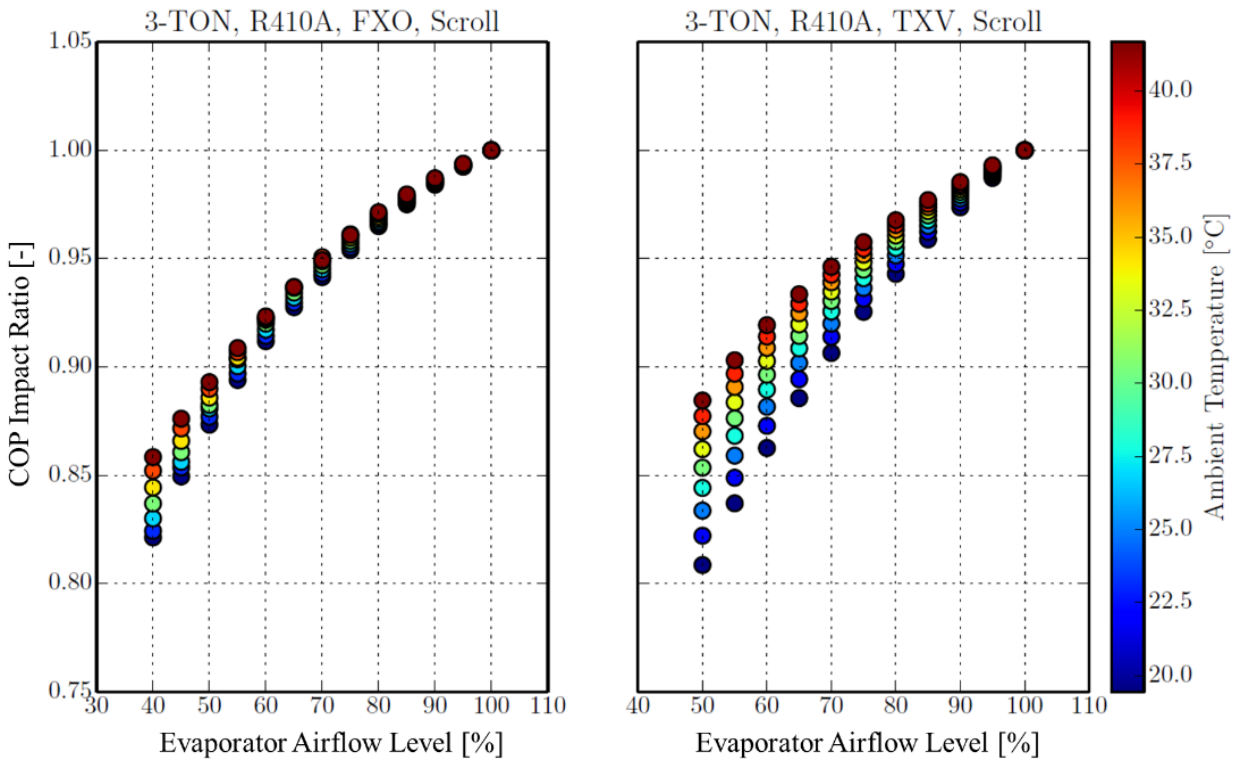

Figure 3.15. Comparison of simulated evaporator fouling fault impacts on *COP* for systems with different expansion devices.

3.5.3 Estimating Sensible Heat Ratio Impacts of Faults

Some common faults not only impact the total cooling capacity of a system, but also impact the ratio of sensible to latent cooling capacity. Examples of these faults include undercharged faults and evaporator fouling which tend to decrease the evaporator saturation temperature. When this occurs, more moisture is removed by the evaporator coil – increasing latent capacity.

In most fixed-speed or multistage RTU applications, the system is controlled solely based on the sensible condition of the zone. Thus, increases in latent capacity due to faults at the expense of sensible capacity results in increased run-time. To quantify this impact, the sensible capacity impact ratio can be used, FIR_{SHR},

$$FIR_{SHR} = \frac{SHR_{actual}}{FIR_{normal}} \tag{3.30}$$

where SHR_{actual} is the actual sensible heat ratio of the system and SHR_{normal} is the normal sensible heat ratio of the system operating without faults.

To estimate actual sensible heat ratio, the sensible capacity, $\dot{Q}_{cool,s}$, and total capacity, $\dot{Q}_{cool,t}$, of the air conditioner must be measured

$$SHR = \frac{\dot{Q}_{cool,s}}{\dot{Q}_{cool,t}}. \tag{3.31}$$

When a system operates with faults, it is impossible to measure the normal *SHR* directly and instead a reference model is required. The methodologies developed previously in Section 3.3 can be used to estimate the normal *SHR*. The combination of using virtual sensors with a commonly used model to reference normal operation enables the estimation of a fault's impacts on *SHR*.

To illustrate the range of impact on *SHR* caused by different types of faults, the component-based physical model was used to simulate improper refrigerant charge levels, condenser fouling, and evaporator fouling for RTUs with FXOs and TXVs. The impacts on SHR caused by low and high charge levels for the two types of systems are shown in Figure 3.16 at different ambient temperatures. When systems are undercharged with refrigerant, the low-side operating pressure tends to decrease significantly but the exit superheat can increase. When this occurs, the cooling coil temperature in the section of the coil with two-phase refrigerant decreases due to lower evaporation temperatures. Lower coil temperatures result in more moisture removal. However, sensible heat transfer is also reduced because less of the coil operates with two-phase refrigerant for lower refrigerant charge levels. These effects can be seen in the results generated in Figure 3.16. Interesting behavior occurs for moderately undercharged TXV systems: *SHR* tends to increase. This may be due to the increases in evaporator inlet quality that occurs with reduced

charge for TXV systems having constant superheat. In this range, the TXV is controlling the refrigerant flow to maintain evaporator superheat and it is likely that there is a reduction in the two-phase heat transfer coefficient that leads to a higher surface temperature and less moisture removal. At sufficiently low refrigerant charge, the TXV becomes fully open and the evaporating pressure and temperature drop leading to greater moisture removal in the two-phase section.

When systems are overcharged, the evaporator pressure and temperature tend to increase. For the FXO system, this leads to increases in SHR. For the TXV, the refrigerant flow rate is likely increasing with refrigerant charge in this range leading to a lower coil surface temperature and lower SHR.

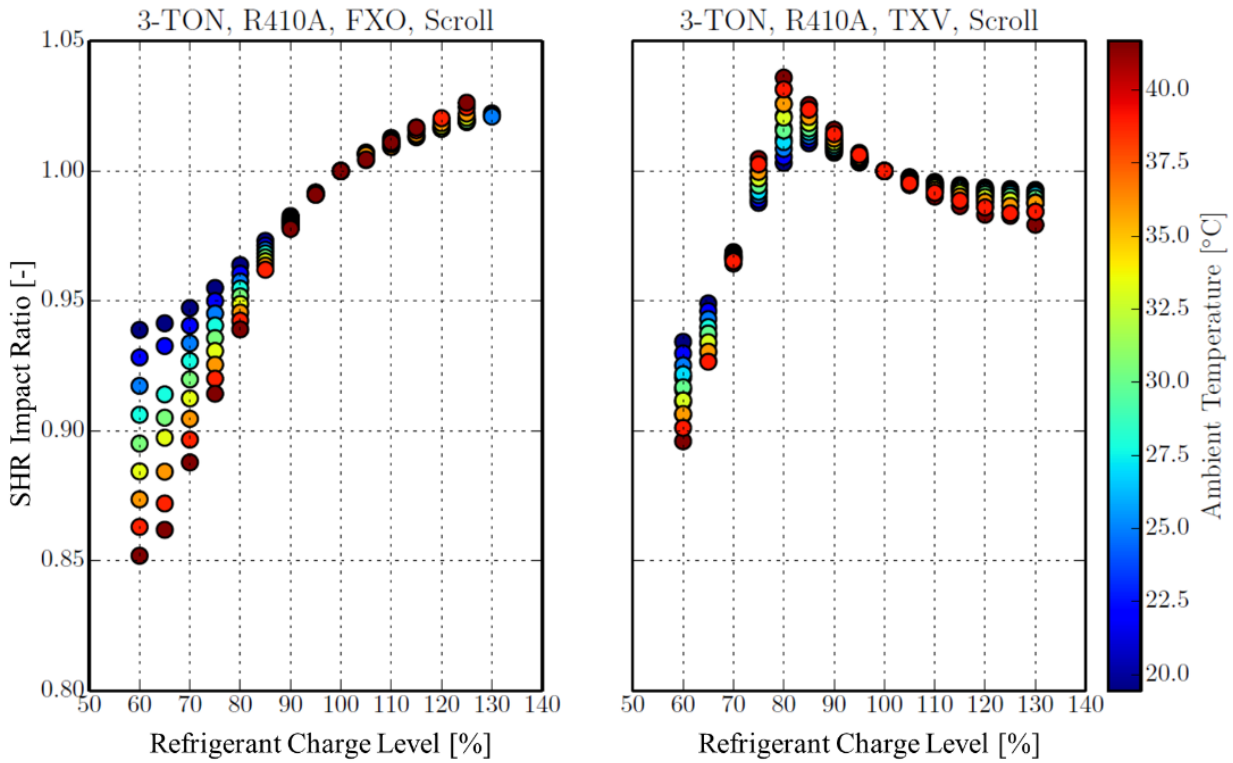

Figure 3.16. Comparison of simulated improper refrigerant charge level fault impacts on *SHR* for systems with different expansion devices.

When condenser coils become fouled, significant increases in high-side pressure result in moderate increases in evaporator pressures. In addition, increases in high-side pressure tend to increase the quality of refrigerant entering the evaporator coil. These two effects combine to increase the coil temperature in the two-phase section and increase the suction line superheat in FXO systems. Simulating these faults, condenser fouling tends to increase the *SHR* in DX systems, shown in Figure 3.17. Smaller impacts on *SHR* are observed in TXV systems – this is due the expansion valve maintaining suction superheat.

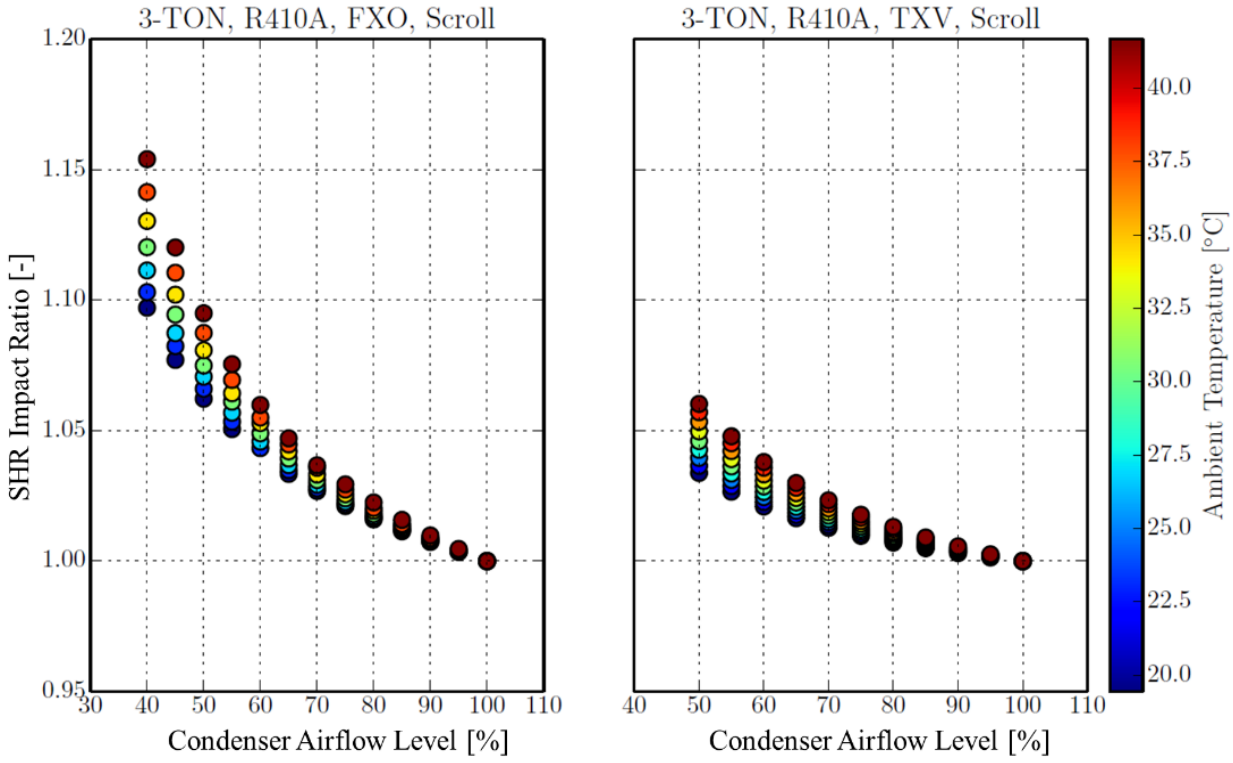

Figure 3.17. Comparison of simulated condenser fouling fault impacts on *SHR* for systems with different expansion devices.

The impacts of evaporator fouling on the *SHR* performance of the RTUs studied are shown in Figure 3.18. With increasing evaporator coil fouling, air flow through the evaporator coil is

decreased. This causes the evaporator pressure and temperature to decrease. Lower coil temperatures tend to increase moisture removal.

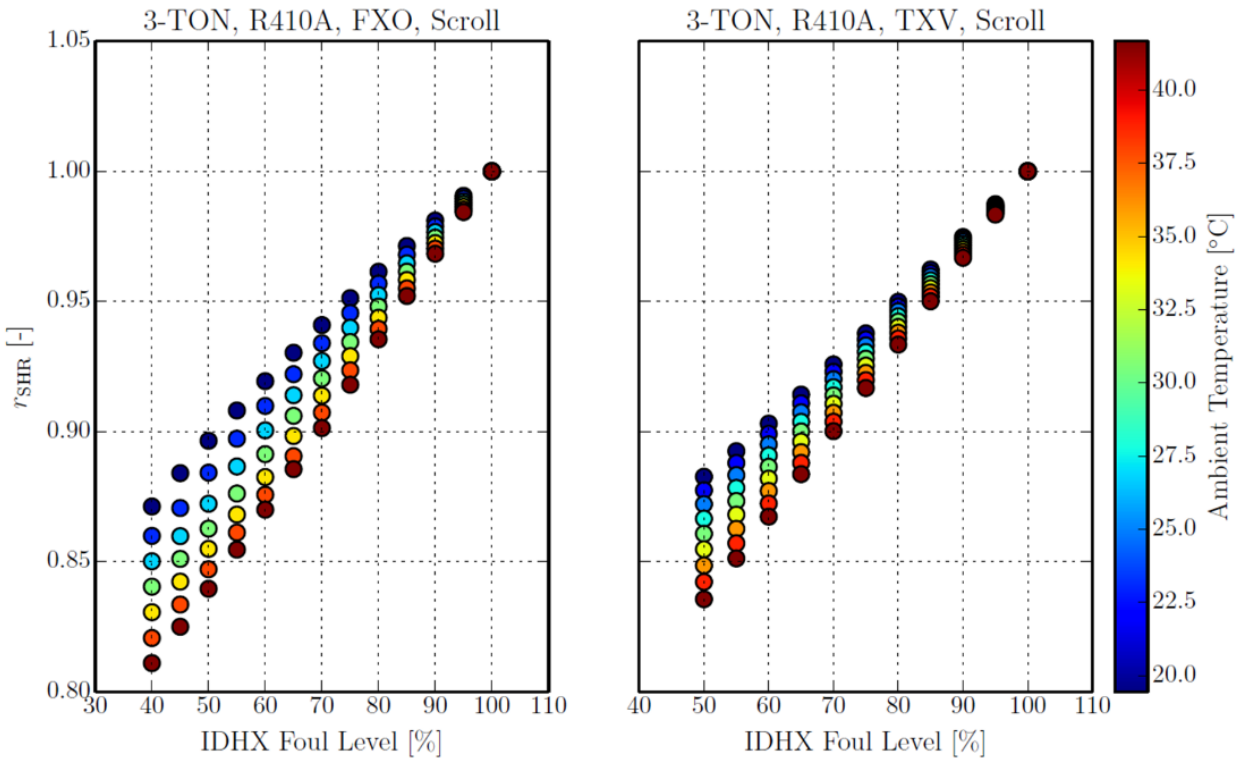

Figure 3.18. Comparison of simulated evaporator fouling fault impacts on *SHR* for systems with different expansion devices.

3.5.4 Estimating Run-time Impacts of Faults

With reduced cooling capacity, an air-conditioning system is required to operate longer to satisfy a given load. This increase in run-time has at least two effects: 1.) increased energy consumption and 2.) decreased equipment life (under the assumption that a typical air conditioner has a finite life). Many fixed-speed and multistage air conditioners are typically operated by a thermostat controlling the dry-bulb temperature of the zone to some set point. Thus, for a given sensible cooling load and sensible capacity, the run-time of the air conditioning equipment, Δt_{run} can be estimated using

$$\Delta t_{run} = \frac{Q_{load,s}}{SHR \cdot \dot{Q}_{cool,t}} \tag{3.32}$$

where $Q_{load,s}$ is the sensible cooling load in the conditioned space. Li and Braun have derived a run-time fault impact ratio, FIR_{run},

$$FIR_{run} = \frac{\left(Q_{load,s} / SHR \cdot \dot{Q}_{cool,t}\right)_{actual}}{\left(Q_{load,s} / SHR \cdot \dot{Q}_{cool,t}\right)_{normal}} = \frac{FIR_{load}}{FIR_{SHR} \cdot FIR_{cool}} \tag{3.33}$$

where FIR_{load} is the impact a fault has on the load in the zone [14]. Equation (3.33) shows that faults cause an increase in equipment run-time when the cooling load increases or when the sensible heat ratio and cooling capacity decrease.

The cooling load impact ratio, FIR_{load}, presents a possible problem for estimating the run-time impact of different faults since it is not feasible to measure the zone load using only equipment measurements. Li and Braun suggest a solution to this problem by considering only cases when $FIR_{load} = 1$ [14]. In other words, faults that affect equipment performance, such as cooling capacity, efficiency, and sensible heat ratio, should not significantly affect the load on the space. When this is the case, Equation (3.33) can be simplified to

$$FIR_{run} = \frac{1}{FIR_{SHR} \cdot FIR_{cool}} \tag{3.34}$$

While there are some faults that will affect the load faced by the equipment, *e.g.* stuck outdoor-air dampers, these types of faults are not considered in this work but can be handled with some modifications [31].

Equation (3.34) presents a method to estimate the additional run-time required for a system due to faults. However, it should be noted this model has been derived considering fixed-speed and multispeed DX systems which cycle on and off in response to loads. Variable-speed

equipment, on the other hand, do not cycle at most loads but modulate system capacity to closely match the cooling load. Application of Equation (3.34) to variable-speed systems may not yield informative results since rather than cycling between on and off, the system may increase compressor speed to maintain load. When this occurs, the run-time of the system will not change.

Run-time impacts caused by improper refrigerant charge levels have been simulated for RTUs with FXOs and TXVs and results are shown in Figure 3.19. From the simulation, it is observed that significant increases in required run-times may be caused by RTUs with low charge levels. This is due to the significant reductions in sensible cooling capacity. For overcharged systems, much less impact on run-time is observed.

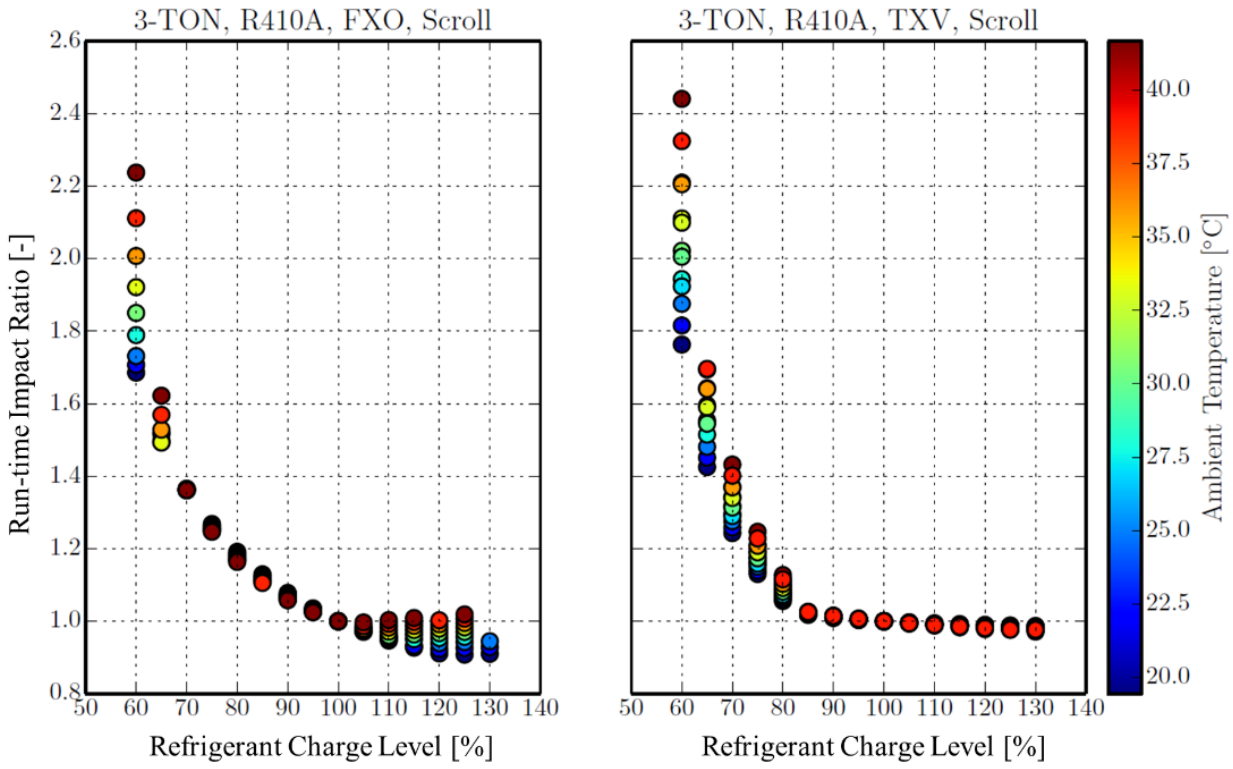

Figure 3.19. Comparison of simulated improper refrigerant charge level fault impacts on run-time requirement to meet equivalent loads for systems with different expansion devices.

In comparison to refrigerant charge run-time impacts, condenser fouling affects system run-time to a much lesser extent. Simulated results of condenser fouling, and the calculated run-time impacts are shown in Figure 3.20. Decreases in system run-time can be observed for systems with FXOs due to the small reductions in total capacity and moderate increases in SHR.

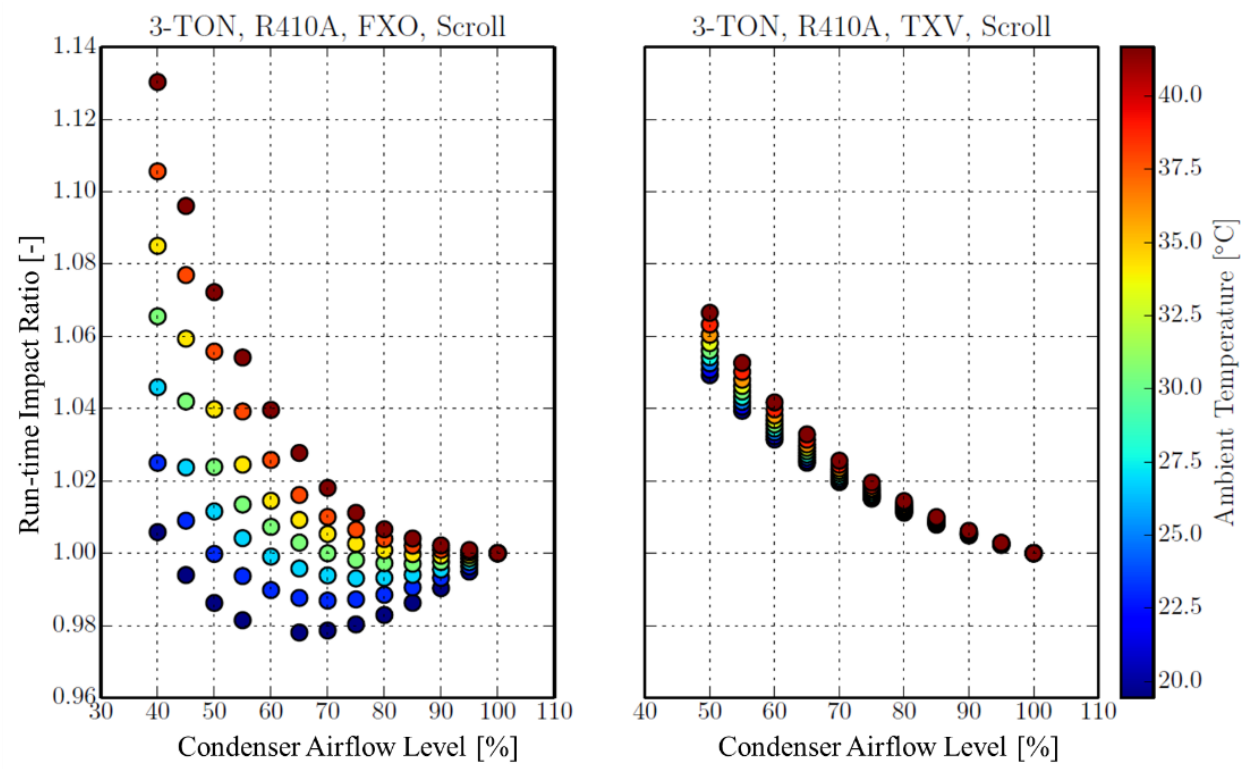

Figure 3.20. Comparison of simulated condenser fouling fault impacts on run-time requirement to meet equivalent loads for systems with different expansion devices.

The net run-time impact of simulated evaporator fouling faults on DX systems are shown in Figure 3.21. Run-time tends to increase moderately in systems that have evaporator fouling due to decreases in sensible cooling capacity. This results in systems having to run longer to maintain equivalent cooling loads. It is possible that after enough evaporator fouling, systems may no longer have sufficient sensible cooling capacity to maintain comfortable conditions.

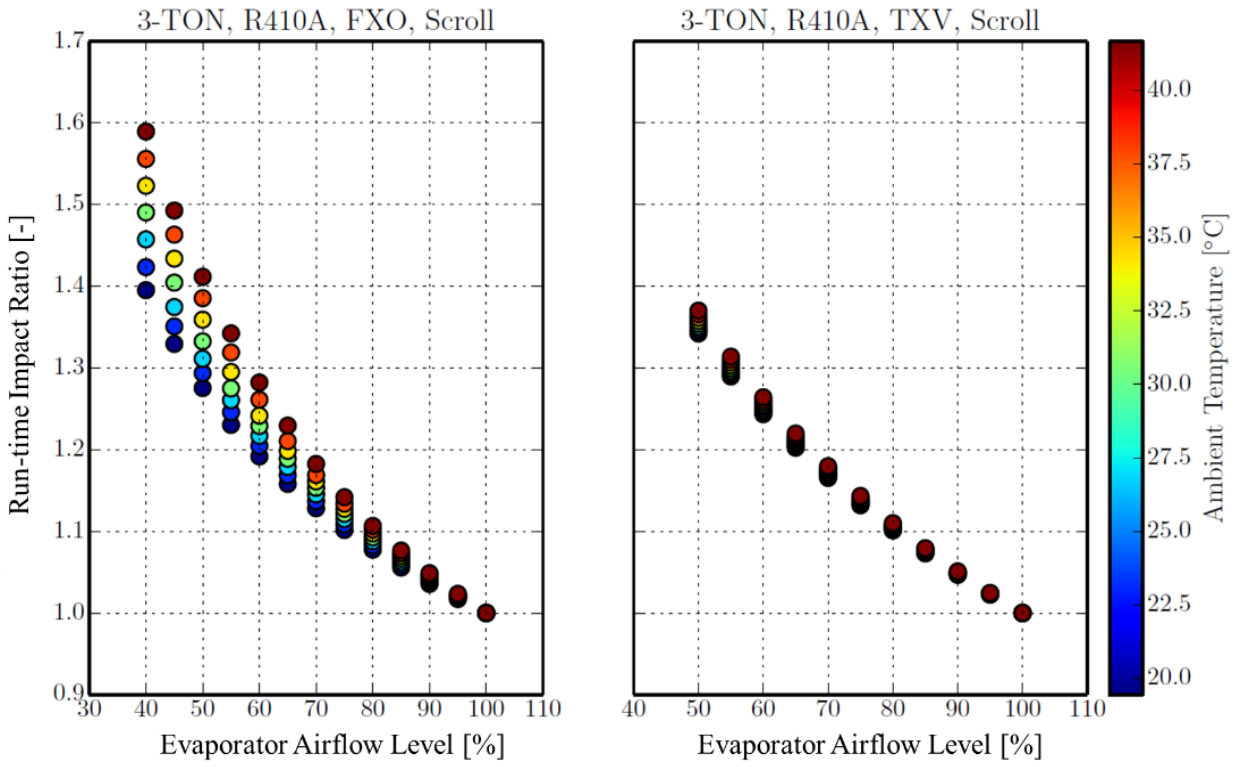

Figure 3.21. Comparison of simulated evaporator fouling fault impacts on run-time requirement to meet equivalent loads for systems with different expansion devices.

3.5.5 Estimating Energy Consumption Impacts of Faults

Ultimately, the most important consideration when it comes to assessing the total impact of a fault on the performance of an air conditioning system may be the impact on energy consumption. This is because increases in energy consumption translate directly to increases in utility costs for a building owner or facility manager which must be paid to keep operations running. And while costs associated with additional run-time are more than likely significant since equipment replacement costs are not trivial, increases in utility costs are easy to understand and are effective supporting material for service decisions.

The energy consumption required by an air conditioning system can be calculated as a function of the cooling capacity of the system, the COP, and the run-time required to satisfy a load,

$$W_{elec} = \frac{\dot{Q}_{cool,t}}{COP} \cdot \Delta t_{run}. \tag{3.35}$$

This shows the energy consumption fault impact, FIR_{elec}, is a function of the fault impacts on cooling capacity, COP, and run-time,

$$FIR_{elec} = \frac{\left(\dot{Q}_{cool,t} \cdot \Delta t_{run} / COP\right)_{actual}}{\left(\dot{Q}_{cool,t} \cdot \Delta t_{run} / COP\right)_{normal}} = \frac{FIR_{cool}}{FIR_{COP}} \cdot FIR_{run} \tag{3.36}$$

The definition of FIR_{elec} in Equation (3.36) shows that it is possible to estimate the total impacts of faults on energy consumption using the combination of virtual sensors and reference models for normal performance. Moreover, the additional energy consumption that was used by a system due to faults, ΔW_{elec}, can be tracked over time by integrating FIR_{elec} with estimates of the actual energy consumed,

$$\Delta W_{elec} = \sum \left(FIR_{elec} - 1\right) \cdot \dot{W}_{elec,actual} \tag{3.37}$$

where $\dot{W}_{elec,actual}$ is the actual electrical energy consumption that can be determined using direct measurements or virtual sensors.

The total energy impact of improper refrigerant charge level faults has been simulated and results are shown in Figure 3.22 for DX equipment with FXOs and TXVs. The results show that significantly more energy is consumed by equipment that are undercharged or overcharged. For systems that are significantly undercharged, energy consumption is nearly doubled. This is caused by significant reductions in sensible cooling capacity and COP.

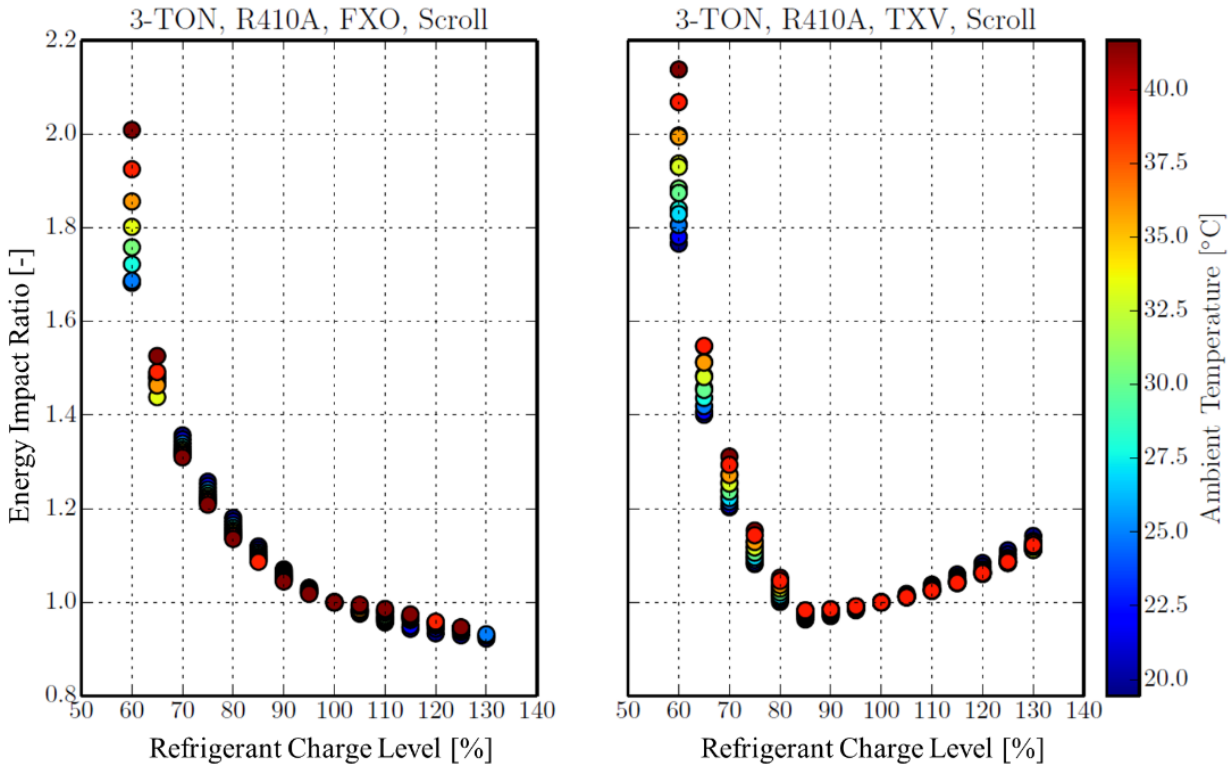

Figure 3.22. Comparison of simulated improper refrigerant charge level fault impacts on energy consumption to meet equivalent loads for systems with different expansion devices.

Similarly, comparisons of the energy impact of different condenser fouling faults are shown in Figure 3.23. These results show that condenser fouling results in large increases in energy consumption over time. Energy consumption tends to increase in systems with condenser fouling primarily due to larger compressor power consumption caused by elevated high-side pressures.

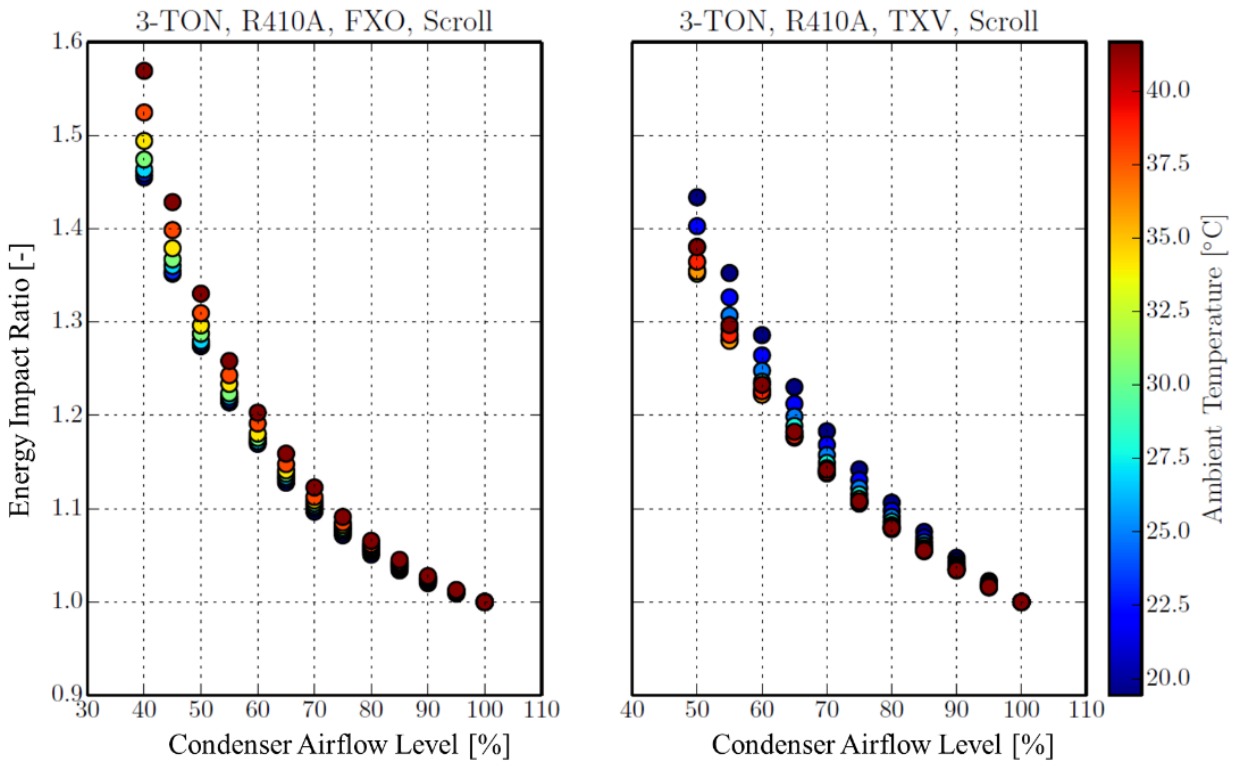

Figure 3.23. Comparison of simulated condenser fouling fault impacts on energy consumption to meet equivalent loads for systems with different expansion devices.

Similar impacts on energy consumption are also caused by evaporator fouling, shown in Figure 3.24. This is caused by reduced sensible cooling capacity which results in longer run-times to maintain equivalent cooling loads. From the simulation outputs, the type of expansion valve used in the equipment did not seem to affect energy impacts significantly.

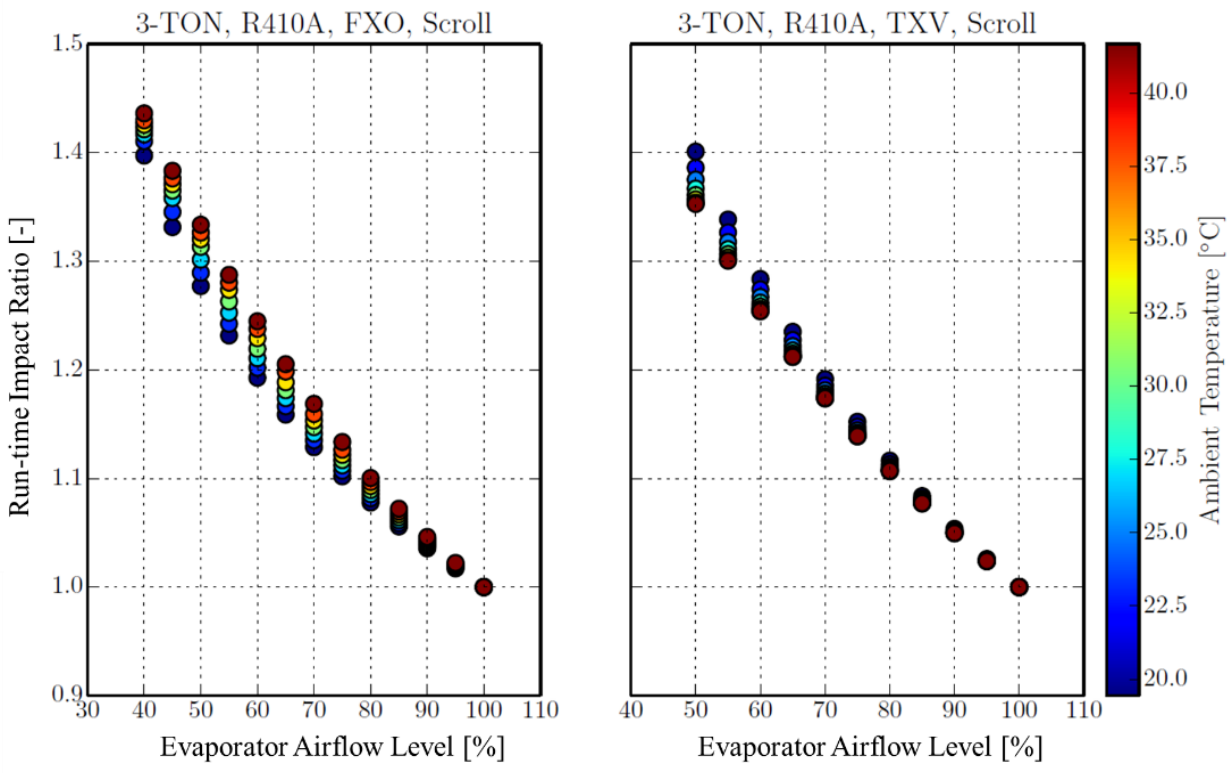

Figure 3.24. Comparison of simulated evaporator fouling fault impacts on energy consumption to meet equivalent loads for systems with different expansion devices.

3.6 Estimating Impacts of Multiple Faults using Semi-Empirical Models

To address constraints associated with the cost of performing experimental testing in psychrometric chambers that would be required to develop detailed models like was proposed by Cheung and Braun or even neural network models, simplified semi-empirical models are an alternative for predicting the impacts of faults on system performance. These models should be able to predict the total cooling capacity, efficiency, and sensible heat ratio of operating systems that are subject to faults. They should also be able to relate these impacts to operating a system without faults so that economical service decisions can be made.

A semi-empirical equipment model that can be used to predict the performance of systems at off-design conditions has been described in Equation (3.1) for fixed-speed systems and Equation (3.13) for variable-speed systems. Recently there has been work published by Du *et al*.

and Cheung and Braun with the intent to develop models for estimating fault impacts in energy simulation programs [63, 64]. The models developed in these works have aimed to estimate single fault impacts using least squares regression models. In the present work, the goal is to estimate total fault impacts when multiple faults are present. One possible extension of the model described in Equations (3.1) or (3.13) based on the work by Du *et al*. and Cheung and Braun can be expressed as,

$$\frac{\dot{Q}_{cool,\,fault}}{\dot{Q}_{cool,normal}} = f_{chrg} \cdot f_{c,\,foul} \cdot f_{e,\,foul} \tag{3.38}$$

where $\dot{Q}_{cool,\,fault}$ is the cooling capacity delivered by a system with faults, $\dot{Q}_{cool,normal}$ is the normal system cooling capacity, and the additional f_i terms are correction factors adjusting the total capacity for the effects of different faults. These correction factors can be tuned for different faults including improper refrigerant charge levels (f_{chrg}), condenser fouling ($f_{c,\,foul}$), and evaporator fouling ($f_{e,\,foul}$).

The impacts of different faults on equipment performance can be estimated using simple polynomial regression models. For example, the impact of different refrigerant charge levels on the cooling capacity can be represented by

$$f_{chrg} = a_0 + a_1 \cdot x_{chrg} + a_2 \cdot x_{chrg}^2 \tag{3.39}$$

where x_{chrg} is the relative amount of refrigerant charge in the system and $a_{0,1,2}$ are empirically determined parameters relating the charge level to the capacity impacts. The relative amount of refrigerant in the system can be expressed as

$$x_{chrg} = \frac{m_{chrg,actual} - m_{chrg,normal}}{m_{chrg,normal}} \tag{3.40}$$

where $m_{chrg,actual}$ is the actual mass of charge in the system and $m_{chrg,normal}$ is the mass of charge recommended by the equipment manufacturer. In practice, the actual amount of refrigerant in a system could be estimated using the virtual refrigerant charge sensor. A value for the normal amount of refrigerant could be programmed by a manufacturer into an equipment controller of the AFDD device.

The impact of condenser fouling can be modeled in a similar manner as given by

$$f_{c,foul} = b_0 + b_1 \cdot x_{ca,flow} + b_2 \cdot x_{ca,flow}^2 \tag{3.41}$$

where $x_{ca,flow}$ is the relative amount of airflow through the condenser and $b_{0,1,2}$ are empirical parameters determined using test data. Condenser fouling can be represented as a function of the relative reduction in condenser airflow

$$x_{ca,flow} = \frac{\dot{V}_{ca,actual} - \dot{V}_{ca,normal}}{\dot{V}_{ca,normal}} \tag{3.42}$$

where $\dot{V}_{ca,actual}$ is the actual airflow rate through the condenser coil and $\dot{V}_{ca,normal}$ is the normal airflow rate through the condenser coil. Using a virtual sensor, the condenser airflow rate of a system can be measured. For fixed-speed systems, the normal condenser airflow rate can be programmed into an equipment controller. For variable-speed systems, a model for the normal airflow could be implemented by the manufacturer that depends on the control input.

The impact of evaporator fouling can be modeled in a similar manner with

$$f_{e,foul} = c_0 + c_1 \cdot x_{ea,flow} + c_2 \cdot x_{ea,flow}^2 \tag{3.43}$$

where $x_{ea,flow}$ is the relative amount of airflow through the evaporator coil and $c_{0,1,2}$ are empirical parameters determined using test data. Evaporator fouling can be represented as a function of the relative reduction in evaporator airflow

$$x_{ea,\,flow} = \frac{\dot{V}_{ea,actual} - \dot{V}_{ea,normal}}{\dot{V}_{ea,normal}} \tag{3.44}$$

where $\dot{V}_{ea,actual}$ is the actual airflow rate through the evaporator coil and $\dot{V}_{ea,normal}$ is the normal airflow rate through the evaporator coil. Using a virtual sensor, the evaporator airflow rate of a system can be measured. For systems that have fixed-speed indoor fans, the normal evaporator airflow rate can be programmed into an equipment controller. For variable-speed systems, a model for the normal airflow could be implemented by the manufacturer that depends on the control input and the supply duct static pressure.

Like total cooling capacity, the electrical input power impacts of different faults can be estimated using a semi-empirical model

$$\frac{\dot{W}_{total,\,fault}}{\dot{W}_{total,normal}} = g_{chrg} \cdot g_{c,\,foul} \cdot g_{e,\,foul} \tag{3.45}$$

where $\dot{W}_{cool,\,fault}$ is the total power consumed by a system with faults, $\dot{W}_{total,normal}$ is the normal system power consumption, and the additional g_i terms are correction factors adjusting the normal power consumption for the effects of different faults. These correction factors can be tuned for different faults including improper refrigerant charge levels (g_{chrg}), condenser fouling ($g_{c,\,foul}$), and evaporator fouling ($g_{e,\,foul}$). Simple quadratic polynomial models can be used to capture the impact of each fault,

$$g_{chrg} = d_0 + d_1 \cdot x_{chrg} + d_2 \cdot x_{chrg}^2 \tag{3.46}$$

$$g_{c,\,foul} = e_0 + e_1 \cdot x_{ca,\,flow} + e_2 \cdot x_{ca,\,flow}^2 \tag{3.47}$$

$$g_{e,\,foul} = k_0 + k_1 \cdot x_{ea,\,flow} + k_2 \cdot x_{ea,\,flow}^2 \tag{3.48}$$

where $d_{0,1,2}$, $e_{0,1,2}$, and $k_{0,1,2}$ are empirical parameters tuned using test data.

One potential pitfall of the methodology proposed in Equations (3.38) and (3.45) is the assumption that the impacts of different faults on system performance are conditionally independent that is predicated by the model form. Using the model proposed, the interactions between different faults cannot be captured since the correction models f_i and g_i are all univariate functions of a specific fault indicator. To assess the potential errors associated with this independence assumption, the detailed fault impact model previously developed by Cheung and Braun has been used to generate training and validation datasets for model evaluation.

3.6.1 Simulated Results using Fault Impact Meta-Model

Before fitting the fault impact models, normal (fault-free) models for the total cooling capacity and power consumption were fit using data from the training set containing no faults. The model form used to fit normal performance is given by Equation (3.1) for total cooling capacity and Equation (3.9) for power consumption. The resulting normal system model fits are plotted against the detailed fault impact model outputs containing no faults in Figure 3.25. Based on the comparison of the outputs from the two models, it was determined that the simplified model accuracy was sufficient to predict normal performance.

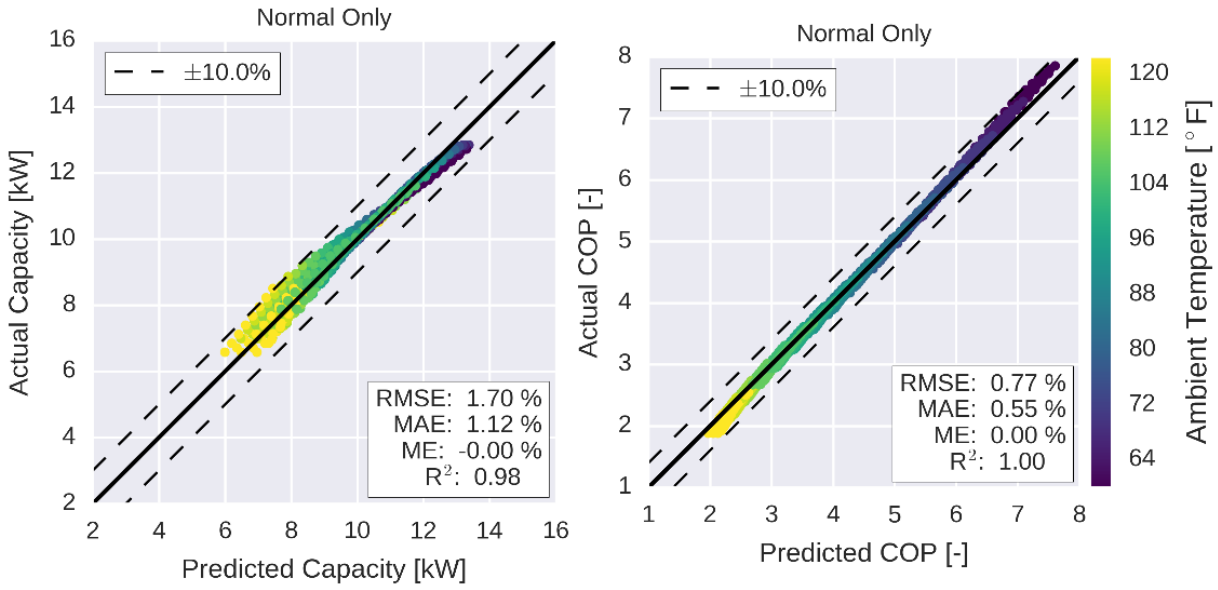

Figure 3.25. Accuracy of models used to predict normal total cooling capacity and COP against neural network meta-model generated outputs.

After fitting the normal performance models, models to predict the impacts of refrigerant charge faults were fit using data from the training dataset containing only normal data and data containing refrigerant charge faults. The model forms used to predict these impacts are given in Equations (3.38) and (3.45) for total capacity and power consumption. After analysis of the prediction errors, it was determined the impacts of refrigerant charge on system performance had a significant correlation with the ambient temperature and entering air wet-bulb temperature. To account for this dependence, the original models for the charge impacts were modified using

$$f_{chrg} = a_0 + a_1 \cdot B_{ID} \cdot x_{chrg} + a_2 \cdot T_{OD} \cdot x_{chrg}^2 \tag{3.49}$$

and

$$g_{chrg} = d_0 + d_1 \cdot B_{ID} \cdot x_{chrg} + d_2 \cdot T_{OD} \cdot x_{chrg}^2 \tag{3.50}$$

where B_{ID} evaporator entering air wet bulb temperature and T_{OD} is the outdoor-air dry bulb temperature. Validation results for the resulting model fits for the total cooling capacity and calculated COP are shown in Figure 3.26.

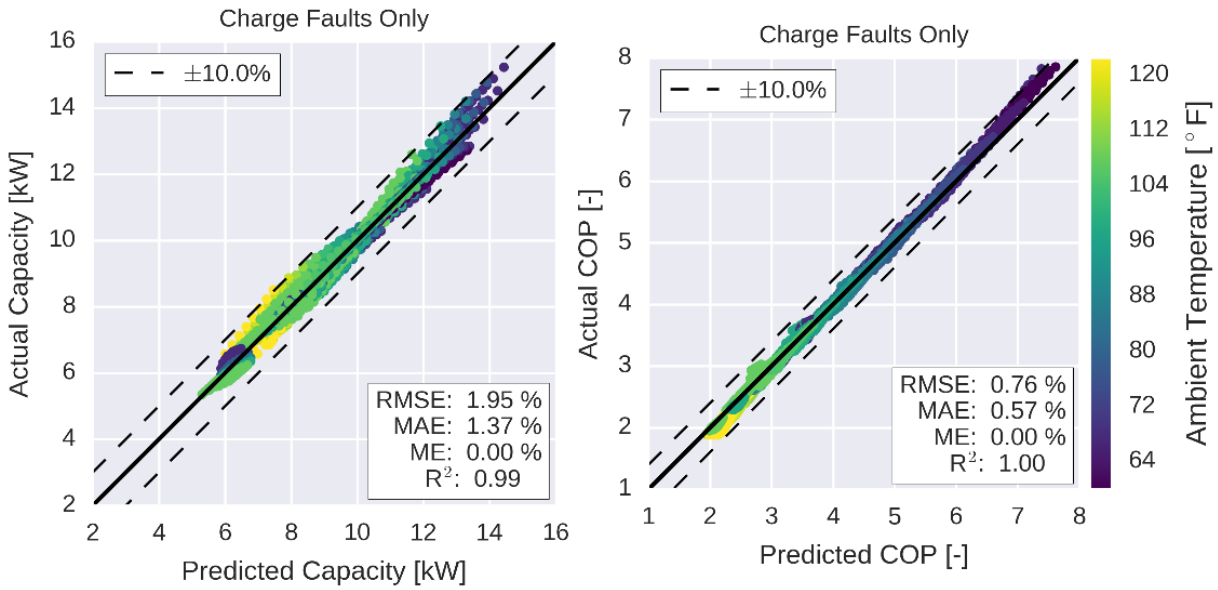

Figure 3.26. Accuracy of simplified charge fault impact models for predicting total cooling capacity and COP of systems with charge faults.

To predict the impacts of condenser fouling on total cooling capacity and COP, condenser fouling fault impact models given by Equations (3.41) and (3.47) were tuned using data generated with the detailed fault impact model. In this training data set, only data at different ambient conditions and condenser fouling levels were used. The predicted capacity and predicted COP are compared with the values calculated by the detailed model in Figure 3.27. The prediction errors of the total cooling capacity were within 10% of the detailed model outputs for all the test cases generated. The prediction errors of the COP were within 5% of the detailed model outputs.

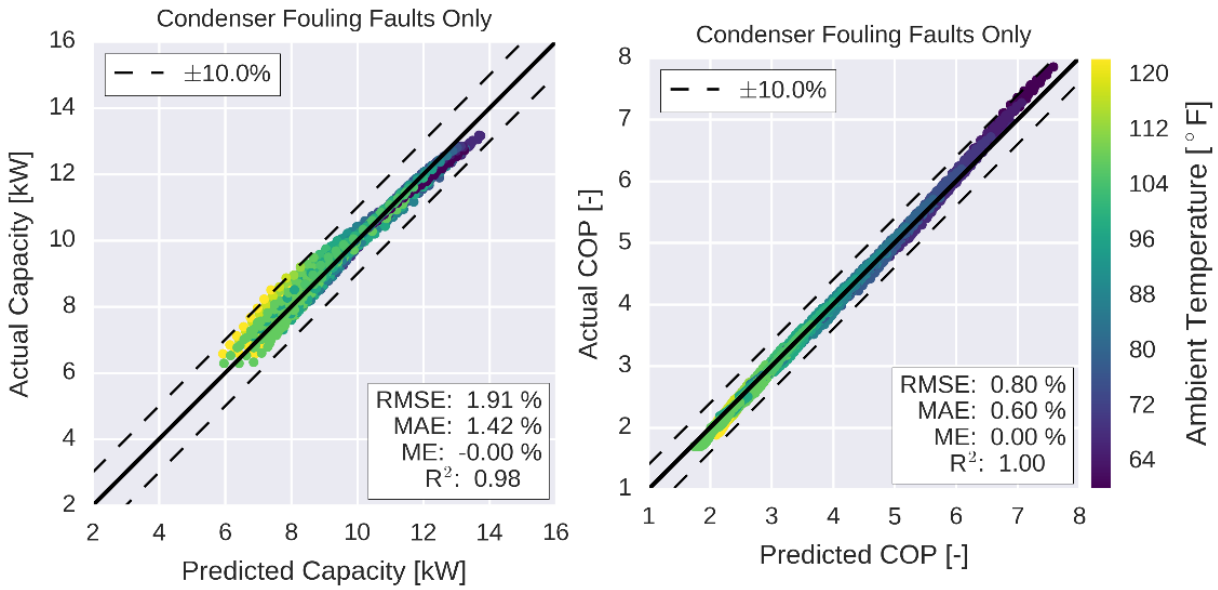

Figure 3.27. Accuracy of simplified condenser fouling fault impact models for predicting total cooling capacity and COP of systems with condenser fouling faults.

The evaporator fouling fault impact models given by Equations (3.43) and (3.48) were tuned using data generated by the detailed fault impact inverse model. To capture the impacts of only evaporator fouling, data containing different ambient conditions and evaporator airflow rates were used to tune the empirical coefficients. The prediction accuracy of the total cooling capacity and calculated COP is compared with the detailed model outputs in Figure 3.28. For the test cases used in the evaluation, the evaporator fault impact predictions were within 10% of detailed model outputs. The COP predictions using the simplified semi-empirical model had errors less than 10% as well.

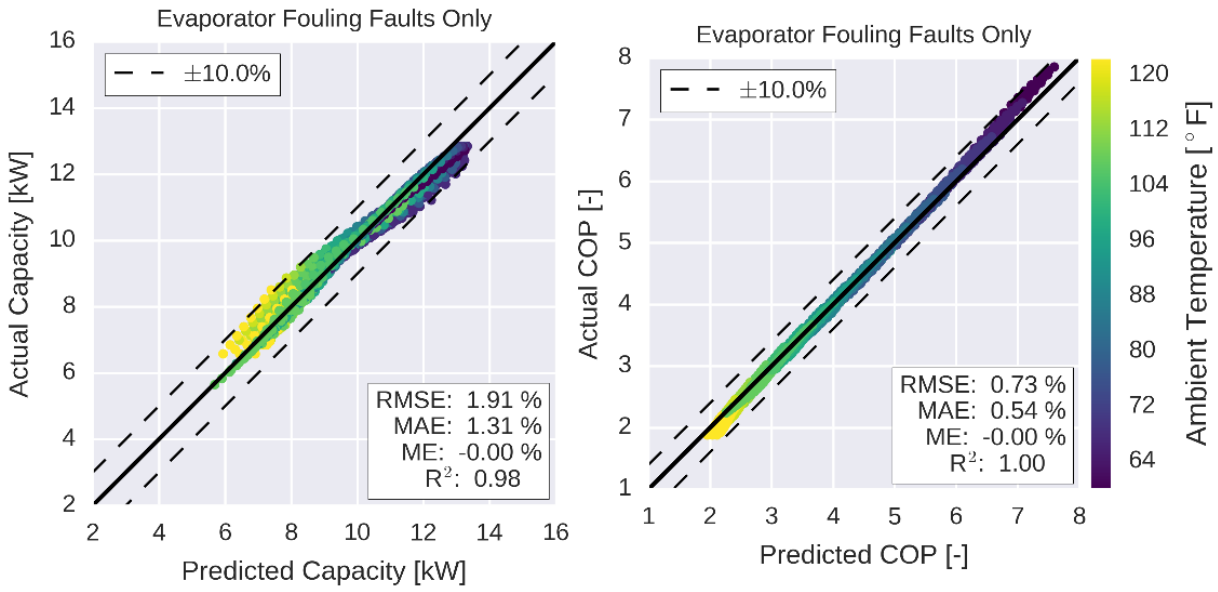

Figure 3.28. Accuracy of simplified evaporator fouling (reductions in evaporator air flow) fault impact models for predicting total cooling capacity and COP of systems with evaporator fouling faults.

Since the goal of these simplified fault impact models is to be used within a method to isolate the relative impacts of multiple faults that occur simultaneously, it is important to test the model accuracy while multiple faults are imposed on the system. Using the semi-empirical models trained for using single fault data sets, the total cooling capacity and COP were predicted for operating conditions with multiple faults present. The first combination of faults studied were condenser fouling and refrigerant charge faults. A comparison of the predicted total cooling capacity and COP with the detailed model outputs is shown in Figure 3.29. The predicted cooling capacity errors were within 10% of the detailed model outputs for 99% of the test cases. The predicted COP errors were within 10% of the detailed model outputs for greater than 99% of the test cases. It should be noted that significant bias between the semi-empirical fault impact models and the detailed model outputs were observed. In consideration of the total cooling capacity, the mean error of all the test cases was approximately -1.99%. This indicates that the semi-empirical models under predict the detailed model outputs by about 2%. This bias may be

the result of the inherent independence assumption between fault impact factors. The detailed model captures the interaction between multiple faults. Similar behavior can be observed for the COP prediction errors – the semi-empirical model tended to under predict the detailed model outputs by approximately 0.83%.

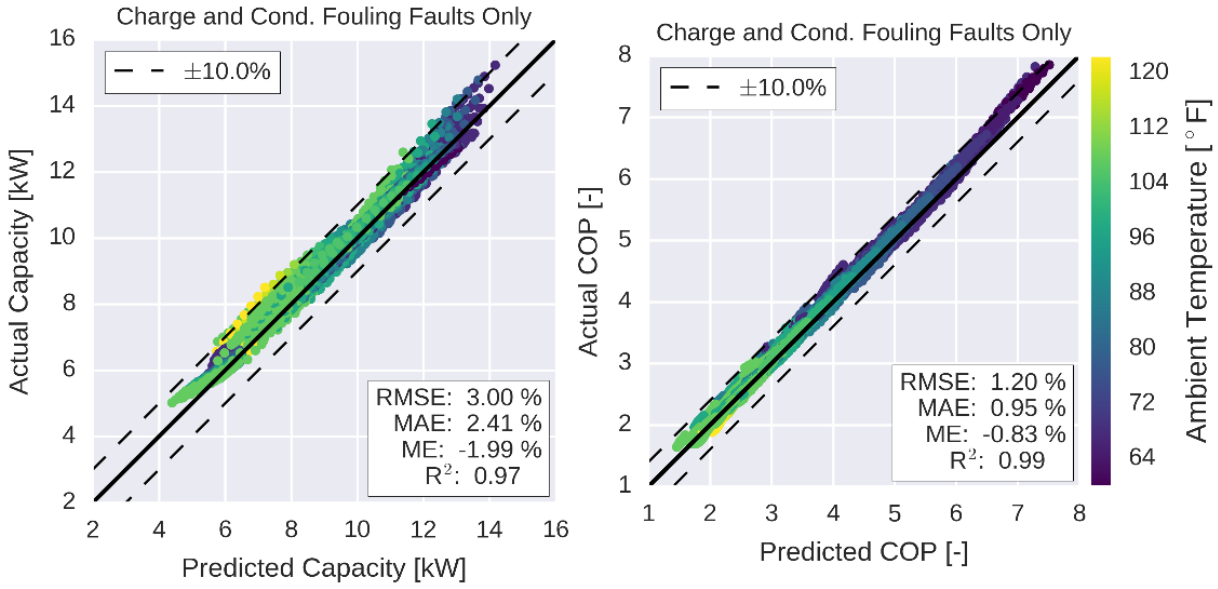

Figure 3.29. Accuracy of simplified fault impact models for predicting total cooling capacity and COP of systems with combinations of refrigerant charge and condenser fouling faults.

Combinations of condenser fouling and evaporator fouling were generated using the detailed fault impact inverse model. Using the semi-empirical model, the total cooling capacity and COP were predicted at the same conditions and were compared with the detailed model outputs in Figure 3.30. The predicted cooling capacity errors were within 10% of the detailed model outputs for 99% of the test cases. The predicted COP errors were within 10% of the detailed model outputs for greater than 99% of the test cases. It should be noted that significant bias between the semi-empirical fault impact models and the detailed model outputs were observed. In consideration of the total cooling capacity, the mean error of all the test cases was approximately -0.98%. This indicates that the semi-empirical models under predict the detailed

model outputs by about 1%. This bias may be the result of the inherent independence assumption between fault impact factors. The detailed model captures the interaction between multiple faults. Similar behavior can be observed for the COP prediction errors – the semi-empirical model tended to under predict the detailed model outputs by approximately 0.50%.

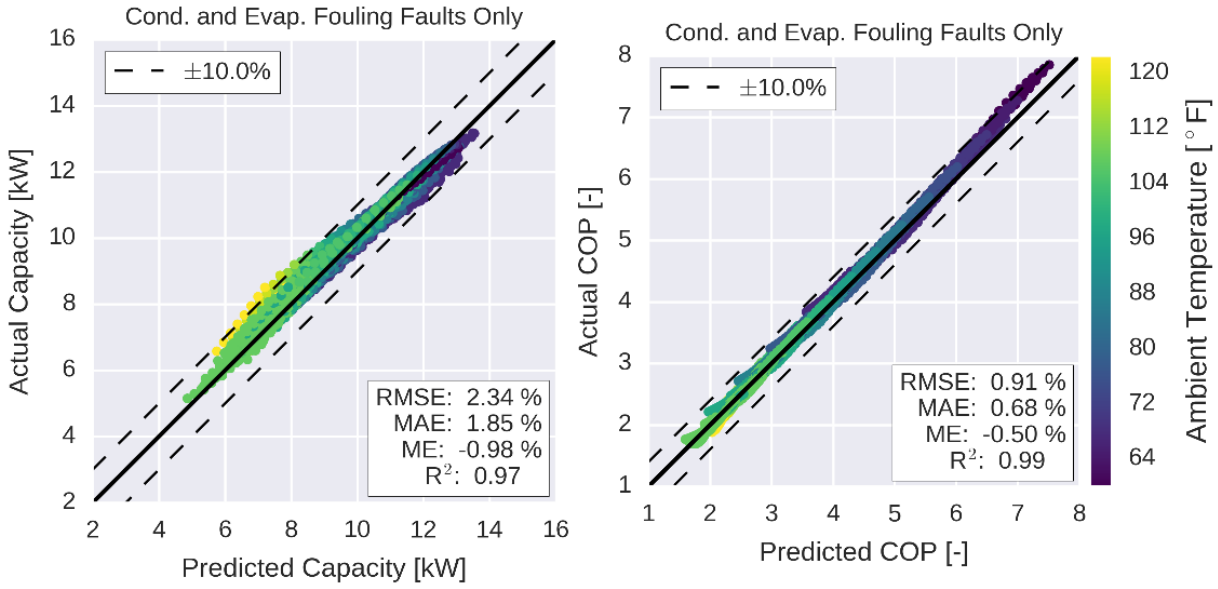

Figure 3.30. Accuracy of simplified fault impact models for predicting total cooling capacity and COP of systems with combinations of refrigerant charge and evaporator fouling faults.

Finally, combinations of the three faults were used to assess the performance of the semi-empirical fault impact models. These results are shown in Figure 3.31. With respect to the total cooling capacity, the model predictions fell within 10% for greater than 95% of the test cases. With regards for the COP, the model predictions fell within 10% for great than 90% of the test cases. Additionally, it should be noted that the model errors are biased by approximately 1.0% in both the total cooling capacity and the COP. This may indicate that the semi-empirical models do not capture impacts that are dependent on the correlation between different faults while the detailed fault impact model does. It appears from the results that the total impact of combinations of faults may become slightly more severe when several faults are present.

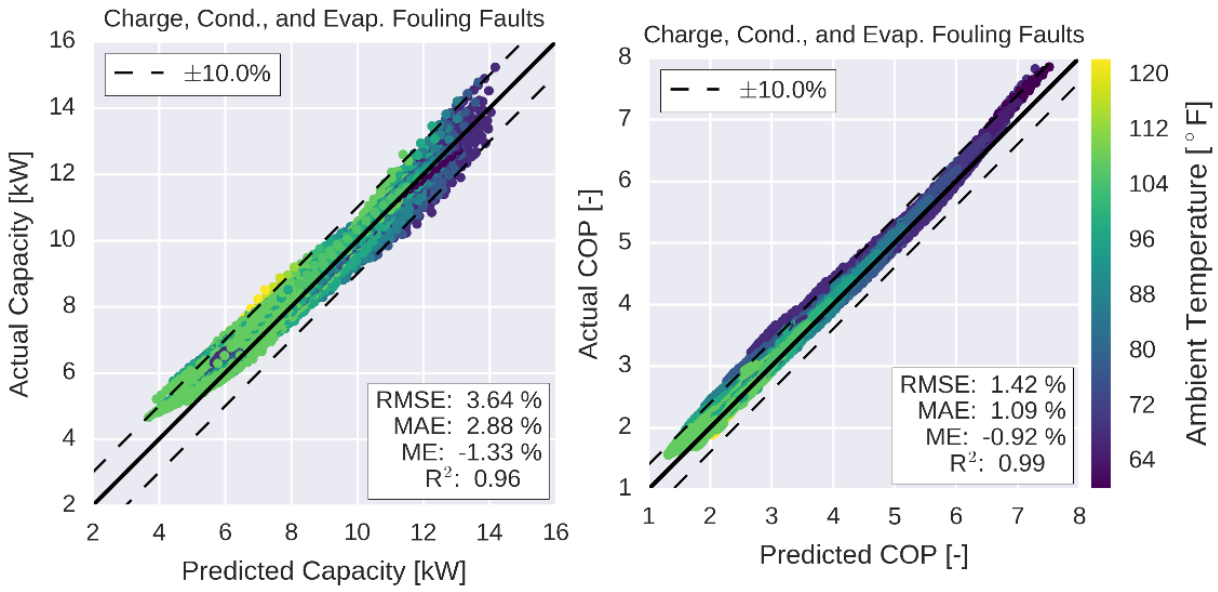

Figure 3.31. Accuracy of simplified fault impact models for predicting total cooling capacity and COP of systems with combinations of refrigerant charge, condenser fouling, and evaporator fouling faults.

3.7 Summary of Fault Impact Estimation Results

Semi-empirical models that can be implemented within microprocessors using relatively modest computing resources have been developed and implemented for the purposes of estimating the performance impacts of different faults. To accomplish this task, the normal cooling capacity, power consumption, and sensible heat ratio of fixed-speed and variable-speed DX systems were modeled. Next, a discussion of how virtual sensors can be used to estimate the actual cooling capacity, power consumption and sensible heat ratio was provided. Following this, a methodology for determining the overall impacts of faults on equipment performance was discussed and example fault impacts at different operating conditions were presented. The fault impact ratios discussed in this section can be used by an automated fault detection and diagnosis tool to recommend service.

In the closing section of this chapter, semi-empirical models intended to estimate the individual impacts of simultaneous faults are developed and implemented. First, the

methodology was evaluated using the outputs of a previously developed fault impact model that incorporates relatively complex component models. Using the simulated data sets, the fault impact models proposed predicted cooling capacity and COP within 10% of the detailed model outputs.

4. MODELING DX EQUIPMENT FAULT IMPACTS USING ARTIFICIAL NEURAL NETWORKS

4.1 Background and Motivation

This chapter develops comprehensive fault impact models for inclusion within the simulation platform described in Chapter 5 so that alternative service maintenance strategies, including optimal service scheduling, could be studied in Chapter 6. A commonly held belief among engineers is that to truly understand the underlying phenomena of a physical system, an engineer needs to formulate a model of such a system based on the known principles of nature governing the processes involved. As the complexity of a system grows, it's generally accepted that the model required for estimating the physics of the must also be extended to account for additional variables that affect its performance. Of course, adding complexity to models can often require ever more increasing engineering and computational effort (with sometimes diminishing returns). Because of this tradeoff, every engineer must determine the balance between complexity and effort that is suited for each project at hand.

One example of a relatively complex approach to model the performance impacts of common faults affecting unitary air conditioning equipment has been described by Cheung and Braun [48, 49]. In this work, Cheung and Braun formulated detailed models of common air conditioning equipment components using engineering principles. Components modeled in this work included compressors, air cooled condensers, expansion devices, cooling coils, as well as refrigerant piping. To tune difficult or unmeasurable parameters used in component models, Cheung and Braun used inverse modeling methods to estimate these parameters using experimental data collected from actual systems. Finally, component models were integrated

into system models which imposed additional constraints to maintain conservation of mass and energy throughout the closed system.

In comparison with other methodologies developed for estimating fault impacts of DX air conditioning equipment, the methodology presented by Cheung and Braun implements one of the most detailed and representative approaches for predicting the impacts of simultaneous faults. When using the model, the user can be assured that energy balances across the heat exchangers are maintained over the range of inputs. Additionally, differences between mass flow rates through the compressor and the expansion valves predicted by the model are driven to zero by the cycle solver. Also, important in consideration of refrigerant charge faults, detailed charge inventory modeling based in part on empirical relationships tuned using experimental data is performed by the cycle solver.

The detailed fault impact model developed by Cheung and Braun is not without one significant cost: the computational effort required for some combinations of faults and operating conditions can be significant. Even with the extreme speed of modern computer processors, simulating equipment performance using the detailed model can be time-consuming. In practice, the model often requires more than 30 seconds and multiple computer processor cores to simulate a single combination of inputs. When infrequent or one-off model evaluations are required, e.g. estimating steady-state fault impacts at a specific operating condition, the amount of computer time required is still insignificant to warrant optimization. If frequent or routine model function calls are required for an application, this time delay can be problematic. For example, if the average time requirement to predict equipment fault impacts is only 5 seconds, an 8760-hourly building simulation would require approximately 12 hours of computer time. As a result, this detailed model is not practical for inclusion within the simulation platform described

in Chapter 5 that is used for assessing optimal and alternative service strategies presented in Chapter 6.

The remainder of this chapter describes a meta-modeling approach used to reduce the computational effort required to simulate DX equipment fault impacts without significantly sacrificing model accuracy. A general discussion of meta-modeling with an emphasis placed on using neural networks is provided in Section 4.2. Following this, a more detailed description of the artificial neural network models used to simulate equipment fault impacts is included. The model training and evaluation methodology is described in Section 4.3 for a RTU with a fixed orifice expansion device. Finally, a review of the important results is provided in the Section 4.4.

4.2 Fault Impact Meta-Modeling Approach

The detailed fault impact model developed by Cheung and Braun was used to generate outputs for training of a meta-model so that fault impact predictions could be provided at a much faster rate. This process, shown in Figure 4.1, is not unlike the process used to develop the semi-empirical models in Chapter 3. First, an extensive list of ambient conditions and fault levels were generated that spanned the range of operating conditions that were desired to be modeled. Next, these model inputs were fed to the detailed fault impact inverse model and one-by-one the outputs were collected. These input and output combinations were collated and consolidated into a training and validation data set.

Next, the internal structure of an artificial neural network model was generated, and the internal weights were tuned using a supervisory learning setting. This means that a portion of the generated dataset was used to optimize the parameters of the neural network by minimizing the mean absolute error of the predicted outputs in comparison to the detailed model outputs. In order to do this in an efficient manner, highly optimized backpropagation algorithms and

software packages were used [65, 66]. In an iterative manner, the resulting model was tested using a reserved portion of the generated data set to evaluate how well the model predicts data points not contained in the training data set. If undesirable or significant errors exist in the resulting model, the model structure was modified, and the process was repeated.

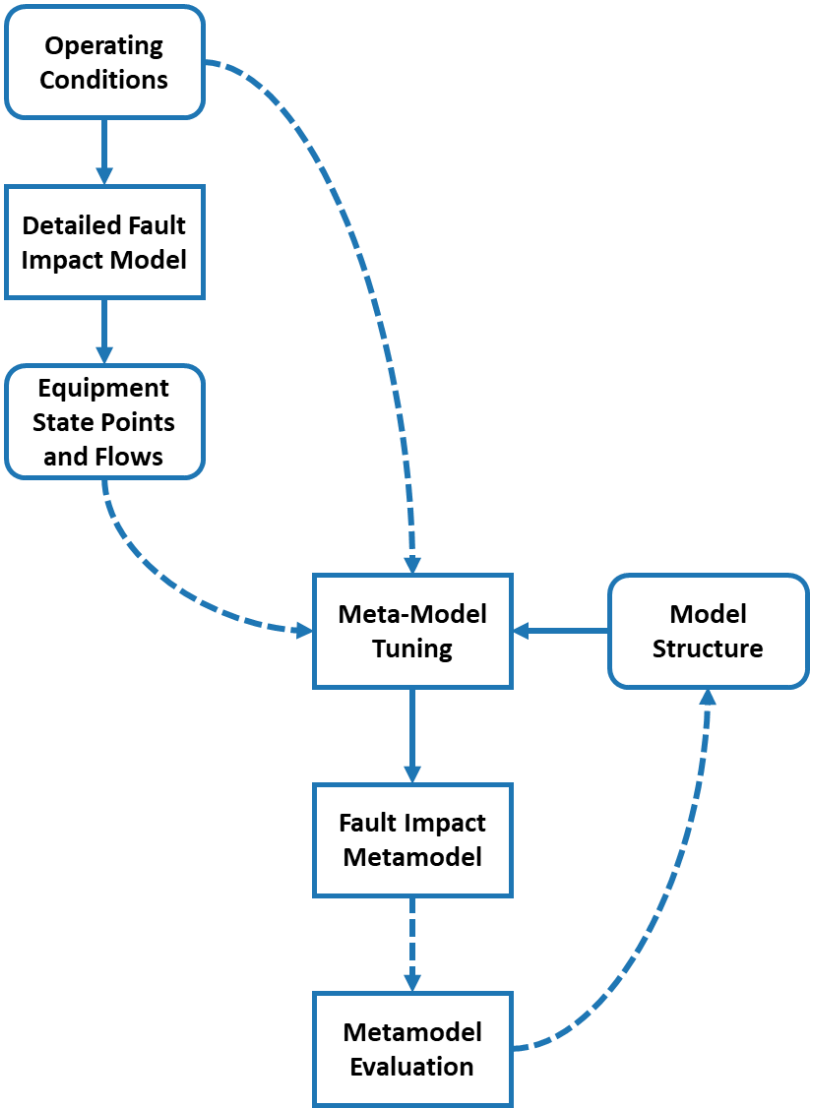

Figure 4.1. Overview of supervisory learning process used to develop fault impact meta-model using detailed fault impact model inputs and outputs.

A simplified schematic representation of the underlying neural network inputs, outputs, and inner layer nodes is shown in Figure 4.2. The inputs to the model are the ambient driving

conditions of the DX system, including the outdoor air dry-bulb, return air dry-bulb, and return-air wet-bulb. In addition, the fault levels of the system are given as inputs to the model. The outputs of the neural network meta-model are the refrigerant- and air-side state points that are determined by the original detailed fault impact model. These include suction pressure and enthalpy, liquid-line pressure and enthalpy, supply air temperature and humidity, etc. Connecting the inputs to the outputs are three layers containing so-called hidden nodes. These nodes are non-linear activation functions that are sequentially connected. The activation functions are typically sigmoidal or functions with an asymptote that are expressed in terms of a weighted combinations of inputs that are connected to it. These functions provide a highly non-linear behavior that can approximate complex interactions between inputs. The function weights used in each of these nodes are tuned during the model training process with backpropagation via stochastic gradient descent.

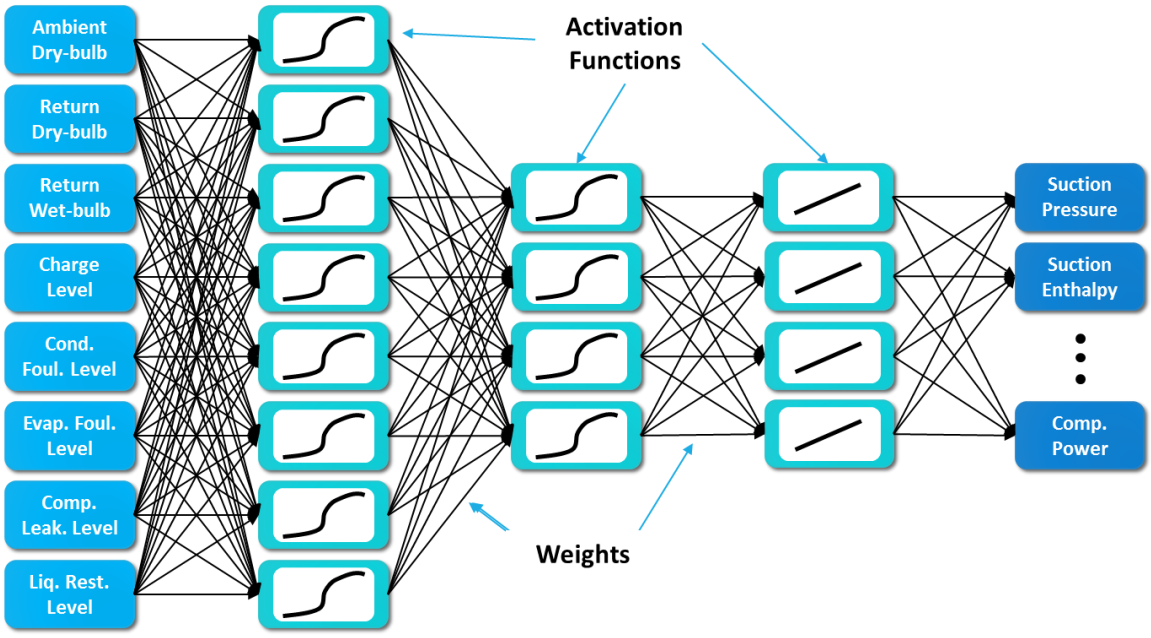

Figure 4.2. Simplified representation of the neural network meta-model used to approximate the detailed fault impact model developed by Cheung and Braun [48, 49].

To train the model while avoiding overfitting, the number of layers and hidden nodes were pruned using the combination of a training data set and a test data set. The activation functions selected for the model included sigmoidal functions in the first two layers and a linear activation layer in the final layer. The training data set consists of 80% of the total cases output from the original detailed model and the test data set was made up of the remaining 20% of the outputs. The parameter tuning process of the model required approximately 90 minutes of computer time. Evaluating the resulting model required on the order of 10 μs for a combination of inputs – significantly faster than the original detailed model.

4.3 Fault Impact Meta-model Training and Results

4.3.1 Overall Prediction Accuracy of Meta-model

Using the RTU with FXO meta-model, refrigerant state points were predicted over the range of operating conditions and fault levels that were simulated using the detailed model implemented by Cheung and Braun [48, 49]. These meta-model predictions were then compared with the detailed model outputs to evaluate the accuracy and reliability of using the meta-model. The suction pressure and enthalpy predicted by both models are compared for the entire dataset in Figure 4.3. In both cases, the agreement between the suction pressure and suction enthalpy were within 10% error for greater than 99.99% of the data set. Furthermore, the suction pressure predictions were within 5% error for greater than 90% of the data set; the suction enthalpy predictions were within 5% error for greater than 95% of the data set. It should also be noted that the range of suction pressures included in the data set was relatively large. The suction pressure ranged from 350 kPa (51 psia) to 1300 kPa (189 psia). At these pressures, the evaporator saturation temperature ranged from -23.1 °C (-9.5 °F) to 16.3 °C (61.4 °F) respectively. At these low suction pressures, an actual RTU would more than likely be

deactivate by low pressure cut-out safety or form a giant block of ice on the evaporator coil. These physics are not captured by the detailed model and the outputs under these conditions should be taken cautiously. While unrealistic, these low pressures were the result of significant evaporator air flow reduction and charge leakage.

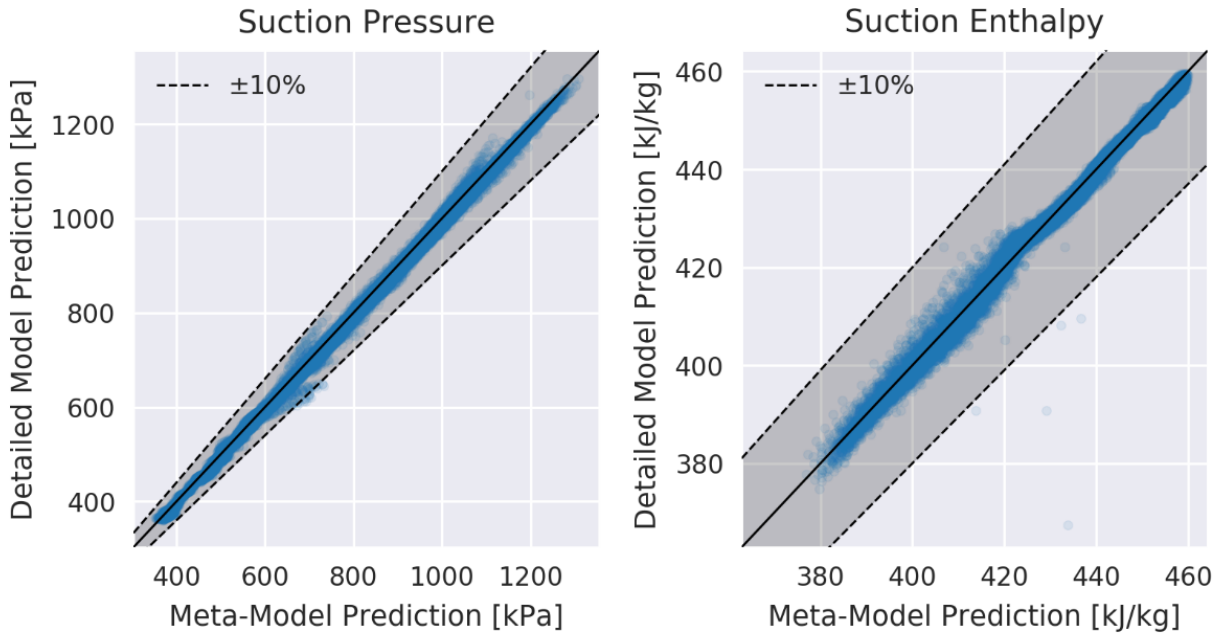

Figure 4.3. Overall comparison of suction pressure and enthalpy predictions from fault impact meta-model and detailed model for system with fixed orifice expansion valve.

The discharge pressure and enthalpy were simulated using both models and compared in Figure 4.4. These results also showed agreement between the discharge pressures and enthalpy within 10% for greater than 99.99% of the complete data set. Like the suction pressure results, the data set showed a relatively large range in discharge pressures: 1700 kPa (247 psia) to 4700 kPa (682 psia). Since the critical pressure of R410A is approximately 4578 kPa (664 psia), these data points should have been discarded from the data set. Moreover, these high pressures are surely outside the operating envelope of the compressor and the system would be disabled on high pressure cut-out safeties.

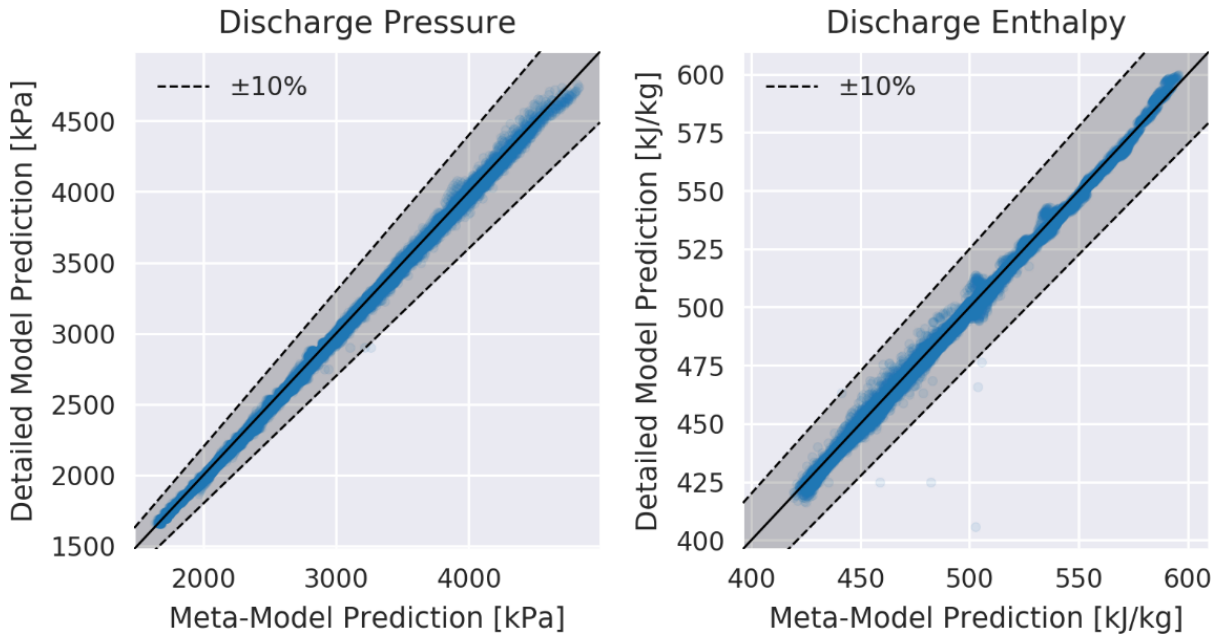

Figure 4.4. Overall comparison of discharge pressure and enthalpy predictions from fault impact meta-model and detailed model for system with fixed orifice expansion valve.

The liquid-line (condenser outlet) refrigerant pressure and enthalpy predicted by both models for the entire data set are compared in Figure 4.5. The predictions generated by each model for the liquid line pressure and enthalpy were within 5% for greater than 99.99% of the entire data set. As with the discharge pressure predictions, some of the pressures resulted in operating pressures much greater than typical equipment operating limits. These outputs should be discarded from the dataset for model evaluation since they would result in unreasonable equipment performance.

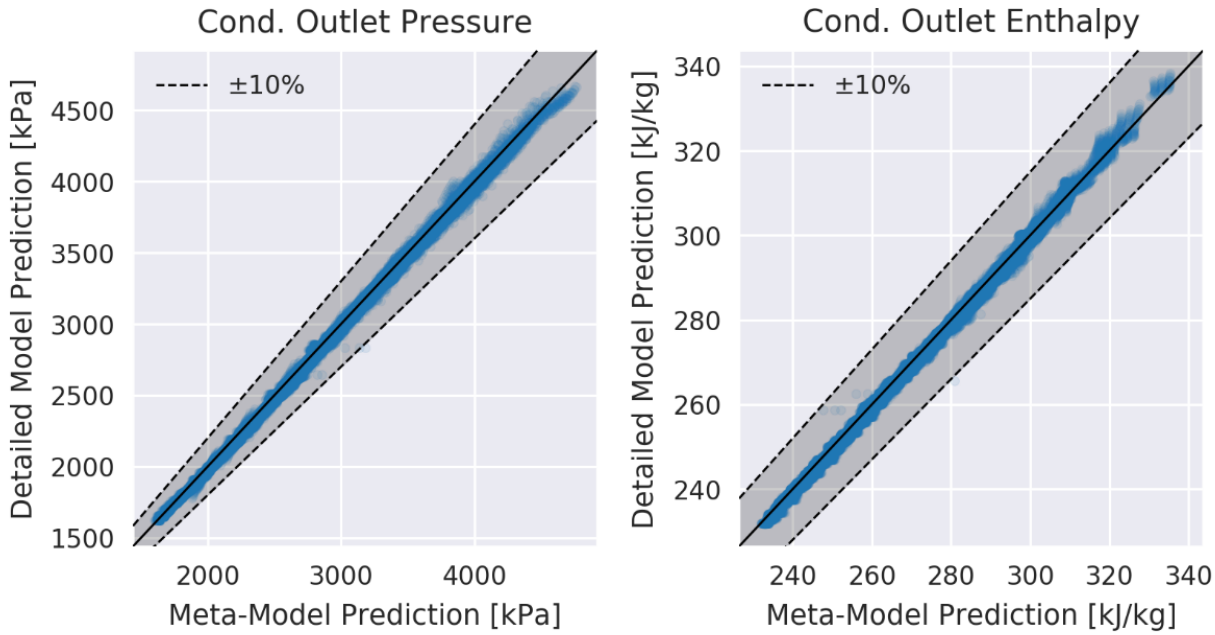

Figure 4.5. Overall comparison of liquid line (condenser outlet) pressure and enthalpy predictions from fault impact meta-model and detailed model for system with fixed orifice expansion valve.

The evaporator inlet refrigerant pressure and enthalpy predicted by both models for the entire data set are compared in Figure 4.6. The predictions generated by each model for the evaporator inlet pressure were within 10% for greater than 99.99% of the entire data set. The predictions generated by each model for the evaporator inlet enthalpy were within 5% for greater than 99.99% of the entire data set. As with the suction pressure predictions, some of the pressures resulted in operating pressures that are much lower than typical operating limits and would result in significant ice formation. Moreover, these data points would result in operation outside of compressor manufacturers recommended operating envelope and premature compressor failure would be likely. For these reasons, data points outside of typical operating limits should be discarded when model evaluation is performed.

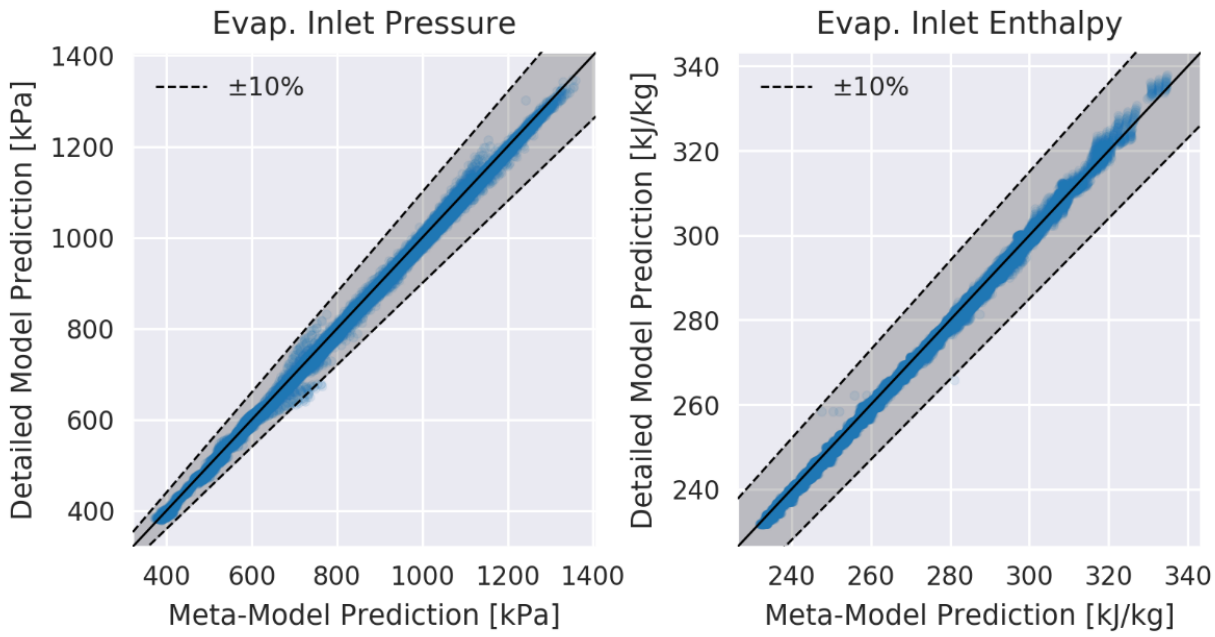

Figure 4.6. Overall comparison of evaporator inlet pressure and enthalpy predictions from fault impact meta-model and detailed model for system with fixed orifice expansion valve.

4.3.2 Refrigerant Charge Fault Impact Prediction Accuracy

Using the tuned neural network meta-model, the performance of the RTU was simulated at different ambient operating conditions and fault levels. The predicted total cooling capacity, *COP*, and *SHR* of the system at different ambient temperatures and fault levels has been plotted in comparison with the outputs of the original detailed model in Figure 4.7. Close agreement between the outputs of the meta-model and the detailed model were observed. More importantly, the meta-model outputs do not display signs of overfitting since the resulting outputs were mostly smooth, even where interpolation and extrapolation were required.

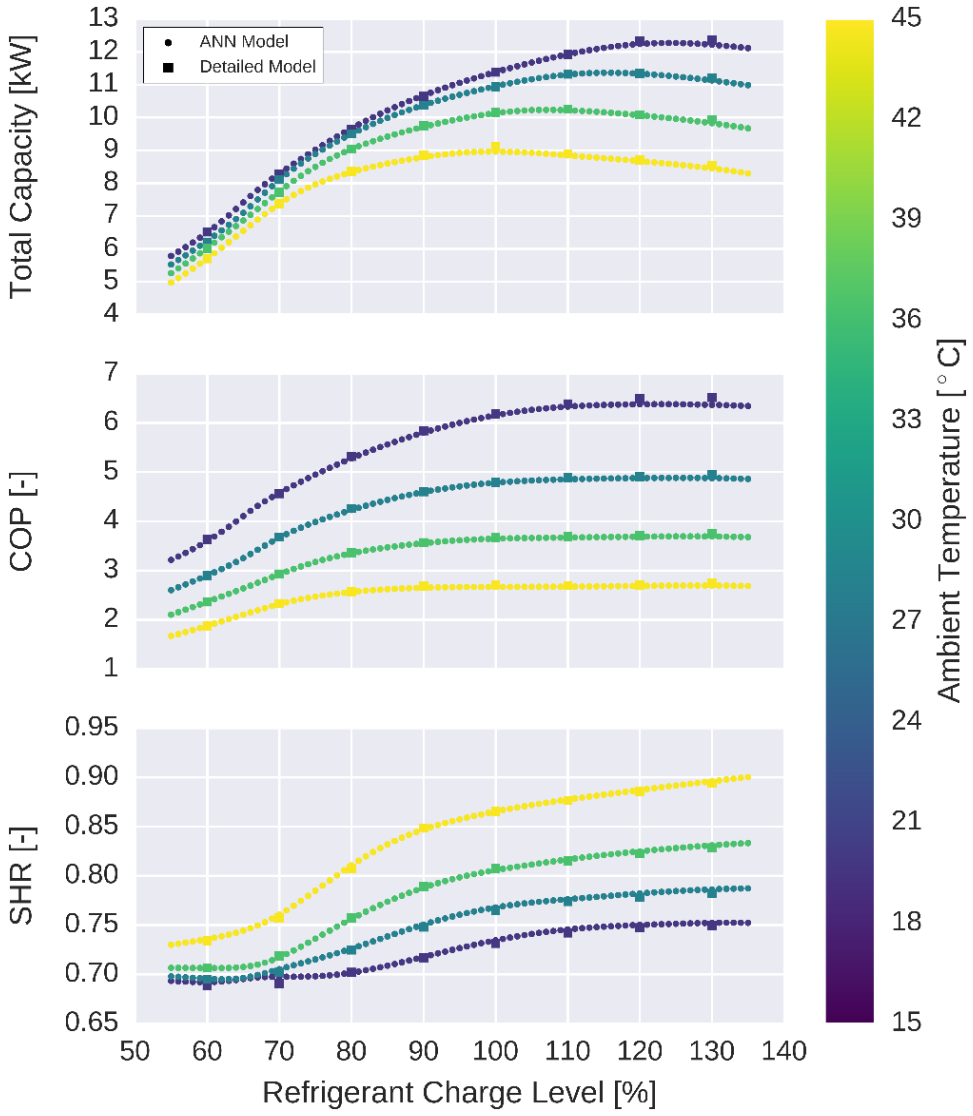

Figure 4.7. Modeled total cooling capacity, cycle *COP*, and *SHR* for RTU with fixed orifice expansion valve at different levels of refrigerant charge.

4.3.3 Condenser Fouling Fault Impact Prediction Accuracy

The meta-model was used to predict the performance of the RTU when the system is subjected to different levels of condenser fouling in Figure 4.8. Like the results shown for varying levels of refrigerant charge, good agreement was seen between the meta-model outputs and the original detailed model for condenser fouling faults at different ambient temperature

conditions. The model also predicted more significant impacts on RTU efficiency than total cooling capacity as would be expected by condenser fouling faults.

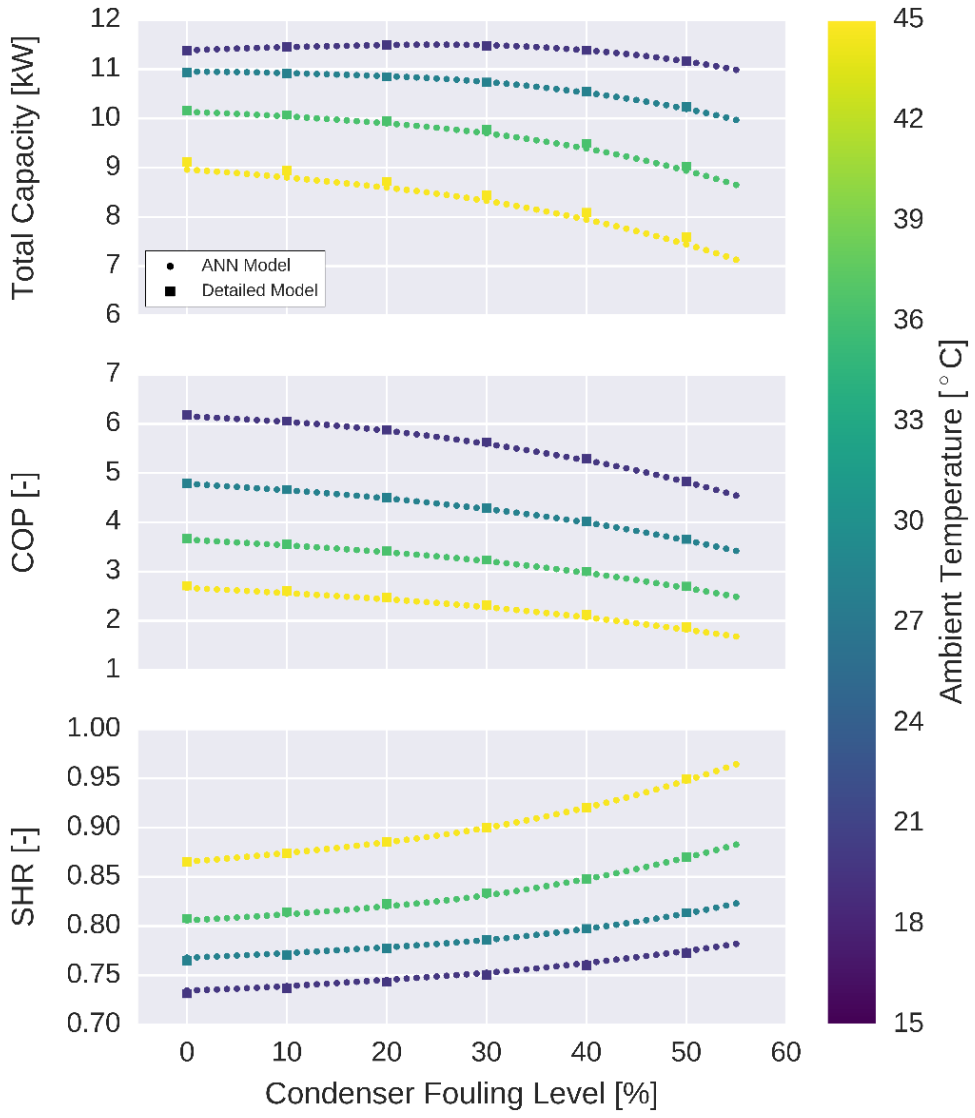

Figure 4.8. Comparison of modeled total cooling capacity, cycle COP, and SHR for RTU with fixed orifice expansion valve over a range of condenser fouling levels.

4.3.4 Evaporator Fouling Fault Impact Prediction Accuracy

A comparison of the predicted RTU performance with different combinations of ambient temperatures and evaporator fouling levels is shown in Figure 4.9. Again, good agreement

between the two models is observed suggesting the neural network model approximates the physics of the detailed model well. Furthermore, it is important to note that both models predict relatively small impacts on cycle efficiency but large impacts on total cooling capacity and sensible heat ratio. This indicates both models produce outputs that are representative of actual evaporator fault impacts.

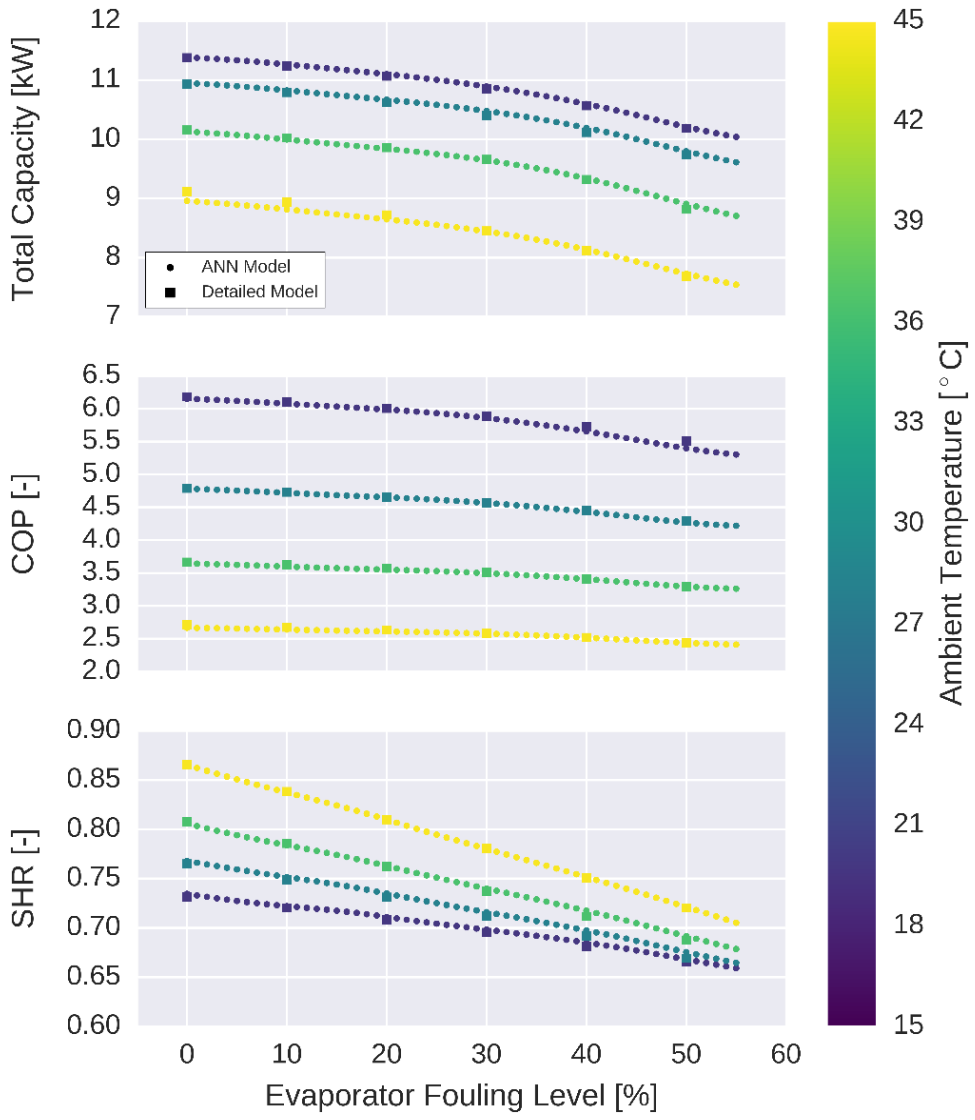

Figure 4.9. Comparison of modeled total cooling capacity, cycle COP, and SHR for RTU with fixed orifice expansion valve over a range of evaporator fouling levels.

4.3.5 Compressor Valve Leakage Fault Impact Prediction Accuracy

The performance impacts of compressor valve leakage faults are compared for both models in Figure 4.10. Good agreement between the neural network meta-model and the detailed fault impact inverse model is observed. The behavior of the SHR of the meta-model is of interest for high levels of compressor valve leakage. In these cases, the elevated evaporator temperature and additional suction superheat tends to promote dry-coil operation. While the change between wet-coil and dry-coil is difficult to capture for linear models, the neural network can accurately capture this phenomenon.

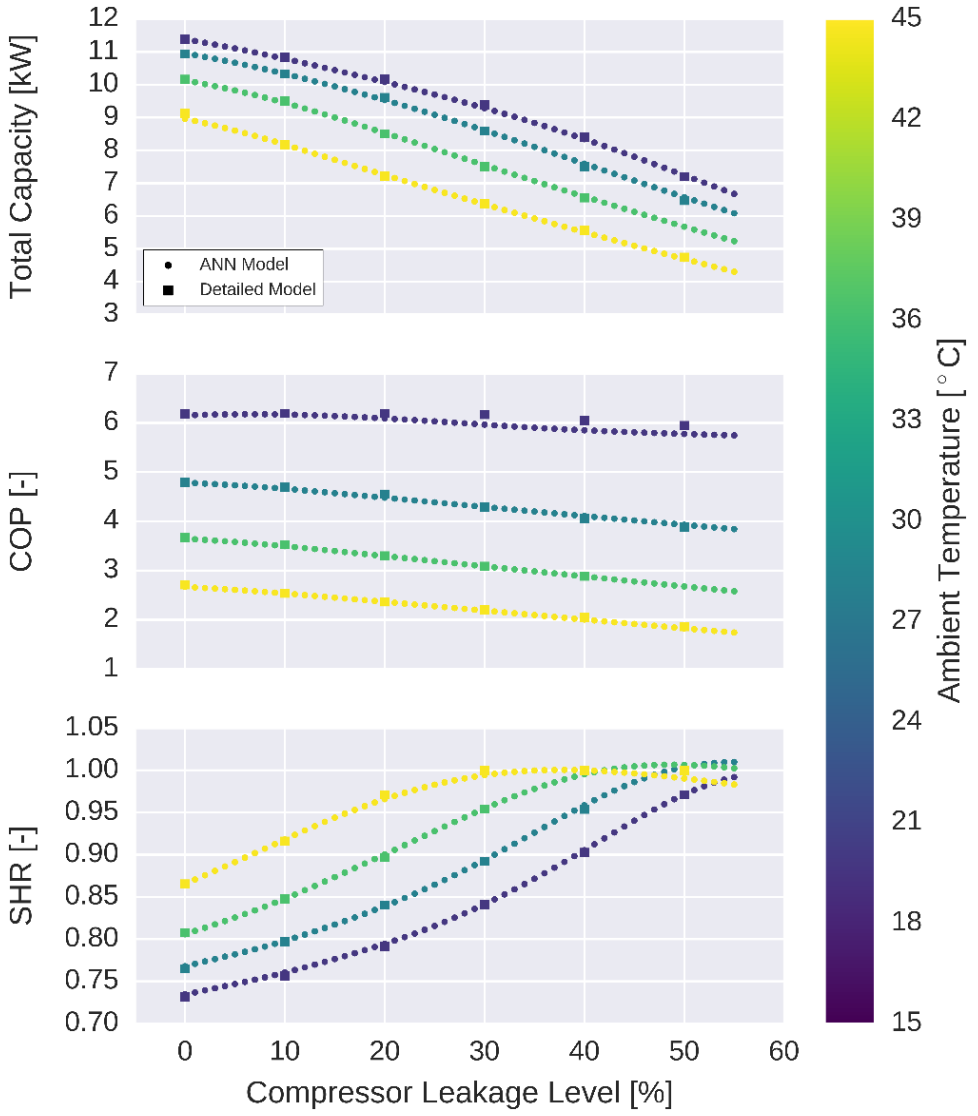

Figure 4.10. Comparison of modeled total cooling capacity, cycle COP, and SHR for RTU with fixed orifice expansion valve over a range of compressor valve leakage levels.

4.4 Review of Important Results

The neural network meta-model has been shown to generate accurate results that approximate the outputs of the grey box fault impact inverse model developed by Cheung and Braun [48, 49]. The outputs of this meta-model overcome one of the major limitations of the original model – the solution time required. One issue remains with both models – the amount of experimental data required to accurately tune the model is relatively high and the resulting model

is not generalizable to all systems. Because of this, it is unlikely that a manufacturer could devote enough testing resources to develop these models.

While it is unlikely that manufacturers would be able to develop sufficient testing resources for developing neural network models of equipment performance to predict the impacts caused by different faults, it should be reiterated that these models are useful for the present work. In an ideal setting, it would be possible to use detailed physical models for analyzing the performance of fault impact models and performing parametric studies of different optimal service scheduling strategies. However, with the current performance limits of computing systems, it would not be possible to adequately analyze all the possible combinations of faults in a reasonable amount of time. These limitations could be overcome using the neural network models described in this work. Using the extensive datasets collected from previous researchers, the detailed fault impact models developed by Cheung and Braun were possible [48, 49]. Rather than applying this model directly, the neural network models were designed to approximate the relationship between the inputs and outputs of the detailed inverse models with much greater speed. The methodology could be applied to any detailed physical model to approximate outputs using much less computing resources. These models can then be used to develop less complex models that manufacturers are able to implement.

In conclusion, the neural network models developed are not intended to be viewed as a solution to be used to estimate fault impacts in practice. Rather, they will be used to assess simplified models and strategies for implementing condition-based maintenance. From the validation presented, the neural network models approximate the detailed fault impact model well, and thus its outputs can be trusted. Furthermore, conclusions derived from the outputs of

the neural network models can be validated in comparison to conclusions derived from the original detailed model.

5. SIMULATING AIR CONDITIONING EQUIPMENT FAULTS IN SMALL COMMERCIAL BUILDINGS

5.1 Background and Motivation

In the previous Chapter, an artificial neural network DX equipment fault impact meta-model was described and evaluated using a generated data set. The goal of this model was to accelerate the predictions of fault impacts so that they could be used in a building simulation program to characterize the long-term impacts of different faults and evaluate different maintenance strategies. In this chapter, the building simulation framework that was implemented for this work is described.

Components that are described in detail include the building sensible and latent load models. These models have been developed based on single-node dynamic sensible and latent energy balances of a building served by an air conditioning system. Descriptions of the weather data used for the simulation and the locations selected for comparison are also included. The operating cost model has also been described, including electrical utility costs, equipment costs, and service costs. Central to the model are fault evolution models that are used to estimate how faults may grow over time. These models are simplistic in nature but provide reasonable estimations of real systems. Finally, demonstration results are included to show how the simulation models interact.

5.2 Building Load Model Description

The sensible and latent dynamics of a building served by a single RTU were simulated using single-node models. The dynamics of the indoor dry bulb temperature were simulated using

$$C_s \frac{dT_{ID}}{dt} = \dot{Q}_{load,s} - \dot{Q}_{cool,s} \tag{5.1}$$

where T_{ID} is the indoor temperature, $\dot{Q}_{load,s}$ is the sensible heat gain in the building, $\dot{Q}_{cool,s}$ is the sensible cooling rate of the RTU, and C_s is the effective thermal capacitance of the building. The sensible heat gain of the building was based on an internally driven component $\dot{Q}_{int,s}$ and an externally driven component $\dot{Q}_{ext,s}$,

$$\begin{aligned}\dot{Q}_{load,s} &= \dot{Q}_{int,s} + \dot{Q}_{ext,s} \\ &= \dot{Q}_{int,s} + UA_s \cdot \left(T_{OD} - T_{ID}\right)\end{aligned} \tag{5.2}$$

where UA_s is the overall heat transfer conductance of the building and T_{OD} is the outdoor air dry bulb temperature. The sensible cooling rate of the system was simulated using the neural network meta-model discussed in Chapter 4. The required inputs of the meta-model were the outdoor dry bulb temperature, indoor dry bulb and wet bulb temperatures, and the fault levels of the equipment.

The internal gain and overall heat transfer conductance used in Equation (5.2) were determined for different climates using the commonly used balance point methodology based on equipment sensible capacity at the warmest condition [61]. The overall conductance was determined using

$$UA_s = \frac{\dot{Q}_{cool,s,design}}{1+f_{os}} \cdot \frac{1}{T_{OD,design} - T_{OD,balance}} \tag{5.3}$$

where $\dot{Q}_{cool,s,design}$ is the design point sensible cooling capacity of the equipment without faults, f_{os} is an equipment oversizing factor, $T_{OD,design}$ is the design point outdoor air dry bulb temperature, $T_{ID,design}$ is the design point indoor dry bulb temperature, and $T_{OD,balance}$ is the balance point outdoor dry bulb temperature when the building has no load. For all climates used for simulations, the design point temperature was chosen by finding the maximum outdoor air

temperature using the weather data. The design point indoor dry bulb temperature setpoint was set to 75 °F. The balance point outdoor temperature was set to 65 °F for all climates. The internal sensible gain is given by

$$\dot{Q}_{int,s} = \frac{\dot{Q}_{cool,s,design}}{1+f_{os}} \cdot \frac{T_{ID,design} - T_{OD,balance}}{T_{OD,design} - T_{OD,balance}}. \tag{5.4}$$

An approximation of Equation (5.1) was implemented using a finite difference formula for the derivative term

$$T_{ID}(k+1) = T_{ID}(k) + \frac{\Delta t_{sim}}{C_s}\left[\dot{Q}_{int,s} + UA_s \cdot \left(T_{OD}(k) - T_{ID}(k)\right) - \dot{Q}_{cool,s}(k)\right] \tag{5.5}$$

where Δt_{sim} is the simulation time step size, $T_{ID}(k)$ is the indoor temperature at the k^{th} simulation step, and $T_{ID}(k+1)$ is the indoor temperature at the $(k+1)^{\text{th}}$ step. The simulation step size used for the model was 60 minutes.

The latent dynamics of the building were modeled using a single node dynamic energy balance,

$$C_l \frac{d\omega_{ID}}{dt} = \frac{1}{h_{fg,w}}\left[\dot{Q}_{load,l} - \dot{Q}_{cool,l}\right] \tag{5.6}$$

where ω_{ID} is the indoor air humidity ratio, $\dot{Q}_{load,l}$ is the latent heat gain in the building, $\dot{Q}_{cool,l}$ is the sensible cooling rate of the RTU, $h_{fg,w}$ is the heat of vaporization of water, and C_l is the effective thermal capacitance of the building. The latent heat gain of the building was based on an internally driven component $\dot{Q}_{int,l}$ and an externally driven component $\dot{Q}_{ext,l}$,

$$\begin{aligned} \dot{Q}_{load,l} &= \dot{Q}_{int,l} + \dot{Q}_{ext,l} \\ &= \dot{Q}_{int,s} + U_l \cdot \left(\omega_{OD} - \omega_{ID}\right) \end{aligned} \tag{5.7}$$

where U_l is the overall latent conductance of the building and ω_{OD} is the outdoor air humidity ratio. The conductance was determined using the equipment latent heat transfer rate at the design point condition,

$$U_l = \frac{\left(1 - SHR_{design}\right) \cdot \dot{Q}_{cool,t,design}}{1 + f_{os}} \cdot \frac{1}{\omega_{OD,design} - \omega_{ID,balance}} \tag{5.8}$$

where $\dot{Q}_{cool,t,design}$ is the design point total cooling capacity of the equipment without faults, SHR_{design} is the design sensible heat ratio of the building, f_{os} is an equipment oversizing factor, $\omega_{OD,design}$ is the design point outdoor air humidity ratio, $\omega_{ID,design}$ is the design point indoor air humidity ratio, and $\omega_{OD,balance}$ is the balance point outdoor air humidity ratio when the building has no latent load. The design point indoor air humidity ratio was calculated using an indoor dry bulb temperature of 75 °F and relative humidity equal to 50%. For all climates used for simulations, the design point humidity ratio was chosen by finding the mean coincident wet bulb of the design outdoor air dry bulb temperature. The balance point outdoor temperature was determined by finding the mean humidity ratio throughout the year using climate data. The internal latent gain is given by

$$\dot{Q}_{int,l} = \frac{\left(1 - SHR_{design}\right) \dot{Q}_{cool,t,design}}{1 + f_{os}} \cdot \frac{\omega_{ID,design} - \omega_{OD,balance}}{\omega_{OD,design} - \omega_{OD,balance}}. \tag{5.9}$$

An approximation of Equation (5.6) was implemented using a finite difference formula for the derivative term

$$\omega_{ID}\left(k+1\right) = \omega_{ID}\left(k\right) + \frac{\Delta t_{sim}}{C_l \cdot h_{fg,w}} \left[\dot{Q}_{int,l} + U_l \cdot \left(\omega_{OD}\left(k\right) - \omega_{ID}\left(k\right)\right) - \dot{Q}_{cool,l}\left(k\right) \right] \tag{5.10}$$

where $\omega_{ID}(k)$ is the indoor air humidity ratio at the k^{th} simulation step, and $\omega_{ID}(k+1)$ is the indoor air humidity ratio at the $(k+1)^{th}$ step.

5.2.1 Climate and Weather Model Description

In order to simulate equipment performance under realistic outdoor ambient conditions, Typical Meteorological Year (TMY) data was used [67]. TMY data sets provide hourly ambient temperature, humidity, barometric pressure, irradiance, wind speed, and other weather data that represent a year of typical climatic conditions for a location [67]. TMY data is commonly used by designers to model HVAC and energy conversion systems since the data provides reasonable driving conditions for assessing and comparing technologies. Three locations were selected for fault impact simulations: Atlanta, GA; Chicago, IL; and Miami, FL. These locations were selected because they have different amounts of cooling and heating requirements. Additionally, the average humidity through the year varies considerably from location to location. For each location, design conditions are summarized in Table 5.1 [61].

Table 5.1. Building HVAC design conditions for selected locations used in fault impact simulation program [61].

Location	HDB (99%)	CDB (1%)	MCWB (1%)	HDD	CDD
Atlanta, GA	-3.1 °C (26.4 °F)	33.1 °C (91.6 °F)	23.3 °C (73.9 °F)	1484 °C-day (2672 °F-day)	1052 °C-day (1894 °F-day)
Chicago, IL	-15.7 °C (3.7 °F)	31.5 °C (88.7 °F)	22.9 °C (73.2 °F)	3449 °C-day (6208 °F-day)	480 °C-day (864 °F-day)
Miami, FL	11.1 °C (52.0 °F)	32.7 °C (90.9 °F)	25.3 °C (77.5 °F)	70 °C-day (126 °F-day)	2521 °C-day (4538 °F-day)

HDB (99%) – 99th percentile design heating dry bulb temperature.
CDB (1%) – 1st percentile design cooling dry bulb temperature.
MCWB (1%) – 1st percentile design mean coincident wet bulb temperature.
HDD – Heating degree days calculated using 18.3 °C (65.0 °F) balance temperature.
CDD – Cooling degree days calculated using 18.3 °C (65.0 °F) balance temperature.

In order to size the building sensible cooling and heating load line in Equation (5.2), the maximum dry-bulb temperature for each location was determined. For Atlanta, GA the maximum dry-bulb temperature occurred on July 8th, shown in Figure 5.1. To determine the latent load line parameters in Equation (5.7), the coincident maximum humidity ratio during the hottest day of the year was determined. For Atlanta, GA the humidity ratio and relative humidity during the warmest day of the year are shown in Figure 5.1. Compared with the other selected locations, Atlanta, GA has relatively warm summer days with mild spring and summers. Atlanta, GA has moderate winter heating requirements compared with the other locations in the study.

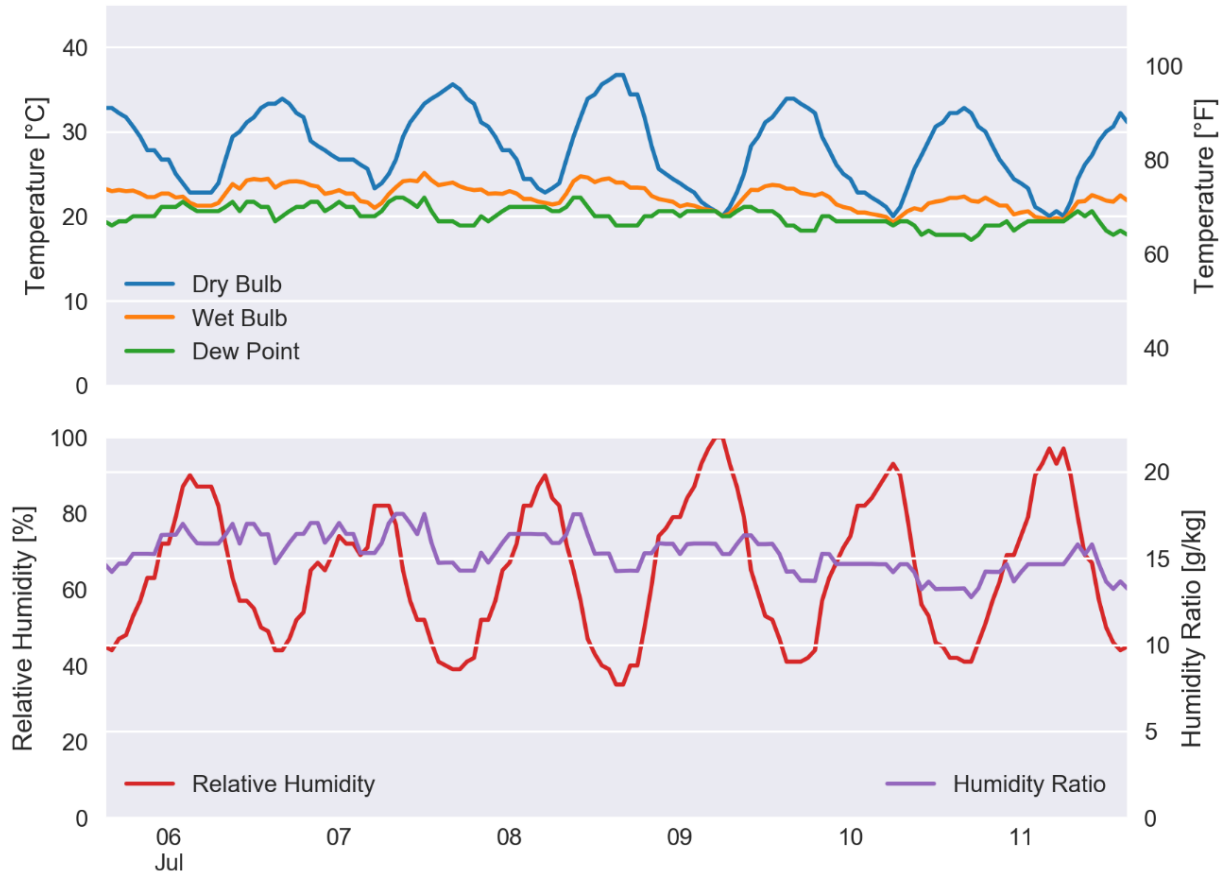

Figure 5.1. Hourly ambient temperatures and humidity for Atlanta, GA TMY3 data set for 72 hours before and 72 hours after peak dry bulb temperature.

The maximum dry-bulb temperature occurred on July 19th in Chicago, IL, shown in Figure 5.2. Compared with the other selected locations, Chicago, IL has a relatively short cooling season compare with the other locations in the study. Chicago, IL has the coldest winter outdoor temperatures and the highest heating requirements throughout the year.

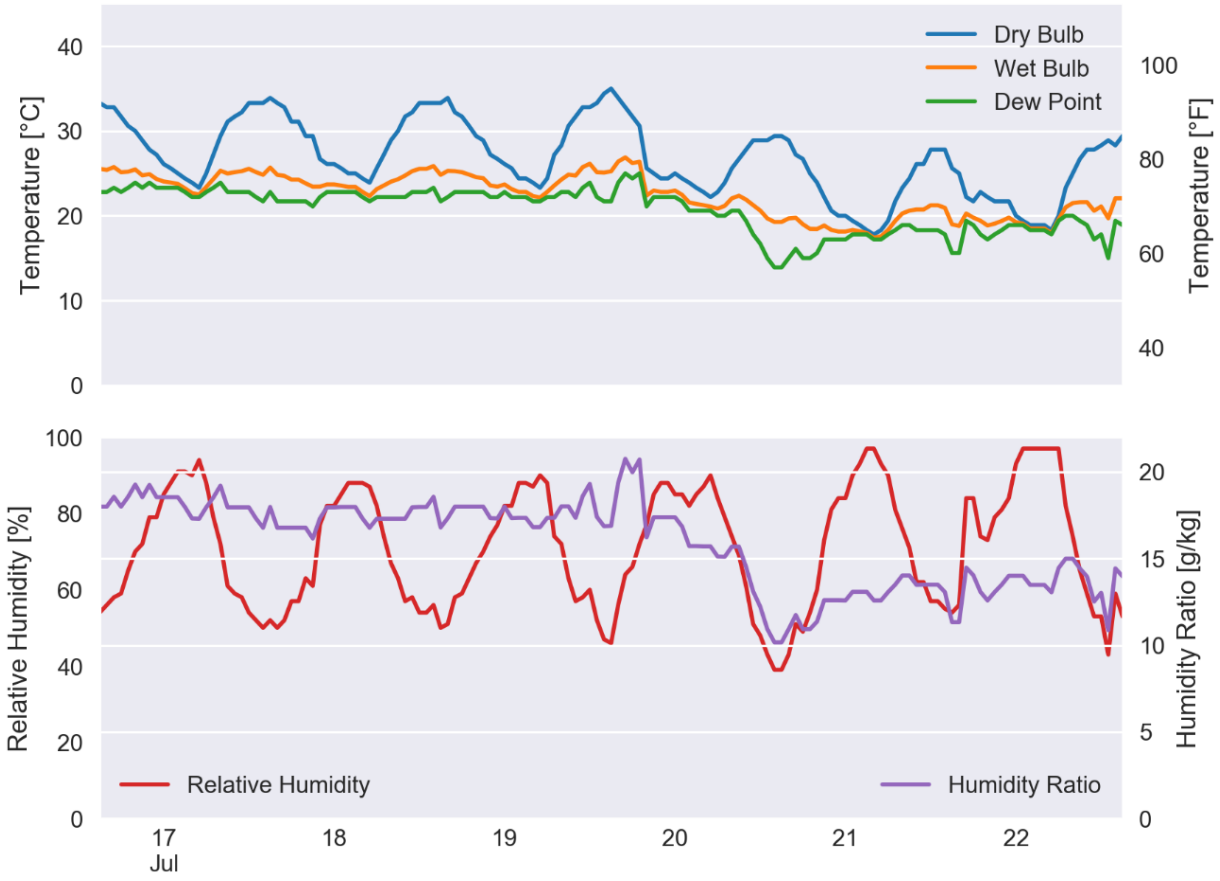

Figure 5.2. Hourly ambient temperatures and humidity for Chicago, IL TMY3 data set for 72 hours before and 72 hours after peak dry bulb temperature.

The maximum dry-bulb temperature occurred on July 29th in Miami, FL, shown in Figure 5.3. Compared with the other selected locations, Miami, IL has a long cooling season compared with the other locations in the study. It is also the one of the most humid climates through the year, compared with the other selections. Miami, FL has the shortest heating season of the locations selected in this study.

Figure 5.3. Hourly ambient temperatures and humidity for Miami, FL TMY3 data set for 72 hours before and 72 hours after peak dry bulb temperature.

The building sensible and latent building load parameters used in the equations implemented for each building location are summarized in Table 5.2. These values were held constant throughout the maintenance simulation study discussed in Chapter 6.

Table 5.2. Building sensible and latent load parameters used in simulation framework for different locations.

Parameter		Miami, FL	Atlanta, GA	Chicago, IL
$T_{OD,design}$ / $B_{OD,design}$	[°C]	32.7 / 25.3	33.1 / 23.3	31.5 / 22.9
$T_{ID,design}$ / $B_{ID,design}$	[°C]	23.9 / 19.2	23.9 / 19.2	23.9 / 19.2
$T_{OD,balance}$ / $B_{OD,balance}$	[°C]	18.3 / 13.3	18.3 / 12.8	18.3 / 10.3
f_{os}	[-]	0.20	0.20	0.20
$\dot{Q}_{cool,design}$	[W]	6330	6390	6480
UA_s	[W/°C]	366	360	409
$\dot{Q}_{s,int}$	[W]	2051	2014	2290
C_s	[kJ/°C]	1000	1000	1000
SHR_{design}	[-]	0.70	0.70	0.70
U_l	[kW/(g.w/g.da)]	198	190	203
$\dot{Q}_{l,int}$	[W]	247	249	253
C_l	[g.w/g.da/s]	5.0	5.0	5.0

5.2.2 Utility, Equipment, and Service Cost Models

A simple consumption utility cost model was used to calculate the electricity cost of operating the air conditioning system. In this model, the utility cost at simulate step *k* was calculated using

$$UC(k) = C_{elec} \cdot \dot{W}_{elec}(k) \cdot \Delta t_{sim} \tag{5.11}$$

where C_{elec} is the cost of one unit of electricity, $\dot{W}_{elec}$ is the average power consumption of the air conditioner during the time step, and Δt_{sim} is the simulation interval. The power consumption of the air conditioner was based on the operating conditions, the fault levels, and the efficiency of the system calculated using the equipment model.

To simulate the economic implications of extended equipment operating times due to reduced system capacity, an equipment cost model developed by Li and Braun was

implemented [14]. The equipment costs at each simulation time set were modeled as a function of equipment runtime,

$$EC(k) = C_{equip} \cdot \Delta t_{run} \tag{5.12}$$

where C_{equip} is the reduction in equipment value per unit time and Δt_{run} is the required runtime of the system during the simulation interval. To determine a reasonable value for C_{equip}, it was assumed that the devaluation of the equipment remains constant over the life of the system

$$C_{equip} = \frac{C_{replace}}{\Delta t_{life}} \tag{5.13}$$

where $C_{replace}$ is the equipment replacement cost and Δt_{life} is the expect number of run-time hours the system will operate over its life. Equation (5.12) adds an economic penalty for operating equipment with faults beyond the impact on utility costs. This is a reasonable assumption since equipment life is finite and operating with faults may reduce times to failure.

Besides the utility cost and equipment costs, which accumulate over time whenever cooling is required, an additional service cost is also considered in the simulation. The service cost is an interval fixed cost that is applied whenever service is performed,

$$SC(t) = \begin{cases} C_{trip} + \sum_{a_i \in A} C_{task}\left(a_i\right) & \text{if service is needed} \\ 0 & \text{otherwise} \end{cases} \tag{5.14}$$

where C_{trip} is a fixed trip cost that is required whenever the service technician must go to a site, C_{task} is the cost to perform a specific task to restore the state of the equipment, and a_i is a Boolean variable representing required service tasks for the set of all service tasks a technician may perform, A.

5.3 Fault Evolution Model Descriptions

The mass of refrigerant charge within the air conditioning system, m_{actual}, was modeled as a linear function of the time since service was last performed, $\Delta t_{service}$,

$$m_{actual} = \begin{cases} \left(1 - \frac{1}{2} \cdot \frac{\Delta t_{service}}{\Delta t_{chrg}}\right) \cdot m_{rated} & \text{for } 0 \le \Delta t_{service} \le \Delta t_{chrg} \\ \frac{1}{2} \cdot m_{rated} & \text{for } \Delta t_{service} > \Delta t_{chrg} \end{cases} \tag{5.15}$$

where m_{rated} is the rated amount of refrigerant in the system and Δt_{chrg} is the simulation time required for the system to leak 50% of the rated charge. In other words, Δt_{chrg} determines the refrigerant leakage rate from the system. Because the meta-model used to predict performance impacts of refrigerant charge faults cannot be trusted below 50% charge levels, the minimum refrigerant charge level in the simulation was constrained to 50%. After the time since last service becomes greater than the leakage time parameter, the refrigerant charge level is fixed at 50%. After refrigerant service is performed, the refrigerant charge in the system is returned to the rated level of charge.

Evaporator fouling was modeled by assuming fouling reduces evaporator air flow rate in the air conditioning system. As a function of the indoor fan run-time, Δt_{IDF}, the evaporator air flow rate, $\dot{V}_{ea,actual}$, was assumed to decrease linearly,

$$\dot{V}_{ea,actual} = \begin{cases} \left(1 - \frac{3}{5} \cdot \frac{\Delta t_{IDF}}{\Delta t_{ea,\,foul}}\right) \cdot \dot{V}_{ea,normal} & \text{for } 0 \le \Delta t_{IDF} \le \Delta t_{ea,\,foul} \\ \frac{2}{5} \cdot \dot{V}_{ea,normal} & \text{for } \Delta t_{IDF} > \Delta t_{ea,\,foul} \end{cases} \tag{5.16}$$

where $\Delta t_{ea,\,foul}$ is the indoor fan runtime required to cause a 60% reduction in air flow rate. When evaporator fouling service is performed in the simulation, the evaporator air flow rate is

returned to the normal level. To prevent equipment model inaccuracies, the minimum air flow rate was restricted to 40% of the normal air flow rate. In the model, the indoor fan is assumed to run 100% of the time during occupied periods and when cooling is required during unoccupied periods.

In a similar manner, condenser fouling was modeled as a function of outdoor fan run-time, Δt_{ODF}, since the last service was performed. When condenser fouling occurs, the air flow through the condenser, $\dot{V}_{ca,actual}$ was reduced linearly according to,

$$\dot{V}_{ca,actual} = \begin{cases} \left(1 - \frac{3}{5} \cdot \frac{\Delta t_{ODF}}{\Delta t_{ca,\,foul}}\right) \cdot \dot{V}_{ca,normal} & \text{for } 0 \le \Delta t_{ODF} \le \Delta t_{ca,\,foul} \\ \frac{2}{5} \cdot \dot{V}_{ca,normal} & \text{for} \quad \Delta t_{ODF} > \Delta t_{ca,\,foul} \end{cases} \tag{5.17}$$

where $\Delta t_{ca,\,foul}$ is the time required for the air flow through the condenser to reduce to 40% of the normal condenser air flow, $\dot{V}_{ca,normal}$. To prevent equipment model inaccuracies, the minimum condenser air flow rate was restricted to 40% of the normal air flow rate. Additionally, it was assumed that the outdoor fan operates only when the compressor operates, thus, the condenser fan run-time is equal to the run-time required for cooling.

In order to simulate comfort disturbances caused by high head pressure cut-out and low suction pressure cut-out safeties, the air conditioning system in the system was required to operate within an acceptable envelope. When the simulated discharge pressure was greater than 3890 kPa (550 psig), the cooling output of the system was assumed to be zero due to high-pressure limit cut-out. High pressure cut-outs most frequently were the result of system operating during warm ambient conditions with significant condenser fouling levels. When the suction pressure of the simulated system fell below 860 kPa (110 psig), no cooling was provided by the system since the low-pressure limit cut-out was assumed to be activated. The safeties

were automatically reset during the next simulation time step, provided the system pressures fell within the operating envelope.

6. COMPARING MAINTENANCE STRATEGIES FOR SMALL COMMERCIAL BUILDING AIR CONDITIONING EQUIPMENT

6.1 Background and Motivation

Like almost any other mechanical system, direct-expansion air conditioners require routine or unscheduled measures to maintain reliable and efficient operation. If an air conditioner is ignored or regular maintenance goes unscheduled, performance of the system will deteriorate over time. Determining the frequency of when to inspect or perform maintenance can be difficult since impacts of faults are relatively difficult to estimate. It is also difficult to quantify the benefits of preforming different maintenance and service tasks. Furthermore, systems can develop multiple different faults at the same time which makes service decisions even more complex.

In some cases, the operation of an air conditioning system may become completely suspended by a fault. For example, an air conditioner may fail to start when a motor capacitor fails or becomes weak over years of operation. When this happens, no cooling will be available until the capacitor is replaced. It is relatively easy to detect faults that totally halt system operation – when cooling is not available, comfort in the condition space cannot be controlled. Additionally, for many applications in commercial buildings the decision about how to handle these faults is easy: the fault must be fixed or the system must be replaced as soon as possible. This is because preserving occupant comfort is usually a high priority in most commercial buildings.

Some other faults do not totally suspend the operation of a system, but rather degrade overall system performance. In other words, faults may decrease the amount of cooling capacity available or the efficiency of the system, but the system can still maintain comfort in the space.

An example of a fault that does not completely disable operation of an air conditioner is condenser fouling. When a condenser becomes fouled, an air conditioner is still able to deliver cooling to a conditioned space; however, it does so less efficiently. These faults are more difficult to detect than faults that halt operation – from an occupants' perspective it may not be noticeable at all. It may also be difficult to determine if faults are present by comparing utility bills as well.

For these types of faults which may degrade or deteriorate performance over time, maintenance decisions are less straightforward. This is because costs required to fix or repair a system may be comparable to the impact that the fault has on utility costs. For some faults, the cost to repair may be more than the benefit incurred. For others, economic benefits for repair may outweigh these service costs, though it still may be difficult to quantify this benefit leading to uncertainty.

6.2 Literature Review of DX Equipment Service Decision Support

Limited work on service decision support systems for DX equipment has been published in the available literature. While there has been extensive work done related to the fields of industrial engineering, much of this work has been focused on infrastructure, manufacturing processes, and large engineered systems and fleets [68-72]. Low-cost systems, like DX equipment have not been as widely studied.

Rossi and Braun developed a near optimal methodology to recommend service for RTUs using historical measurements of utility costs and a model for normal equipment operation [73]. The approach was implemented with a building energy simulation program and a parametric study of the impacts of different individual faults was conducted. The proposed approach was compared with optimal service schedules determined using dynamic programming, the results

showing the simplified approach reduced operating costs to within 2% of the optimal solution. Additionally, costs of performing periodic service and emergency service policies were compared with the near optimal policies. Significant potential operating costs savings were estimated for the use of AFDD and the service decision approach proposed.

Li and Braun developed a methodology to determine the economic impact of different faults affecting DX cooling equipment [14]. The overall economic performance degradation index (EPDI) developed in this work accounted for impacts of utility costs and equipment costs. Utility costs impacts were defined as additional costs for electricity incurred through the loss of equipment efficiency. Equipment costs impacts were defined as additional costs associated with replacement of systems due to increased equipment run-time that would be required to operate systems with faults. In the work, it was proposed that the EPDI could be used to recommend when it becomes economical to fix different faults.

In the literature, no previous work has been found that describes making service decisions when multiple faults affect a system simultaneously. In these cases, decisions become more complex since it requires estimations about how much each fault affects system performance and whether it is more economical to fix none, one, or multiple faults at once.

6.3 Optimal Service Policies

6.3.1 Optimal Service Decision Problem Formulation

The goal of service and maintenance optimization is to determine an optimal sequence of service decisions that minimizes life-cycle operating costs for a system or group of systems while maintaining constraints on occupant comfort, safety, or environmental impact. For direct-expansion cooling equipment, significant life-cycle costs include utility costs (cost to consume

energy, usually electricity), equipment replacement costs (due to premature failures or scheduled upgrades), and maintenance costs.

More explicitly, the goal of maintenance and service optimization is to determine an optimal sequence of service tasks (a) from the set of permissible service tasks (A) that minimizes the total operating costs (OC) of a system over a life-cycle $\tau_{life} = t_{life} - t_0$,

$$J_0 = \min_{a \in A} \left\{ \int_{t_0}^{t_l} OC(t)\,\mathrm{d}t \right\} \tag{6.1}$$

such that the temperature in the conditioned space (T_z) remains comfortable,

$$\tau_l \le T_z(t) - T_{z,sp}(t) \le \tau_u \quad \forall t \in [t_0, t_l] \tag{6.2}$$

where $T_{z,sp}$ is the space air temperature setpoint, and τ_l, τ_u are the minimum and maximum allowable deviations in space air temperature that maintain comfort. As previously stated, operating costs (OC) at some time t can be estimated as the sum of utility costs (UC), equipment costs (EC), and service costs (SC),

$$OC(t) = UC(t) + EC(t) + SC(t) \tag{6.3}$$

The components of total operating costs (UC, EC, and SC) have been discussed previously in Chapter 5.

A description of the underlying methodology for determining optimal solutions for sequential decision problems will be presented in this section. Classic texts of Denardo and Kirk present the underpinnings of the description presented here [74, 75]. The optimal solution is one way to measure the performance of any maintenance plan or policy. In other words, it is a useful benchmark that can be determined initially that provides some measure of how good or bad sub-optimal maintenance policies are. One simple benchmark that can be used to designate one extreme outcome is to determine the worst possible solution, i.e. the solution that results in the

most energy usage, requires the most resources, or has the largest total cost. Alternatively, the solution that results in the least energy usage, requires the least resources, or has the lowest total costs could be determined. This optimal solution, and the methodology used for its determination, is presented in this section.

Maintenance decision problems can be described as deterministic or as probabilistic. This classification stems predominantly from the type of model used to describe the timing and outcome of maintenance. In a deterministic model, the timing of maintenance and the results are known with certainty. A system that is subject to increasing operating costs as a function of the amount of equipment operation is a typical candidate for a deterministic model. When a probabilistic (or stochastic) model is used, the timing and result of maintenance depends on chance. A light bulb is a simple example of a stochastically failing system since its failure time is usually described using a probability distribution. Another complicating factor of probabilistic models is that the actual state of the system or equipment is not known with certainty. While probabilistic models offer interesting insights into maintenance decisions of different systems, the models used in this work will be deterministic in nature.

The goal of the maintenance decision problem is to determine when the best times to perform maintenance occur. This sequence of decisions is the target. Moreover, in the solution procedure, a set of possible solutions are generated that can be used to decide what decision should be made, regardless of the time or equipment condition. This is the maintenance policy that guarantees, under the assumptions of the underlying model, optimal operating costs.

Let the state at time t be x_t. For a decision that begins at $t = 0$, we take as given the initial state x_0. At any time, the set of possible actions depends on the current state; we can write this as $a_t \in A(x_t)$, where a_t represents one or more control variables. In the case of maintenance

decision-making, a_t represents any of the different service tasks we can perform on the air conditioner. We also assume the state changes from x to a new state $T(x,a)$ when action a is taken, and the current payoff (or cost) from taking action a in state x is $F(x,a)$. Finally, we may assume the value of money now is greater than in the future using a discount factor $0<\beta<1$.

Under these assumptions, an infinite-horizon decision problem takes the following form:

$$V(x_0)=\min_{\{a_t\}_{t=0}^{\infty}}\left\{\sum_{t=0}^{\infty}\beta^t F\left(x_t,a_t\right)\right\} \tag{6.4}$$

subject to the constraints

$$\begin{aligned} a_t &\in A\left(x_t\right) \\ x_{t+1} &= T\left(x_t,a_t\right) \\ t &= 0,1,2,\ldots \end{aligned}$$

Notice that we have defined notation $V\left(x_0\right)$ to denote the optimal value that can be obtained by minimizing (or maximizing) the objective function subject to the assumed constraints. This is the value function. It is a function of the initial state variable x_0, since the best value obtainable depends on the initial situation.

The dynamic programming method breaks a decision problem into smaller sub-problems. Bellman's Principle of Optimality describes how to do this:

> *Principle of Optimality: An optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision* [76].

As suggested by the Principle of Optimality, we will consider the first decision separately, setting aside all future decisions. Collecting the future decisions in brackets on the right, the previous problem is equivalent to

$$V(x_0) = \min_{a_0} \left\{ F(x_0, a_0) + \beta \left[\min_{\{a_t\}_{t=1}^{\infty}} \sum_{t=1}^{\infty} \beta^{t-1} F(x_t, a_t) \right] \right\} \tag{6.5}$$

subject to the initial constraints

$$a_0 \in A(x_0)$$
$$x_1 = T(x_0, a_0).$$

Here we are making one decision: choosing a_0, knowing that our choice will cause the state at $t=1$ to be $x_1 = T(x_0, a_0)$. That new state will then affect the decision problem from time $t=1$ on. So far it seems we have only made the problem more difficult by separately the initial decision from the future decisions. However, we can simplify the above equation by noticing that what is inside the square brackets on the right is the value of the decision problem starting from state $x_1 = T(x_0, a_0)$. Therefore, we can rewrite the problem as a recursive definition of the value function

$$V(x_0) = \min_{a_0} \left\{ F(x_0, a_0) + \beta \cdot V(x_1) \right\} \tag{6.6}$$

subject to the initial constraints

$$a_0 \in A(x_0)$$
$$x_1 = T(x_0, a_0).$$

This is known as the Bellman Equation [74, 75]. It can be simplified further if we drop the time subscripts and substitute in the value of the next state,

$$V(x) = \min_{a \in A(x)} \left\{ F(x, a) + \beta \cdot V\left[T(x, a) \right] \right\}. \tag{6.7}$$

The Bellman Equation is classified as a functional equation because it involves solving for an unknown function V, which is the value function. Recall that the value function describes the best possible value of the objective, as a function of the state x. By calculating the value

function, we will also find the function $a(x)$ that describes the optimal action as a function of the state – this is called the *policy function*.

Equation (6.7) describes how to calculate the value of following an optimal policy, however the sequence of actions must still be determined. One way to do this for deterministic problems is by a method call backwards-induction. In this methodology, the product of the set of possible states and set of possible actions at each state are determined for the next-to-last decision. Each of these combinations is analyzed and the action at each state that maximizes (or minimizes) the next-to-last value is assigned the optimal policy for each state. Following this, the procedure is repeated for the preceding decision stage. More explicitly, the product of the possible states and possible actions is determined and the action leading to the optimal ending value is recorded. The usefulness of this methodology is that since an optimal decision at the next-to-last decision stage has already been determined, no further optimization needs to be performed. In this way, the optimal policy function can be built piece-by-piece, i.e. for each decision stage back through time. Additionally, this procedure requires the least amount of computations since only feasible states and actions are analyzed at each step.

6.3.2 Demonstration Results for Optimal Maintenance Schedules

For different locations and fault rates, dynamic programming was used to determine the optimal maintenance policy function for the scenario tested. The policy functions determined for different locations and fault rates are compared for refrigerant charge faults, condenser fouling faults, and evaporator fouling faults. In these simulations (and all remaining throughout this chapter), the costs to perform different service tasks are listed in Table 6.1. When multiple service tasks are performed, a 20% discount on the total cost is applied. The electric utility rate

used throughout the simulations was \$0.15/kWh. This rate was used at all the building locations considered and no peak demand charges were used in the electrical cost calculation.

Table 6.1. Summary of service task costs used in equipment fault and maintenance simulations. When multiple tasks are performed, a 20% discount to the total cost is applied.

Task	Cost
Add Refrigerant Charge	\$100 + \$50/lb refrigerant
Clean Condenser Coil	\$300
Clean Evaporator / Filter	\$80

Service decisions were considered on weekly intervals for all simulation scenarios considered and presented in this chapter. The expected equipment life was also assumed to be fixed at 15 years for all locations and scenarios. Thus, for each simulation, there were $52 \times 15 = 780$ decision stages where service tasks could be performed or not. The fault levels considered at each decision stage were discretized to 1% intervals and linear approximation was used to determine the dynamic programming decision at each stage. No discounting was used in the model, $\beta = 1$.

6.3.2.1 Refrigerant Charge Faults

For a building in Miami using an air conditioner that leaks charge at a rate of 5% per year and has a maximum capacity that is 20% greater than the maximum load during the year, the optimal policy function is shown over the life of the equipment in Figure 6.1. The policy function can be used to determine the optimal service decision for any equipment state and time during the life of the system. As a demonstrative example, the chart in Figure 6.1 can be used as follows. First locate the vertex corresponding to the age of the system (on the horizontal axis) and the amount of refrigerant charge (on the vertical axis). If the vertex selected is within the blue shaded region, refrigerant should be added to the system since the cost of service can be

repaid by future operational cost savings. On the other hand, if the vertex is located outside the shaded region, service should not be performed.

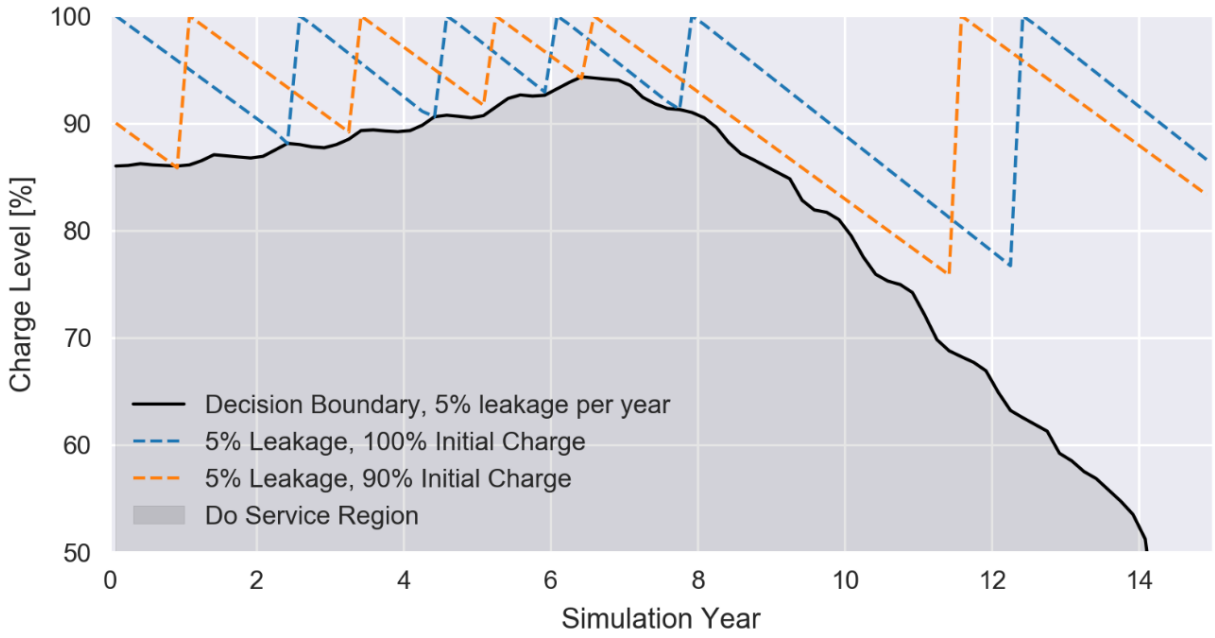

Figure 6.1. Optimal service decision policy for refrigerant charge faults in Miami, FL for system that is 20% oversized to maximum annual load ($f_{os} = 0.2$) and refrigerant leakage rate equal to 5% per year.

Also plotted on Figure 6.1 are the trended refrigerant charge levels for two systems with the same charge fault (5% leakage per year) but different initial charge levels (100% and 90%). The amount of refrigerant in the system decreases linearly with time, as described by the governing refrigerant charge leakage model previously in Chapter 5. When the amount of refrigerant in either system crosses the optimal service decision boundary, the optimal service policy would recommend adding refrigerant to the system. Applying this decision rule at any instance in time, the optimal lifetime operating costs under the assumptions of the simulation will be obtained.

It is interesting to compare the time-varying trends in refrigerant charge for each system shown in Figure 6.1. Because it started with lower refrigerant charge, the system with 90%

initial charge requires service before the system that starts with 100% charge. It is also noteworthy that each system is serviced exactly five time throughout the life of the simulation, regardless of the initial charge level. Furthermore, it can be shown that for any refrigerant charge level between 0% to 100%, a system will be service exactly five times throughout its life following the optimal service policy with 5% annually leakage. This result is an effect of the shape of the optimal decision boundary.

Early in the simulation lifetime, the optimal boundary between adding refrigerant to the system and not performing service is largely a tradeoff between integrated energy and equipment cost impact and the cost to perform additional service tasks. Performing service tasks more often would save on energy and equipment costs, at the expense of much higher lifetime service costs. On the other hand, performing service less frequently would saving on service costs by possibly performing less service tasks throughout the life, at the expense of much higher energy costs.

The decision boundary changes over time for two reasons. First, there is a small seasonality component to the decision which causes the small ripple in the decision boundary with a period of 12 months. Because refrigerant leaks throughout the year and there is less cooling load in Miami in the winter months, it is slightly better to wait until the warmer months to do service. The decision boundary for Miami reaches a peak around the 7^{th} year. This is an effect on optimizing the lifetime operating costs over a finite interval. After the 7^{th} year, the costs for performing service must be balanced by diminishing utility costs savings since there is no reward for finishing the simulation with an air conditioner with more charge (no salvage value). In other words, the optimal service policy tolerates more leakage since the possible future utility costs savings are less than earlier in the simulation.

Finally, it should be noted that each system shown in Figure 6.1 has charge added around the 11^{th} and 12^{th} years much before the systems' charge levels intersect with the decision boundary. These service decisions are the result of comfort violations – the significant reduction in refrigerant charge resulted in insufficient cooling capacity to maintain the space temperature of the building. Because there is a constraint to maintain comfort with the building, service must be performed.

The optimal decision boundary depends on the rate of refrigerant leakage during the simulation. Two optimal decision boundaries are shown for two systems located in Miami leaking refrigerant at different rates in Figure 6.2. It is observed that when the leakage rate is greater, more service is required over the life of the system. However, the decision boundaries between the 5% per year and 15% per year leakage rates are different. This is because the added costs required for additional service calls are expensive relative to the improvement in performance that can be gained by adding refrigerant. It is also observed from the results shown in Figure 6.2 that service is performed for a greater percentage of the equipment life for the larger leakage rate. This is because the higher refrigerant leakage rates lead to increases in utility costs per unit time and makes performing service later in the simulation more economical.

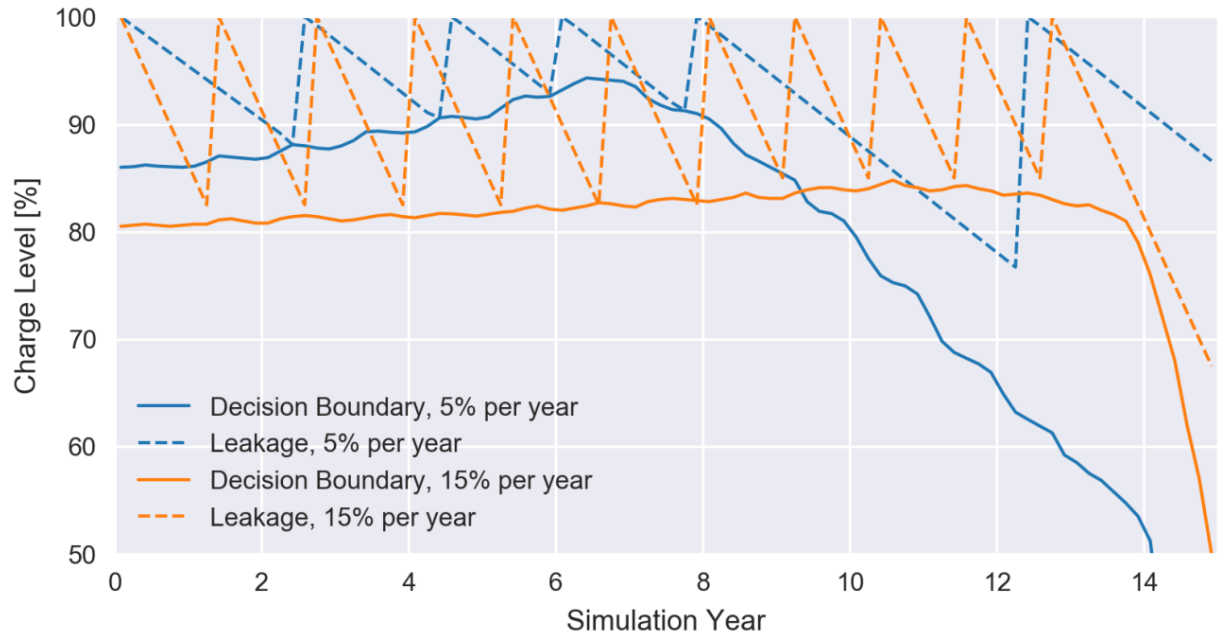

Figure 6.2 Comparison of optimal decision boundaries for refrigerant charge faults in Miami, FL for system that is 20% oversized to maximum annual load ($f_{os} = 0.2$) and fault leakage rate equal to 5% and 15% per year.

The same refrigerant charge fault (5% leakage per year) has been simulated for a building located in Atlanta, GA that has a similarly sized system installed (scaled for the maximum cooling load during the year). The optimal decision policy for this system is shown in Figure 6.3. In comparison to the Miami, FL optimal policy, a somewhat unexpected result can be observed. In this scenario, the decision boundary has peaks during the summer and valleys during the winter. This is due to the winter season, when the Atlanta building has a heating load, rather than a cooling load. In the Miami building, cooling was required all year round to some extent. This is because refrigerant leakage is a function of simulation time. If charge is added at the end of the fall season, the refrigerant would leak during the winter while never being used for cooling.

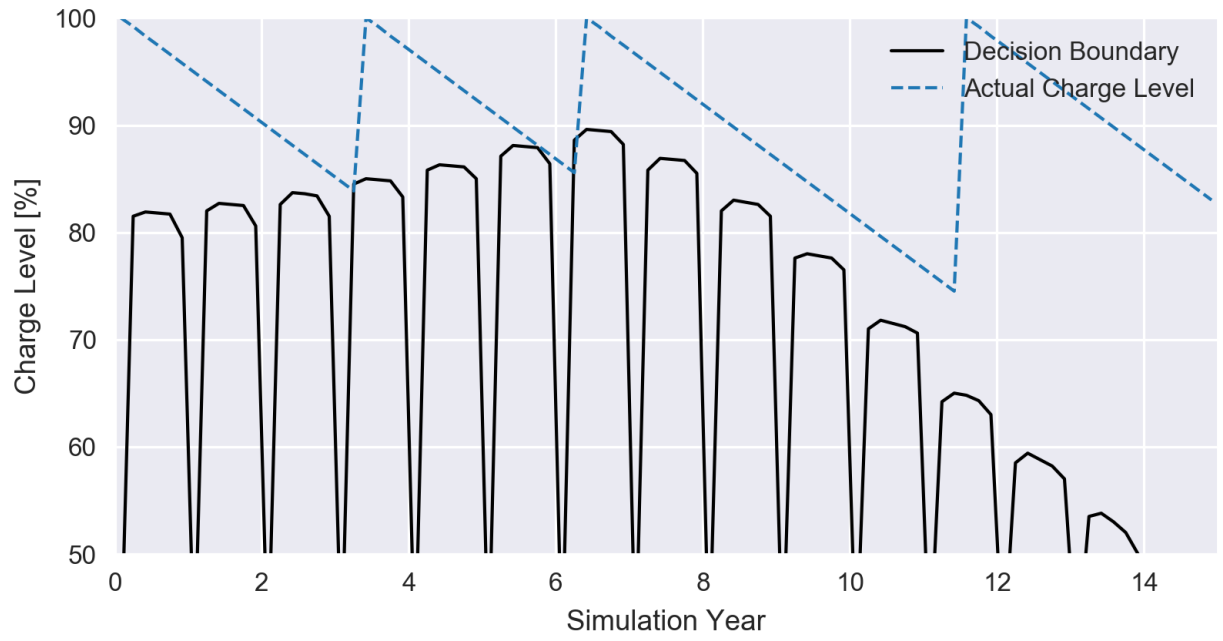

Figure 6.3. The optimal service schedule for Atlanta, GA when the system is assumed be initially fully charged. When service is performed, the charge level is returned to 100% (20% oversized to maximum annual load; fault leakage rate equal to 5% per year).

Assuming the system starts from a fully charged initial condition, the optimal refrigerant charging schedule is shown in Figure 6.3. From the trends shown, refrigerant is typically added during the early spring months, when it is most economical to do so. In comparison with the Miami, FL location, refrigerant charging service is required less often. This is because the cooling season and total cooling requirement in Atlanta is less than in Miami. Thus, refrigerant charging service is costlier in comparison to the utility costs savings possible. Another common trend can be observed between the Miami and Atlanta simulations: near the end of the simulation, service is not performed. Because there is no salvage value assumed in the underlying optimization model, there is no economic benefit in performing service near the end of the equipment life. If there was some monetary return depending on the condition of the equipment at the end of the simulation, there dynamic programming methodology would extend the decision boundary.

The same fault (5% leakage per year) has been simulated for a building located in Chicago, IL with a similarly sized air conditioning system (scaled for the maximum cooling load during the year). The optimal decision policy for this system is shown in Figure 6.4. In this scenario, the decision boundary has peaks during the summer and valleys during the winter. This is due to the winter season, when the Chicago building has a heating load, rather than a cooling load. In comparison to the Atlanta building, the Chicago building has wider valleys and narrow peaks due to the longer heating season and shorter cooling season.

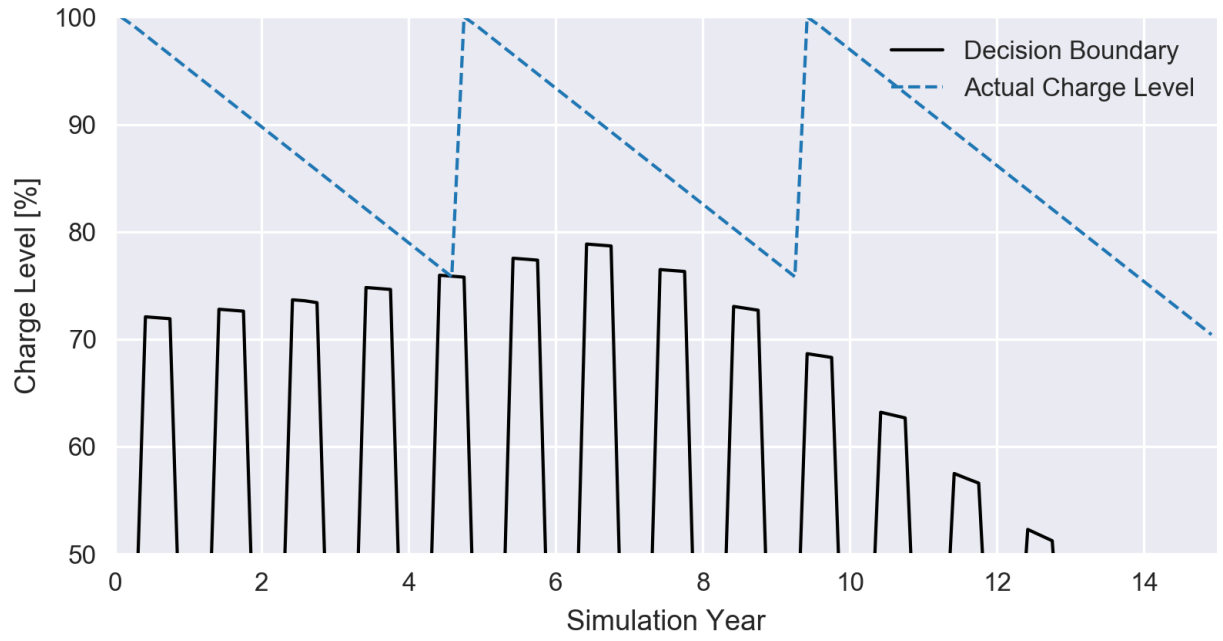

Figure 6.4. An example optimal service schedule for Chicago, GA when the system is assumed is initially fully charged. When service is performed, the charge level is returned to 100% (20% oversized to maximum annual load; fault leakage rate equal to 5% per year).

Assuming the system starts from a fully charged initial condition, the optimal refrigerant charging schedule is shown in Figure 6.4. From the trends shown, refrigerant is added to the system twice during its expected life. In comparison with the Miami, FL and Atlanta, GA locations, refrigerant charging service is required less often. This is because the cooling season

in Chicago is relatively short and a significant amount of leakage occurs during the winter months. Thus, refrigerant charging service is costlier in comparison to the utility costs savings that are possible and less effective since more charge leaks during the winter months when no cooling is required.

6.3.2.2 Condenser Fouling Faults

For a building in Miami using an air conditioner that has a reduction in condenser airflow at a rate of 5% per 5000 hours of run-time due to condenser fouling and has a maximum capacity that is 20% greater than the maximum load during the year, the optimal service scheduling policy function is shown over the life of the equipment in Figure 6.5. The policy function can be used to determine the optimal service decision for any equipment state and time during the life of the system.

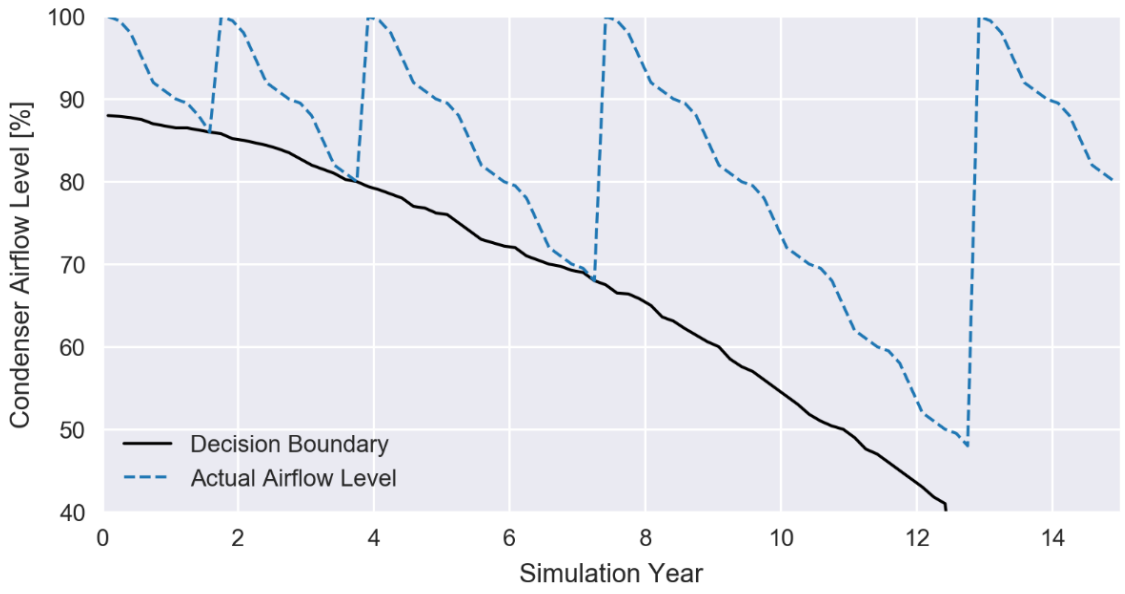

Figure 6.5. An optimal service schedule for Miami, FL when the system starts with a clean coil (no reduction in condenser airflow). When service is performed, the condenser airflow level is returned to 100%, however a service cost is incurred (20% oversized to maximum annual load; condenser fouling rate equal to 5% per 5000 hours of run-time).

Assuming the system starts with an initially clean condenser coil, the optimal condenser cleaning schedule is shown in Figure 6.5. From the trends shown, the condenser is cleaned multiple times throughout its life – however the interval becomes larger with each cleaning. This is because of end of life effects due to the finite interval the simulation is performed. Since at the end of the simulation there is no salvage value for the equipment based on its condition, performing service towards the end of the simulation doesn't not yield large benefits. This is also true because utility cost savings due to condenser coil cleaning see diminishing returns as the simulation approaches the end. It should be noted the last service decision shown in Figure 6.5 was not effect of the optimal policy, but rather a comfort violation. The event service performed between the 12^{th} and 14^{th} year was caused by high discharge pressure limits. When the compressor head pressure exceeds 550 psig, the cooling system is disabled causing a comfort violation until the condenser is cleaned.

The same condenser fouling fault has been simulated for a building located in Chicago, IL that has been similarly over-sized by 20% (scaled for the maximum cooling load during the year). The optimal service decision policy for the system is shown in Figure 6.6. Unlike the refrigerant charge leakage faults, there is less seasonal dependence on the condenser fouling service decision policy. This is because condenser fouling has been modeled as a linear function of equipment run-time. During the winter heating season since the equipment is not a heat pump, no equipment run-time is required, and condenser fouling does not increase. Additionally, more significant reductions in condenser airflow are required for service in comparison with the Miami, FL policy function. This is because Chicago has significantly less annual cooling load and condenser fouling service and requires more equipment utility cost savings to break even.

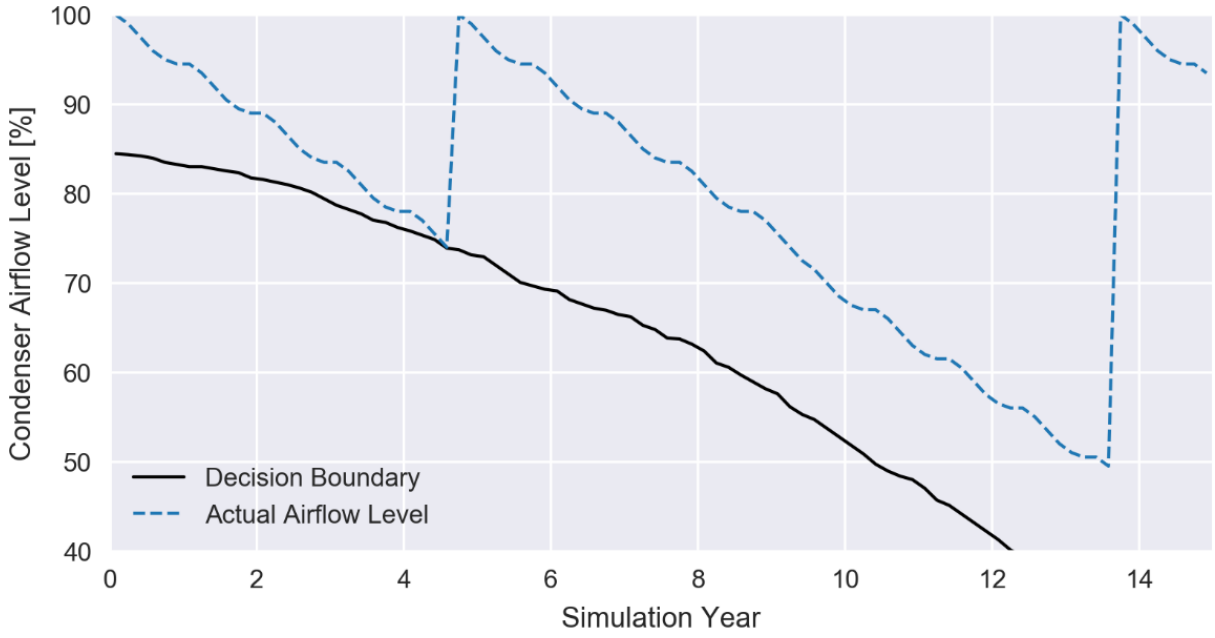

Figure 6.6. An optimal service schedule for Atlanta, GA when the system starts with a clean condenser coil (no reduction in condenser airflow). When service is performed, the condenser airflow level is returned to 100%, however a service cost is incurred (20% oversized to maximum annual load; condenser fouling rate equal to 5% per 5000 hours of run-time).

Assuming the system starts with an initially clean condenser coil, the optimal condenser cleaning schedule for the system installed in Chicago IL is shown in Figure 6.6. From the trends shown, the condenser is cleaned two times throughout its life. In comparison to the optimal schedule for the system installed in Miami that had the same condenser fouling rate, the optimal schedule for Chicago requires service less often. This is true for two reasons. First the annual cooling load in Chicago is much less than the annual cooling load in Miami. This makes performing condenser coil cleaning less economically effective since operating cost savings are smaller during the shorter cooling seasons. The second reason is also due to the shorter cooling season which reduces the annual equipment runtime requirement. Since there are less hours of equipment run-time in Chicago, it takes longer for the condenser coil to foul – decreasing the performance impacts.

6.3.2.3 Evaporator Fouling Faults

For a building in Miami using an air conditioner that has a reduction in evaporator airflow at a rate of 5% per 5000 hours of run-time due to evaporator fouling and has a maximum capacity that is 20% greater than the maximum load during the year, the optimal service scheduling policy function is shown over the life of the equipment in Figure 6.7. The policy function can be used to determine the optimal service decision for any equipment state and time during the life of the system. For evaporator fouling service, only small levels of evaporator fouling are tolerated by the optimal service schedule. This is largely due to the relatively small costs for evaporator fouling service in comparison to annual cooling costs for the building in Miami, FL.

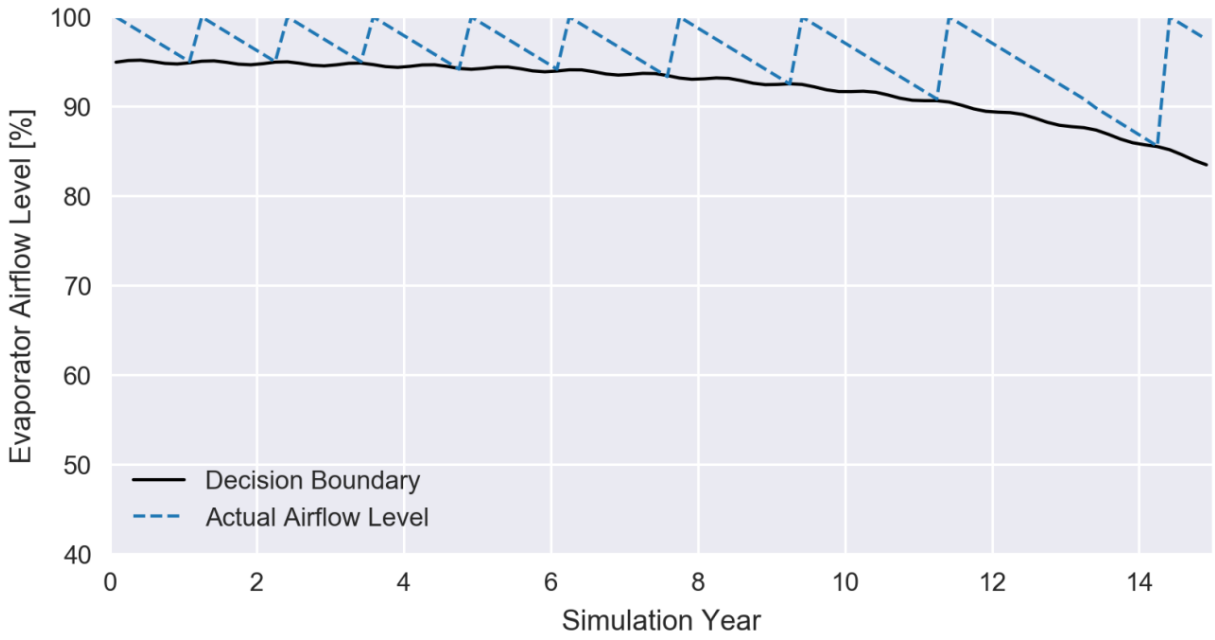

Figure 6.7. An optimal evaporator cleaning service schedule for Miami, FL when the system starts with a clean evaporator coil (no reduction in evaporator airflow). When service is performed, the evaporator airflow level is returned to 100%, however a service cost is incurred (20% oversized to maximum annual load; evaporator fouling rate equal to 5% per 5000 hours of run-time).

Assuming the system starts with an initially clean evaporator coil, the optimal evaporator cleaning schedule is shown in Figure 6.7. From the trends shown, the evaporator coil is cleaned approximately once per year throughout its life – however the interval between service times becomes larger with each cleaning. This is because of end of life effects due to the finite interval the simulation is performed. Since at the end of the simulation there is no salvage value for the equipment based on its condition, performing service towards the end of the simulation doesn't yield large benefits. This is also true because utility cost savings for evaporator coil cleaning see diminishing returns as the simulation approaches the final time step. In comparison to refrigerant charge service and condenser cleaning, these end of life impacts on the schedule are less important since the cost of evaporator filter replacement is significantly less than the other service tasks.

For comparison, the same evaporator fouling fault has been simulated for a building located in Chicago, IL that has been similarly over-sized by 20% (scaled for the maximum cooling load during the year). The optimal service decision policy for the system is shown in Figure 6.8. Like the refrigerant charge leakage faults, there is an obvious seasonal dependence on the evaporator fouling service decision policy. This is because evaporator fouling has been modeled as a linear function of indoor fan run-time. Due to ventilation requirements of the commercial building, this required the fan to operate whenever the building was occupied. During the winter heating season, the indoor fan must also run since it is traditionally used to blow air through the furnace for space heating. However, the utility cost penalty of this fault on operation of the vapor compression cooling equipment is only felt in the cooling season. In comparison to the system in Miami, the optimal service schedule permits slightly more evaporator fouling before service is required. This is because the operating costs calculated in the optimization problem only

consider performance impacts on cooling energy consumption. Since the annual cooling season is much shorter in Chicago than Miami, service can be performed less frequently.

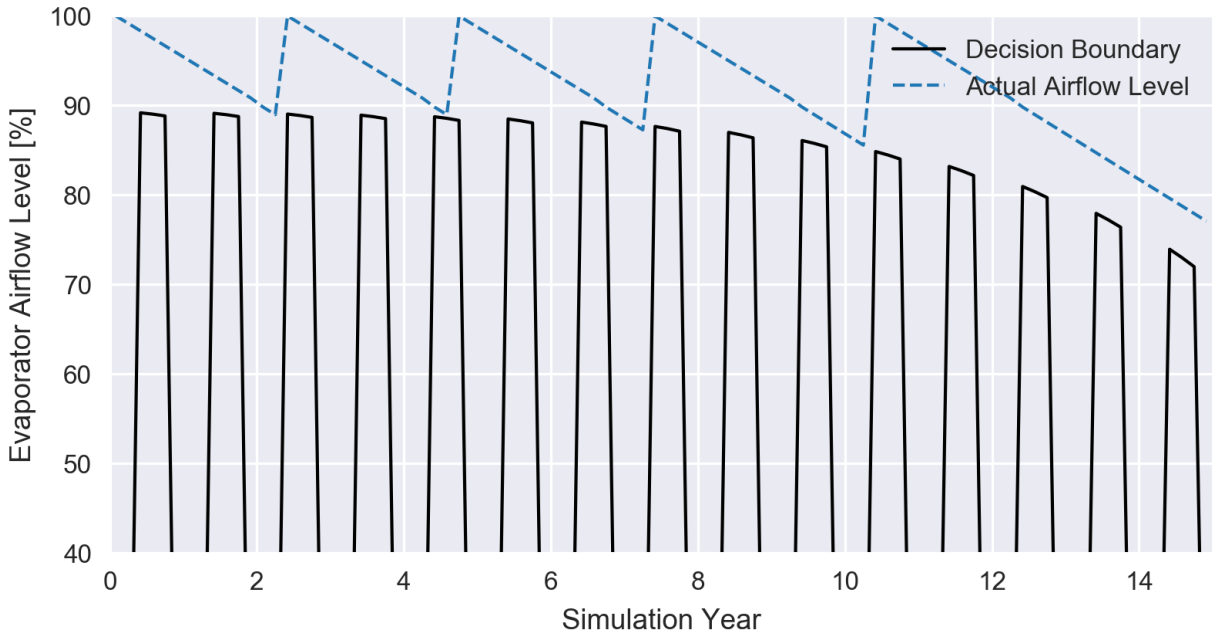

Figure 6.8. An optimal evaporator cleaning service schedule for Chicago, IL when the system starts with a clean evaporator coil (no reduction in evaporator airflow). When service is performed, the evaporator airflow level is returned to 100%, however a service cost is incurred.

Assuming the system starts with an initially clean evaporator coil, the optimal evaporator cleaning schedule for the system in Chicago is shown in Figure 6.8. From the trends shown, the evaporator foil is cleaned approximately once every two years throughout its life – however the interval between service times becomes larger with each cleaning. This is because of end of life effects due to the finite interval the simulation is performed. Since at the end of the simulation there is no salvage value for the equipment based on its condition, performing service towards the end of the simulation doesn't not yield large benefits. This is also true because utility cost savings because of evaporator coil cleaning see diminishing returns as the simulation approaches the final time step. In comparison to refrigerant charge service and condenser cleaning, these

end of life impacts on the schedule are less important since the cost of evaporator filter replacement is significantly less than the other service tasks.

6.3.2.4 Combinations of Faults

The optimal service schedule for a building located in Miami, FL that has multiple faults evolving over time was determined using dynamic programming. The fault rates for each fault are summarized in Table 6.2. The refrigerant charge in the system leaked 5% per year of simulation time. The condenser airflow rate was reduced by condenser coil fouling at a rate of 5% per 5000 hours of condenser fan runtime. The evaporator airflow rate was reduced by evaporator coil fouling at a rate of 5% per 5000 hours of evaporator fan runtime.

Table 6.2. Summary of fault evolution rate parameters used in first multiple fault simulation and service scheduling optimization.

Task	Cost
Refrigerant Charge Leakage Rate	5% per year
Condenser Fouling Rate	5% per 5000 hours condenser fan runtime
Evaporator Fouling Rate	5% per 5000 hours evaporator fan runtime

The optimal service schedule for the system in Miami, FL with the faults described in Table 6.2 is shown in Figure 6.9. Because the optimal decision boundaries for each service task also depend on the other fault levels, it is not possible to show the optimal policy decision threshold as was done previously. Instead, the trended fault levels are plotted in Figure 6.9. It can be observed that the optimal service schedule tends to group multiple service tasks into each service interval. This occurs because of the 20% discount applied to the service costs when multiple faults are serviced at the same time.

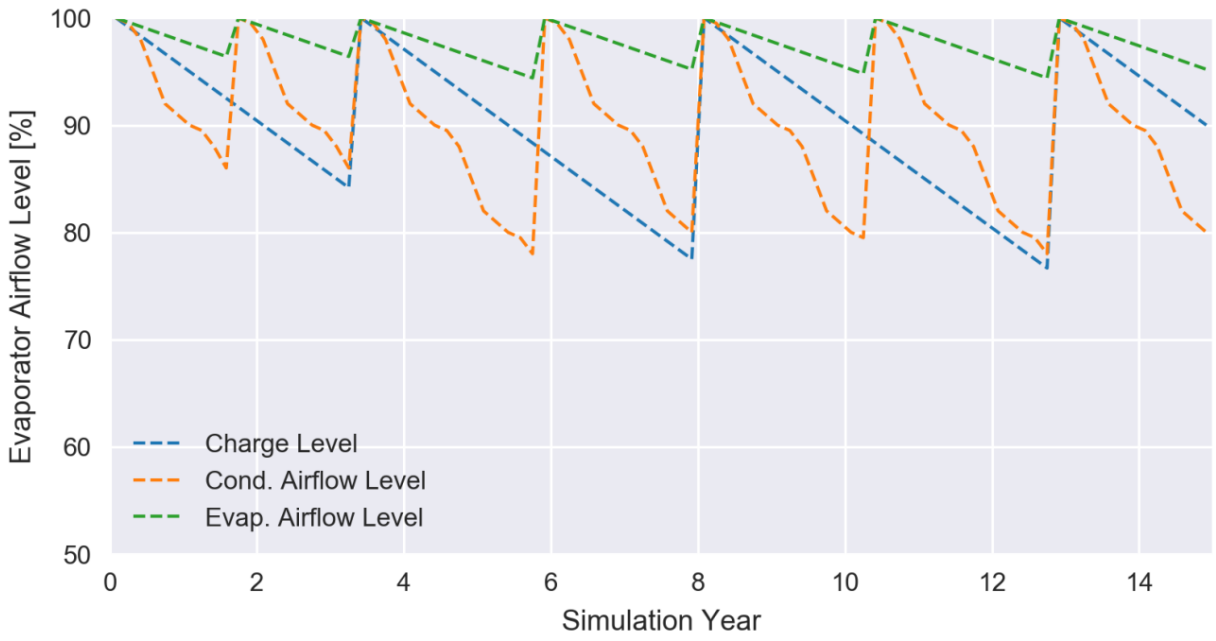

Figure 6.9. Optimal service schedule for system located in Miami, FL with multiple faults (described in Table 6.2). The optimal schedule tends to group multiple service tasks at each service interval.

For comparison, the same building located in Miami, FL was simulated with different fault rates. The fault rates for each fault used in the second scenario are summarized in Table 6.3. The refrigerant charge in the system leaked 5% per year of simulation time (the same rate as the previous scenario). The condenser airflow rate was reduced by condenser coil fouling at a rate of 2% per 5000 hours of condenser fan runtime. The evaporator airflow rate was reduced by evaporator coil fouling at a rate of 10% per 5000 hours of evaporator fan runtime.

Table 6.3. Summary of fault evolution rate parameters used in second multiple fault simulation and service scheduling optimization.

Task	**Cost**
Refrigerant Charge Leakage Rate	5% per year
Condenser Fouling Rate	2% per 5000 hours condenser fan runtime
Evaporator Fouling Rate	10% per 5000 hours evaporator fan runtime

The optimal service schedule for the second scenario in Miami, FL with the faults described in Table 6.3 is shown in Figure 6.10. Like in the previous scenario, the optimal service schedule tends to group multiple service tasks into each service interval. Moreover, evaporator fouling faults are serviced at each service interval. Because the evaporator fouling service task is the least expensive and the cost is relatively small compared to annual cooling energy costs, performance benefits from evaporator cleaning can overcome any service cost relatively quickly. At the end of the simulation, the refrigerant charge and condenser fouling decreased significantly since service was not performed.

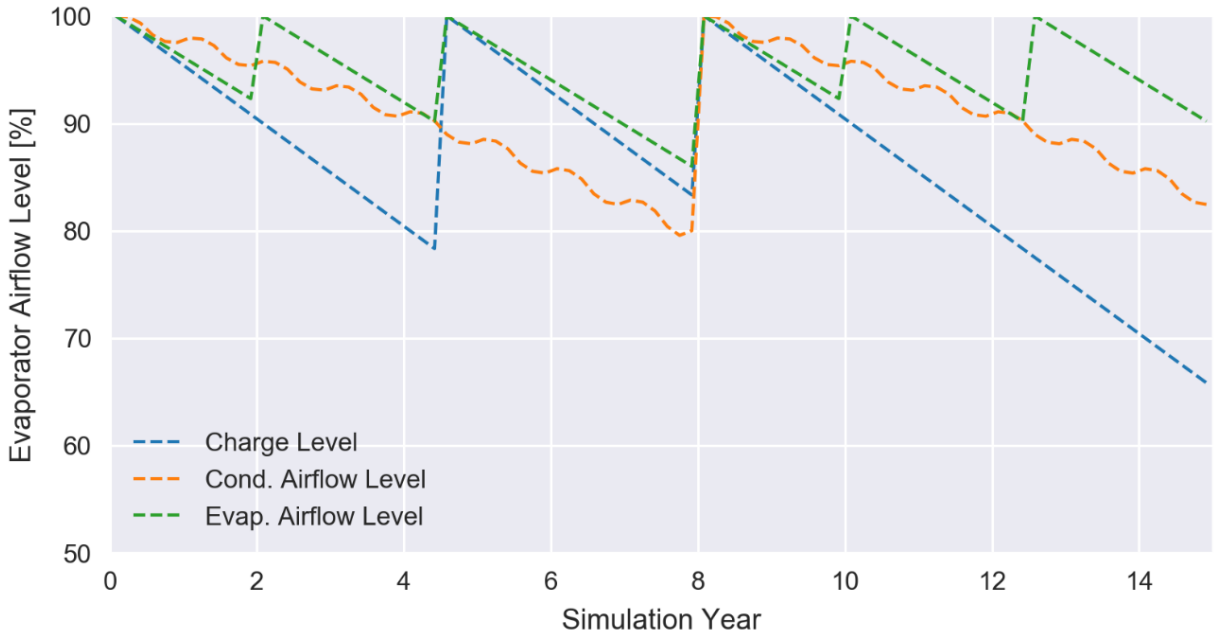

Figure 6.10. Optimal service schedule for system located in Miami, FL with multiple faults (described in Table 6.3). The optimal schedule tends to group multiple service tasks at each service interval.

6.4 Periodic Service Policies

6.4.1 Description and Implementation of Periodic Service Policies

Implementing periodic service policies is relatively straightforward and can easily be compared with optimal service policies using the simulation framework. In these policies, a service technician is assumed to visit the air conditioning system at regular intervals and perform any preventative maintenance that is needed. In many commercial buildings, service contracts between the building owner and HVAC service providers are often implemented which approximate periodic service policies. In these contracts, the service provider generally agrees to visit the site a fixed number of times per year and perform a variety of preventative maintenance tasks in return for some fixed costs paid by the building owner. While service may not be performed at exact intervals (i.e. every six months), service time between service visits is approximately constant (i.e. annually, biannually, or quarterly).

In this work, an additional assumption about how periodic service is performed may not be exactly true in a real application. The periodic service contract that is implemented within the simulation requires each service task to be completed at every visit. In other words, evaporator cleaning, condenser cleaning, and refrigerant charge adjustment is performed whenever the service technician visits the site if the fault levels are not normal (i.e. service to repair a fault is not performed if it is not considered in the simulation). In a real scenario, the service technician may not perform all tasks during every visit. Rather, the technician may only inspect the system to determine if maintenance is needed based on experience. If these inspections are permitted within the service contract, rather than requiring that each task is performed per visit, it may decrease the service costs.

Inspection policies were not investigated or implemented in this work since it requires some assumptions about how the service technician perceives the equipment state and when service is

needed. In real scenarios, performing on-site inspections is generally a good idea since it may provide valuable insights into how a system is performing. For instance, a service technician may clean the condenser coil if they notice it is covered in debris. On the other hand, minor condenser fouling would be ignored if the condenser looks mostly clean. In the simulation, these considerations are not modeled. Rather, the service technician will perform different service tasks regardless of the severity of the faults.

6.4.2 Demonstration Results of Periodic Service Policies

Using the periodic service policy implementation, the simulation framework was used to compare the operating costs of periodic service schedules with optimal service schedules. These comparisons were performed for the different buildings considered in this work with equipment that was oversized by 20% for the maximum cooling load over the year in each location. Different fault rates were also considered to determine the sensitivity of fault rates on annual operating costs using periodic service schedules.

A comparison between the trended refrigerant charge levels for a system that is serviced annually and biennially leaks 5% of its refrigerant charge annually for a system installed in Miami, FL is shown in Figure 6.11. In comparison to the optimal decision boundary, Figure 6.11 shows that the annual and biennial service schedules perform service too often for a 5% leakage fault. Using these schedules, annual utility costs are less than the utility costs obtained using the optimal schedule. However, service costs are much greater since service is performed more often over the life of the equipment. Figure 6.11 also shows that periodic service schedules do not consider the payback time required to breakeven towards the end of equipment life. This accounts for much higher lifetime service costs.

Figure 6.11. Comparison of lifetime refrigerant charge levels for system located in Miami, FL with a leakage rate of 5% percent per year using biennial and annual periodic service schedules.

Integrating the operating costs over the simulation time, the average annual operating costs of the two different periodic service schedules for a system installed in Miami, FL are shown in Figure 6.12 for different refrigerant leakage rates. Compared to the optimal service schedules, the periodic service schedules are often at least 10% greater than the optimal average annual costs. The annual and biennial intersect when the annual refrigerant leakage rate is equal to 23%. When this occurs, the additional cost of performing service annually is equal to the additional utility costs when the performing service biennially.

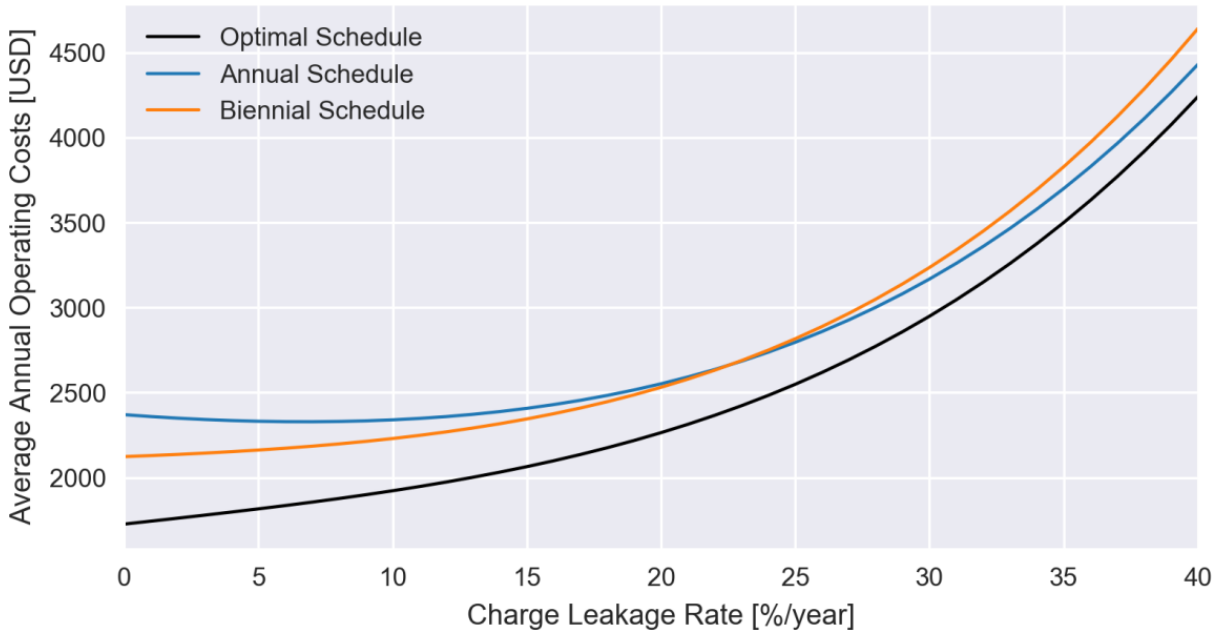

Figure 6.12. Comparison of average annual operating costs for system located in Miami, FL with different refrigerant charge leakage rates using biennial and annual periodic service schedules.

The additional annual operating costs relative to the optimal schedules for the two periodic service policies at a building located in Miami, FL are shown for different leakage rates in Figure 6.13. As the leakage rates become greater, the periodic service intervals tend to approach the optimal service schedules. This is because the service interval approaches the service optimal service interval. This is also partly due to the end-of-life effects becoming less significant since the payback time required to breakeven on service costs becomes shorter.

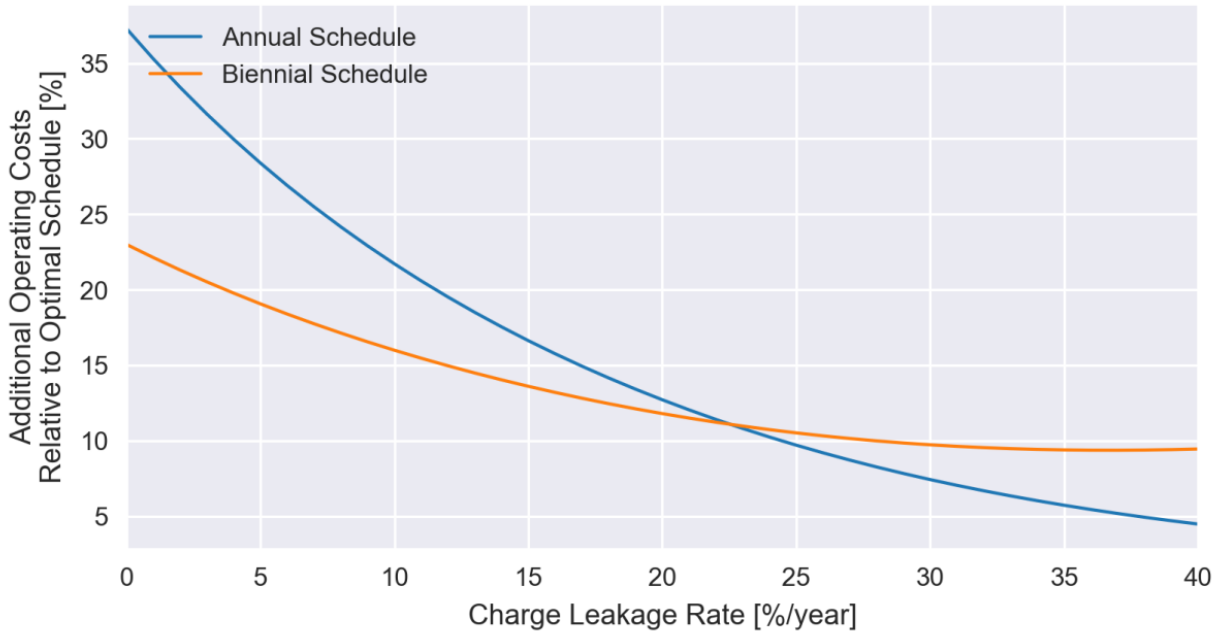

Figure 6.13. Additional lifetime operating costs comparison between annual and biannual service schedules for building located in Miami, FL with different refrigerant leakage rates.

A comparison between the trended refrigerant charge levels for a system that is serviced annually and biennially and that leaks 5% of its refrigerant charge annually for a system installed in Chicago, IL is shown in Figure 6.14. In comparison to the optimal decision boundary, Figure 6.14 shows that the annual and biennial service schedules perform service too often for a 5% leakage fault. In comparison to much greater annual cooling loads in Miami, FL, the periodic service intervals occur significantly more frequently than would be optimal. Using these schedules, annual utility costs are less than the utility costs obtained using the optimal schedule. However, service costs are much greater since service is performed more often over the life of the equipment. Figure 6.14 also shows that periodic service schedules do not consider the payback time required to breakeven towards the end of equipment life. This accounts for much higher lifetime service costs.

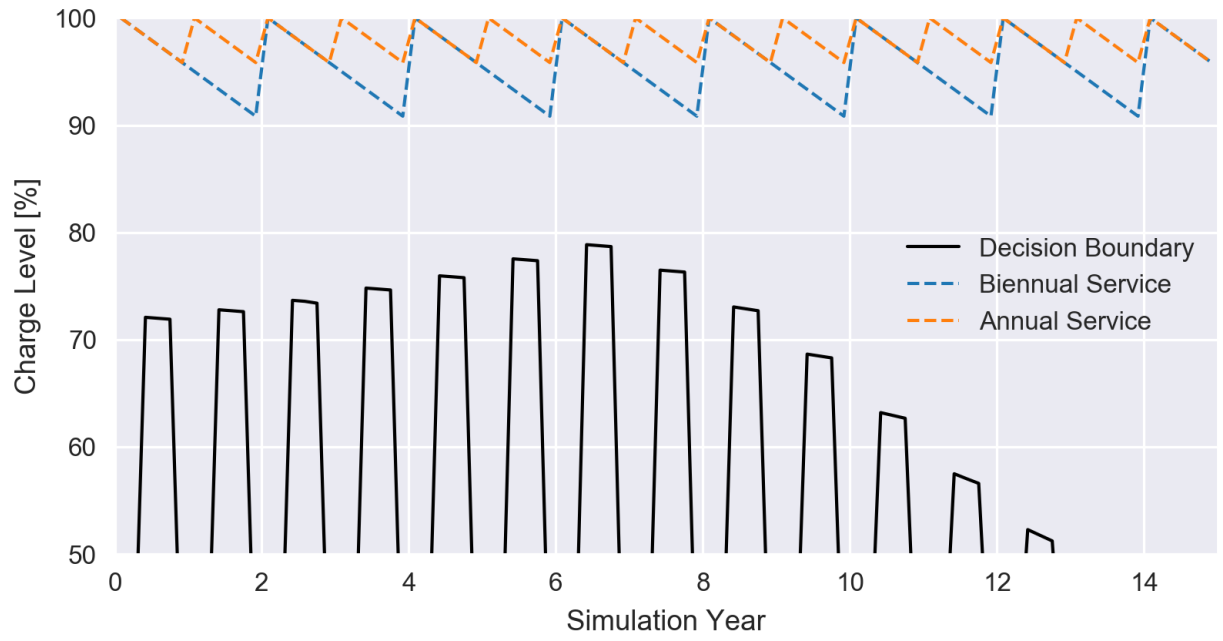

Figure 6.14. Comparison of lifetime refrigerant charge levels for system located in Chicago, IL with a leakage rate of 5% per year using biennial and annual periodic service schedules.

The operating costs were integrated over the simulation time to determine the average annual operating costs of the two different periodic service schedules. These average costs over a range of refrigerant leakage rates are shown for the building system located in Chicago, IL in Figure 6.15. Because annual energy costs in Chicago are much lower than in Miami, the impact of annual and biennial service schedules is relatively greater.

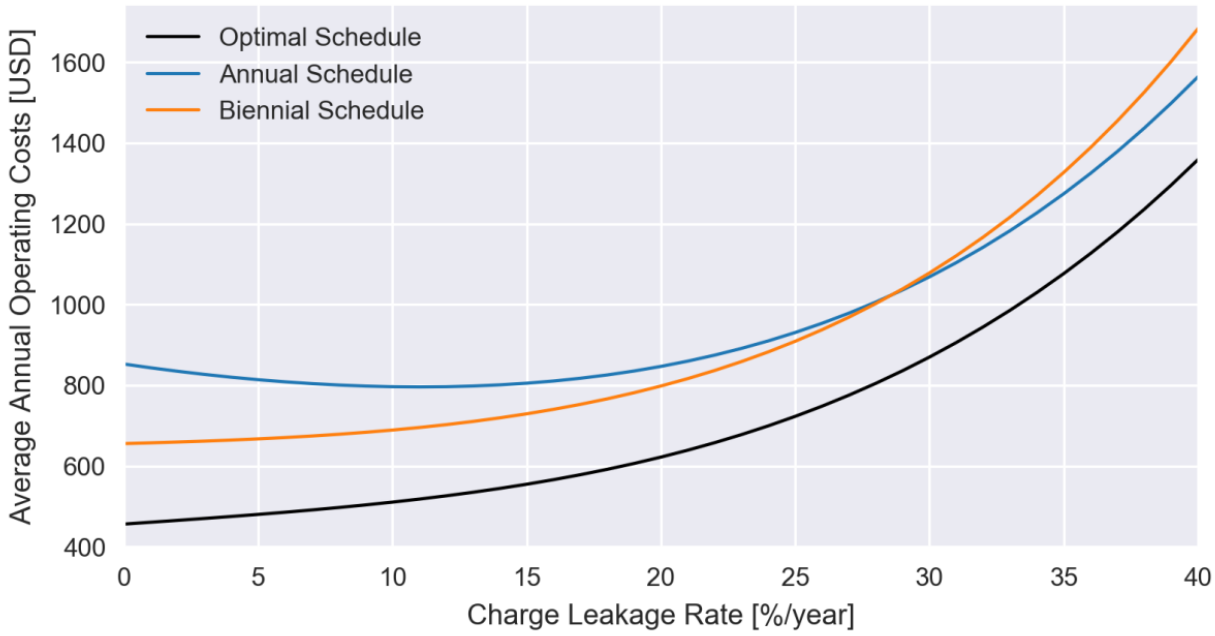

Figure 6.15. Lifetime average annual operating cost comparison between annual and biannual service schedules for building located in Chicago, IL with different refrigerant charge leakage rates.

The additional annual operating costs relative to the optimal schedules for the two periodic service policies at a building located in Chicago, IL are shown for different leakage rates in Figure 6.16. As the leakage rates become greater, the periodic service intervals tend to approach the optimal service schedules. This is because the service interval approaches the service optimal service interval. This is also partly due to the end-of-life effects becoming less significant since the payback time required to breakeven on service costs becomes shorter. In comparison to the Miami, FL results, annual and biennial periodic service schedules are costlier relative to the optimal service schedule for the Chicago application. This is because the annual cooling loads are much less (25% of the cooling loads in Miami). For constant service costs between the two locations, this makes service costs more expensive relative to the utility costs in each location.

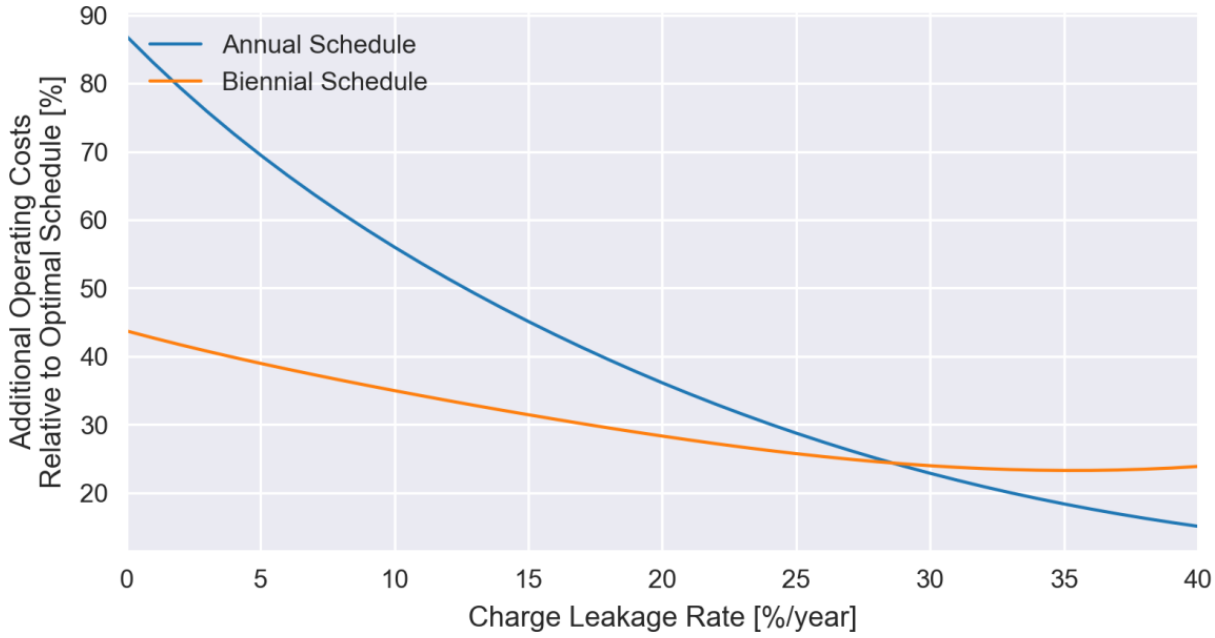

Figure 6.16. Additional lifetime operating costs comparison between annual and biannual service schedules for building located in Chicago, IL with different refrigerant leakage rates.

A comparison between the trended condenser airflow levels for a system installed in Miami, FL that is serviced annually and biennially is shown in Figure 6.17. To generate this comparison, a condenser fouling rate of 5% per 5000 hours of condenser fan runtime was simulated. In comparison to the optimal decision boundary, Figure 6.11 shows that the annual and biennial service schedules perform service too often. Using these schedules, annual utility costs are less than the utility costs obtained using the optimal schedule. However, service costs are much greater since service is performed more often over the life of the equipment. Figure 6.17 also shows that periodic service schedules do not consider the payback time required to breakeven towards the end of equipment life. This accounts for much higher lifetime service costs.

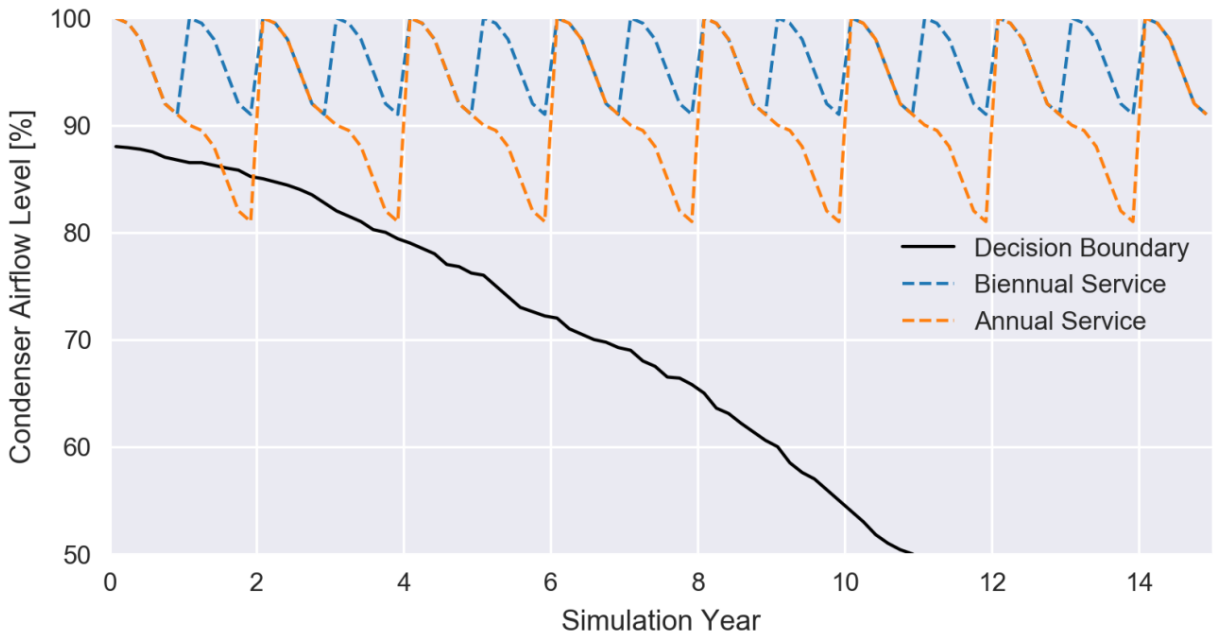

Figure 6.17. Comparison of lifetime condenser fouling levels for system located in Miami, FL with a condenser fouling rate of rate of 5% percent per 5000 hours of condenser fan runtime using biennial and annual periodic service schedules.

Integrating the operating costs over the simulation time, the average annual operating costs of the two different periodic service schedules for the system installed in Miami, FL are shown in Figure 6.18 over a range of condenser fouling rates (per 8760 hours of condenser fan runtime). Compared to the optimal service schedules, the periodic service schedules are always greater than the optimal average annual costs. The annual and biennial schedules intersect when the annual condenser fouling rate is equal to 22% per year of condenser run-time. When this occurs, the additional cost of performing service annually is equal to the additional utility costs when performing service biennially.

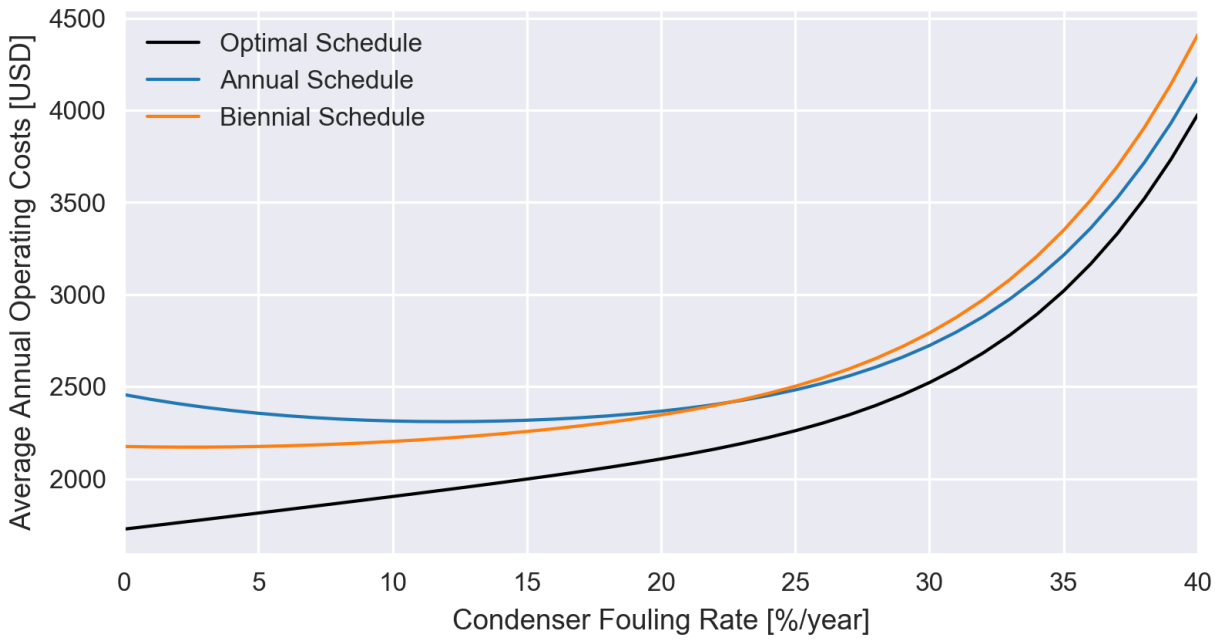

Figure 6.18. Comparison of average annual operating costs for system located in Miami, FL with different condenser fouling rates using biennial and annual periodic service schedules.

The additional annual operating costs relative to the optimal schedules for the two periodic service policies at a building located in Miami, FL are shown for different condenser fouling rates in Figure 6.19. As the condenser fouling rates become greater, the periodic service intervals tend to approach the optimal service schedules. This is because the service interval approaches the service optimal service interval - for faster condenser fouling, the additional energy consumed makes more frequent condenser cleaning economical. This is also partly due to the end-of-life effects becoming less significant since the payback time required to breakeven on service costs becomes shorter.

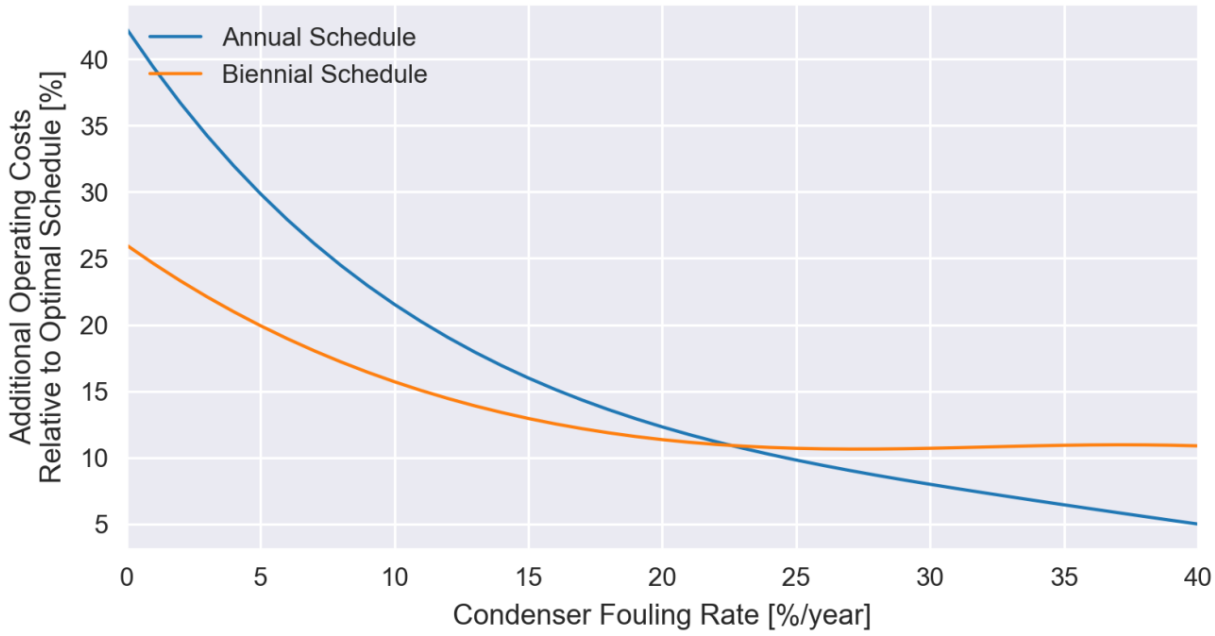

Figure 6.19. Additional lifetime operating costs comparison between annual and biannual service schedules for building located in Miami, FL with different condenser fouling rates.

To understand the sensitivity to periodic service costs for different annual cooling loads, the same condenser fouling faults were simulated for a building in Chicago, IL. The trended condenser airflow levels for the system that is serviced annually and biennially are shown in Figure 6.20. In comparison to the optimal decision boundary, Figure 6.20 shows that the annual and biennial service schedules perform service too often. In comparison to much greater annual cooling loads in Miami, FL, the periodic service intervals occur significantly more frequently than would be optimal. Using these schedules, annual utility costs are less than the utility costs obtained using the optimal schedule. However, service costs are much greater since service is performed more often over the life of the equipment. Figure 6.20 also shows that periodic service schedules do not consider the payback time required to breakeven towards the end of equipment life. This accounts for much higher lifetime service costs.

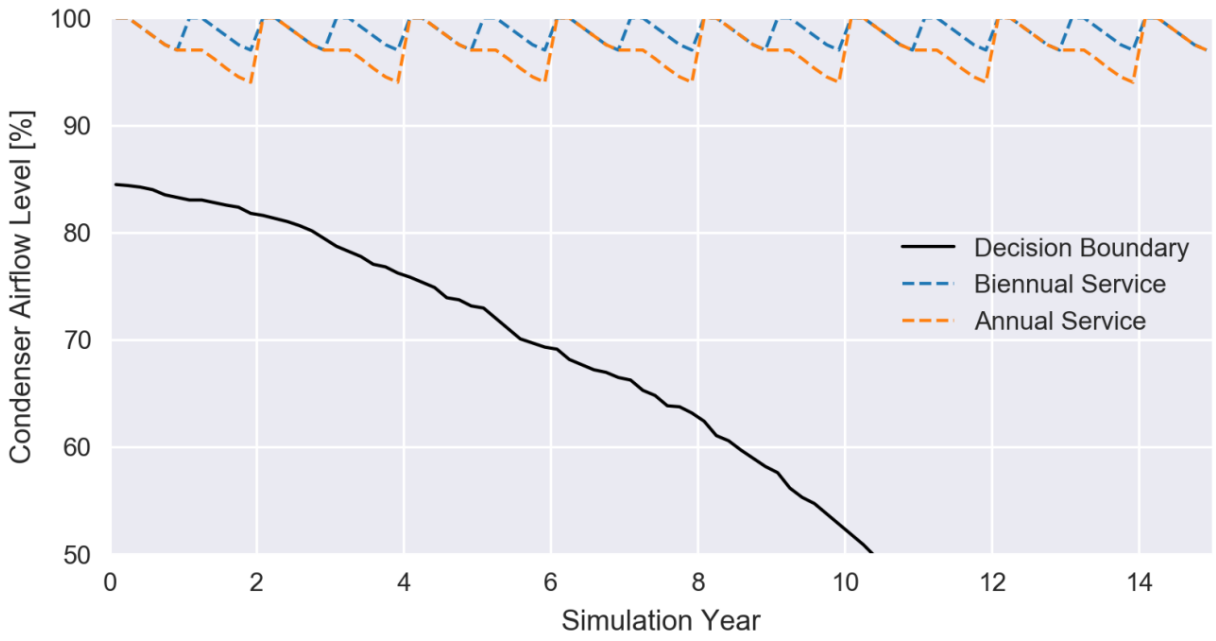

Figure 6.20. Comparison of lifetime condenser airflow levels for system located in Chicago, IL with a condenser fouling rate of 5% 5000 hours of condenser fan runtime using biennial and annual periodic service schedules.

The operating costs were integrated over the simulation time to determine the average annual operating costs of the two different periodic service schedules. These average costs over a range of condenser fouling rates are shown for the building system located in Chicago, IL in Figure 6.21. Because annual energy costs in Chicago are much lower than in Miami, the impact of annual and biennial service schedules is relatively greater.

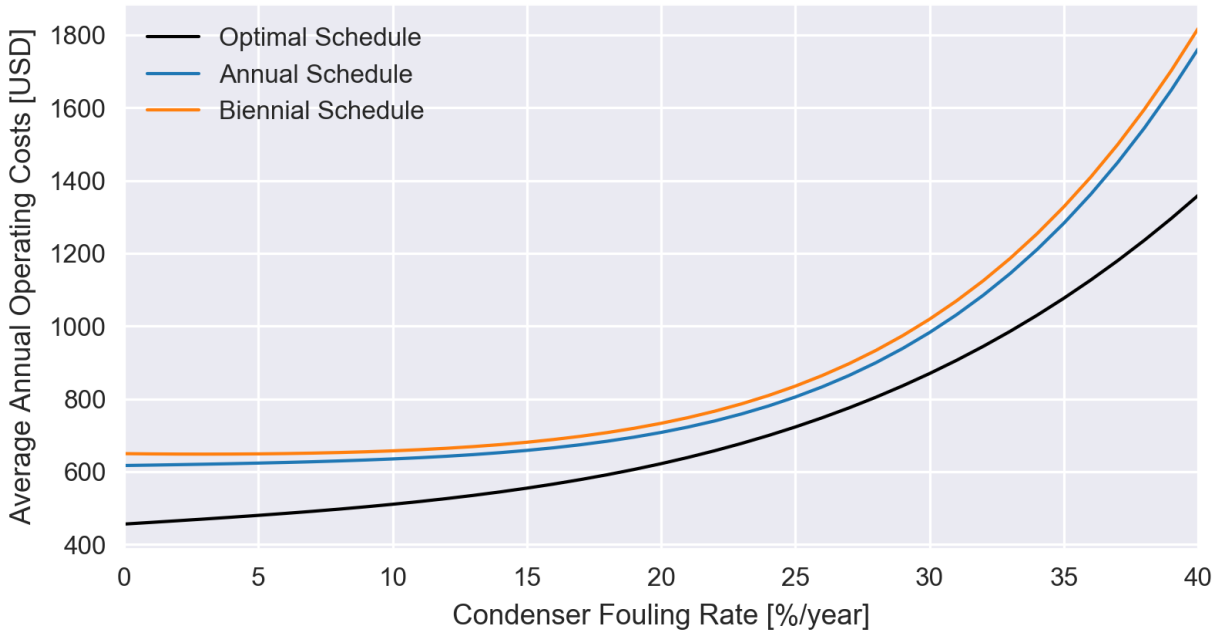

Figure 6.21. Lifetime average annual operating cost comparison between annual and biannual service schedules for building located in Chicago, IL with different condenser fouling rates.

The additional annual operating costs relative to the optimal schedules for the two periodic service policies at a building located in Chicago, IL are shown for different condenser fouling rates in Figure 6.22. In comparison to the Miami, FL results, annual and biennial periodic service schedules are costlier relative to the optimal service schedule for the Chicago application. This is because the annual cooling loads are much less (25% of the cooling loads in Miami). For constant service costs between the two locations, this makes service costs more expensive relative to the utility costs in each location. For condenser fouling service, it is also observed that the annual and biennial additional operating costs do not intersect. This indicates the additional energy savings that might be obtained by cleaning the condenser coil twice per year does not make up for the additional service cost.

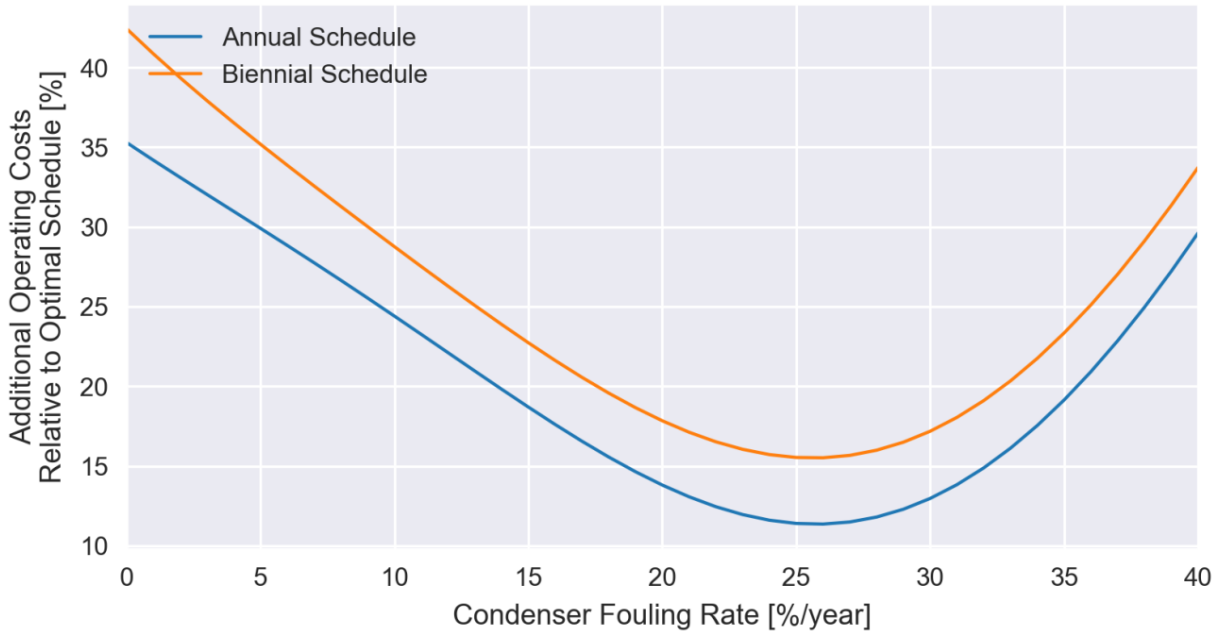

Figure 6.22. Additional lifetime operating costs comparison between annual and biannual service schedules for building located in Chicago, IL with different condenser fouling rates.

6.5 Emergency Service Policies

6.5.1 Description and Implementation of Emergency Service Policies

Whereas periodic service policies can be viewed as proactive, emergency service policies can be considered reactive. In this policy, service is performed only when a comfort violation occurs due to insufficient cooling capacity provided by the air conditioning equipment. In this work, comfort is violated when the temperature in the space exceeds the setpoint by 1.1 °C (2.0 °F) for a continuous interval of 4 hours or longer. When comfort is violated, service is performed immediately – evaporators are cleaned, condensers are cleaned, and refrigerant charge is adjusted. In real situations, a time lag between when comfort is violated and when service is performed may be significant. This is especially true depending on the time of year service is needed: service technicians may be very busy during peak cooling months, while they may be more

available during the shoulder seasons. The seasonal availability and costs are not considered in this work.

Emergency service policies may contribute to large operating costs for oversized systems since capacity violations may not occur until faults have degraded performance significantly. Additionally, emergency policies may incur higher operating costs than periodic service policies for faults that have limited impacts on cooling capacity, e.g. condenser fouling. When condenser fouling occurs, the significant impacts are increased head pressure and energy consumption. Comfort violations caused by condenser fouling may never occur and can lead to significant time between service intervals. One measure that has been implemented within the simulation that triggers comfort violations because of condenser fouling is a high-pressure limit. In normal systems, high pressure limit switches are typically installed to protect the compressor from operating outside the manufacturer's suggested operating envelope. When the head pressure exceeds the high-pressure limit, the air conditioner is disabled until service is performed.

6.5.2 Demonstration Results of Emergency Service Policies

The trended refrigerant charge level for a system located in Miami, FL that is serviced when comfort is violated is shown in Figure 6.23. In comparison to the optimal decision boundary, the emergency service policy allows the refrigerant in the system to go well below the optimal decision boundary. This results in higher rates of utility consumption due to degraded cooling capacity and system efficiency. Service is required when the cooling capacity is degraded enough to cause comfort violations. For this scenario, this occurs when the capacity degrades by approximately 20% (though can be more during the off-peak seasons). Because service is performed only when needed, lifetime service costs are generally considerable lower than policies that require service annually or biennially.

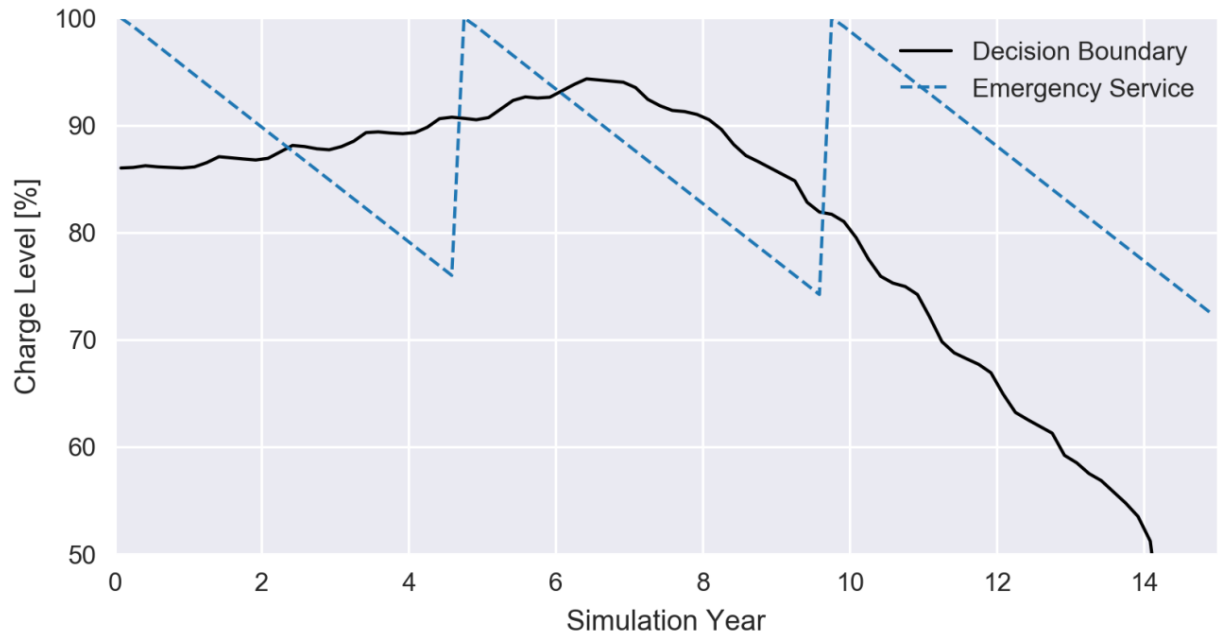

Figure 6.23. Trended refrigerant charge for an air conditioner with refrigerant leaking at a rate of 5% per year located in Miami, FL over life of equipment when an emergency service policy is used. Service is performed with insufficient capacity causes comfort violations.

The operating costs were integrated over the simulation time to determine the average annual operating costs for the emergency service policy. This was repeated for different refrigerant leakage rates and the results are shown in Figure 6.24. The results show that emergency service policies for refrigerant charge leakage tend to have higher operating costs. Additionally, there is much more variability in the trend when compared with the periodic maintenance policies. This is because of the somewhat random timings of comfort violations. Additionally, it should be noted that emergency policies approach the optimal service costs when refrigerant leakage rates become small. This is expected since a system without faults would never require service.

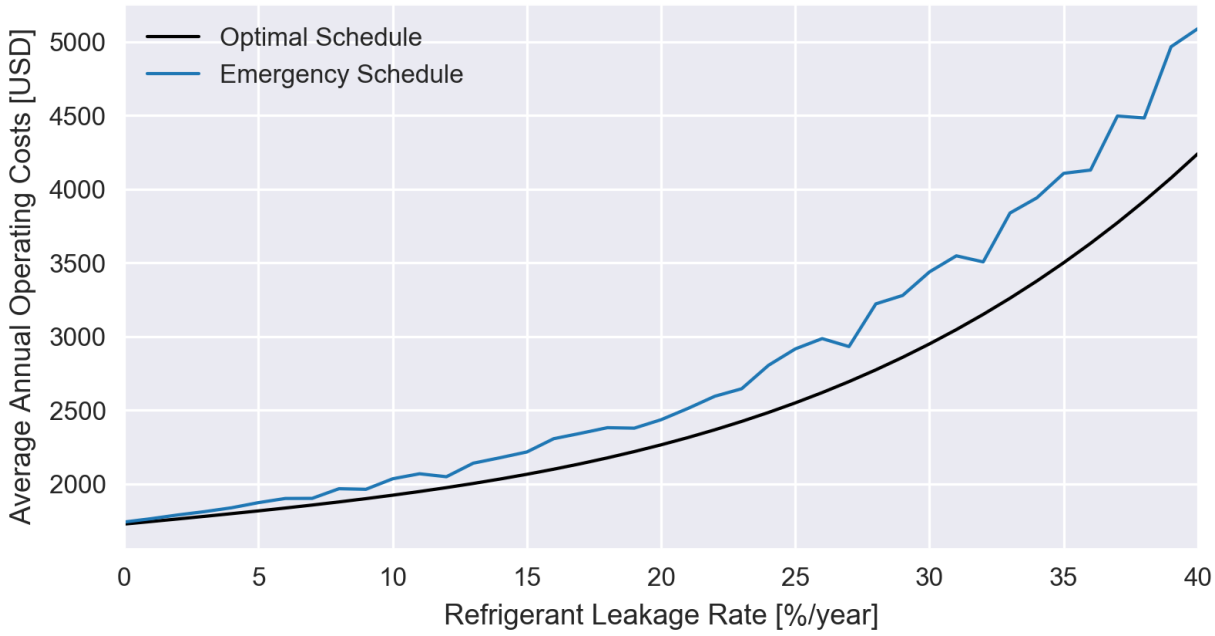

Figure 6.24. Comparison of average annual service costs for system in Miami, FL with different refrigerant charge leakage rates using optimal and emergency service policies.

For comparison, the same fault was simulated for a system located in Chicago, IL. The trended refrigerant charge level of the system that is serviced when comfort is violated is shown in Figure 6.25. In comparison to the optimal decision boundary, the emergency service policy allows the refrigerant in the system to go below optimal decision boundary, but not as much compared to the decision boundary for the Miami location. This will result in slightly higher rates of electric energy consumption due to degraded cooling capacity and system efficiency. Service is performed when the cooling capacity is degraded enough to cause comfort violations. Since the cooling loads are much less than the loads in Miami, the service costs required for the Chicago location are less.

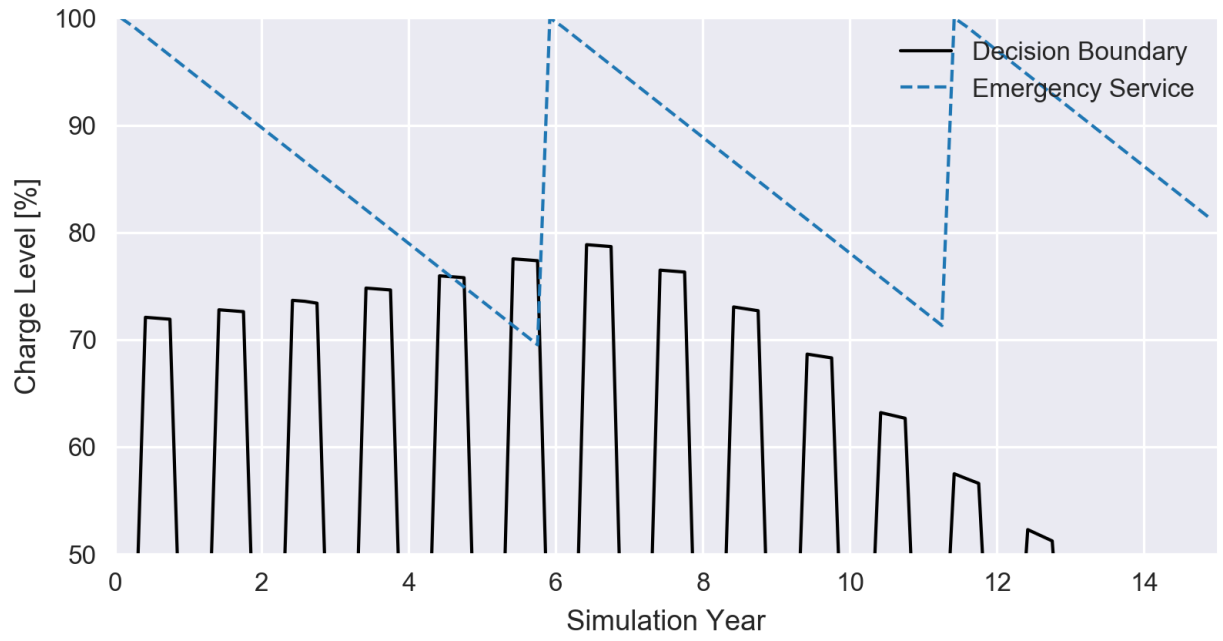

Figure 6.25. Trended refrigerant charge for an air conditioner with refrigerant leaking at a rate of 5% per year located in Chicago, IL over life of equipment when an emergency service policy is used. Service is performed with insufficient capacity causes comfort violations.

The operating costs were integrated over the simulation time to determine the average annual operating costs for the emergency service policy. This was repeated for different refrigerant leakage rates and the results are shown in Figure 6.26. The results show that emergency service policies for refrigerant charge leakage tend to have higher operating costs. Compared to the results obtained for the system in Miami, there is less variation in the average annual operating costs for the system in Chicago. This is because the annual cooling costs in Chicago are relatively small in comparison to the service costs. Thus, the increases in annual operating costs is largely dependent on the number of time the equipment is serviced throughout it life, which is strongly dependent on the leakage rate.

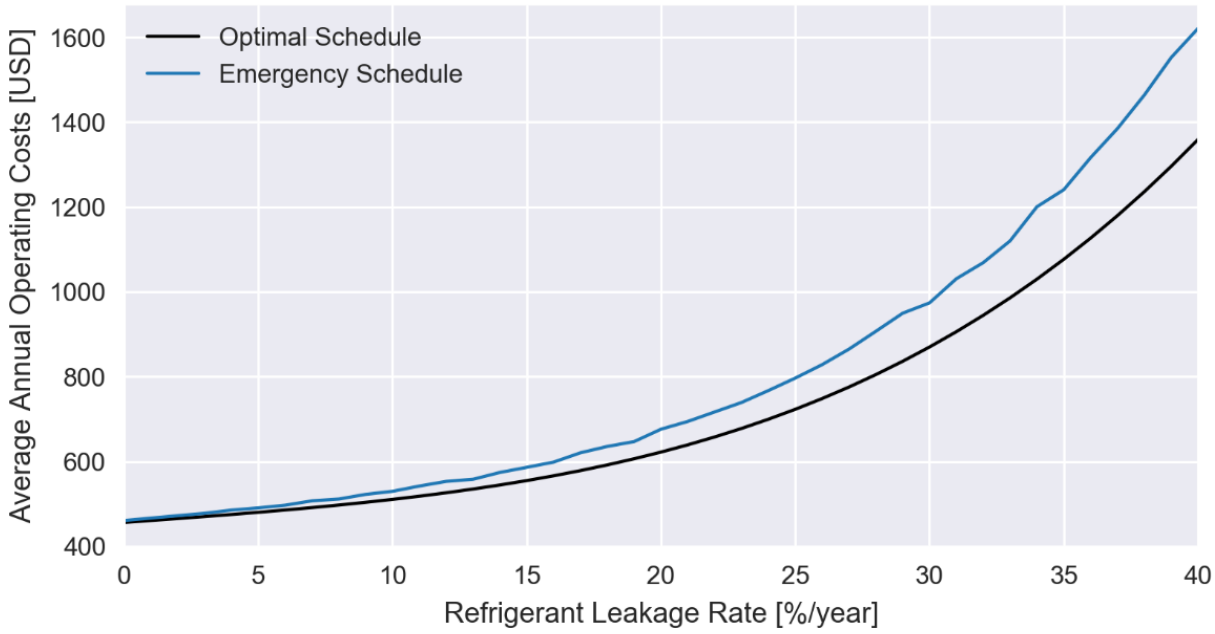

Figure 6.26. Comparison of average annual service costs for system in Chicago, IL with different refrigerant charge leakage rates using optimal and emergency service policies.

6.6 Condition-based Service Policies

Rather than performing service at fixed intervals or solely when comfort violations occur, an alternative policy based on the actual condition of the air-conditioner could be used instead. For example, service could be performed only when a significant fault is present if the air conditioner has an automated fault detection and diagnostics system installed. Furthermore, instead of calling for service when a fault is detected or diagnosed, virtual sensors could be used to call for service when performance has degraded past a certain point. For example, the virtual cooling capacity sensor could be used to monitor capacity degradation and call for service when system capacity decreases below 10% of the normal capacity.

Using real-time data to prioritize maintenance by observing the state of the system is known as condition-based maintenance. In comparison to emergency service policies, condition-based maintenance may reduce comfort violations since maintenance could be performed before

capacity is degraded significantly. In comparison to periodic maintenance policies, condition-based maintenance may reduce operating costs by requiring service time only when significant faults exist. A problem with periodic service policies is that the underlying assumptions about the building loads and rate of performance degradation must be estimated a-priori or re-evaluated annually (or at some other interval). If building loads change or the rates of degradation change, the condition-based maintenance policy would be able to adapt.

Three condition-based service policies were implemented using the simulation framework to determine how operating costs are affected. The first policy performed service when the total cooling capacity degradation exceeded a threshold, δ_{cool},

$$FIR_{cool} \underset{\omega_0}{\overset{\omega_1}{\gtrless}} \delta_{cool} \tag{6.8}$$

where FIR_{cool} is the ratio of actual cooling capacity to normal cooling capacity at the current operating condition (described in Chapter 3), ω_1 is the decision rule to perform service, and ω_0 is the decision rule to not perform service. The second policy performed service when the COP was degraded more than a threshold, δ_{COP},

$$FIR_{COP} \underset{\omega_0}{\overset{\omega_1}{\gtrless}} \delta_{COP} \tag{6.9}$$

where FIR_{COP} is the ratio of actual COP to normal COP at the current operating conditions (described in Chapter 3). The last policy considered the impact on energy consumption and performed service when the electrical energy consumed by the system exceeded a threshold, δ_{elec},

$$FIR_{elec} \underset{\omega_0}{\overset{\omega_1}{\gtrless}} \delta_{elec} \tag{6.10}$$

where FIR_{elec} is the ratio of actual energy consumed by the system to the normal energy consumption of the system at the same operating condition. Demonstrative results for these strategies will be discussed in Section 6.7 for multiple simultaneous faults.

6.7 Operating Cost-based Service Policies

To this point, several types of maintenance strategies have been implemented and examined using the simulation framework developed in the work. Reactive strategies (emergency service policies) were studied and were shown to be problematic when equipment is significantly oversized relative to the building load and require comfort violations before service is requested. Proactive maintenance strategies, like periodic service, may reduce comfort violations and decrease utility cost impacts. However, if periodic service is scheduled too often, additional service costs may outweigh any utility cost savings accrued by keeping equipment in tip-top shape. Periodic service intervals should be adjusted if building loads change or equipment starts to degrade at different rates overtime. To account for these changes, condition-based service strategies may be applied to equipment with automated fault detection and diagnostics systems. Comparing the actual performance of the equipment with a model of normal performance, service decisions can be requested when performance has been degraded significantly. Identifying the optimal degradation threshold is not trivial and depends on the equipment sizing and rate of degradation over time.

To overcome these limitations, automated fault detection and diagnostics systems can be extended even further to account for operating cost impacts of running equipment with faults and performing service over time. In other words, heuristics or simplifications to the underlying maintenance decision problem formulation can be applied to approximate the optimal solution in real-time. Rossi and Braun have developed an operating cost based service policy previously

(described in Section 6.7.1) [73]. One limitation of the Rossi and Braun's simplified approach is that it cannot directly handle multiple simultaneous faults. An extension of this policy has been developed and implemented that can be used to determine when to perform service of multiple faults in Section 0.

6.7.1 Rossi and Braun's Near Optimal Service Policy

Rossi and Braun previously developed a service decision rule based on the cost of service (C_s) and electricity cost (C_u) given by

$$C_u \int_{t_0}^{t} \bar{h}_t(x_t, f_t)\, dt \gtrless_{\omega_0}^{\omega_1} C_s \tag{6.11}$$

where t_0 is the time since service was last performed, and the net accumulated energy consumption benefit to perform service task a_i is given by

$$\bar{h}_t(x_t, f_t) = \bar{h}_{t-1}(x_{t-1}, f_{t-1}) + \frac{\Delta t_{run}}{\tau_h}\left[h_t(x_t, f_t) - \bar{h}_{t-1}(x_{t-1}, f_{t-1})\right] \tag{6.12}$$

where x_t represents the external driving conditions that affect system performance, f_t is a vector containing each fault level, and τ_h is the time constant of a low pass filter used to reduce the diurnal fluctuations on system performance and ensure $\bar{h}$ is an increasing function.

The estimated power consumption savings for performing service at any time are calculated using

$$h(x_t, f_t) = P(x_t, f_t) - P(x_t, f^0) \tag{6.13}$$

where $P(x_t, \cdot)$ is a function that estimates the power consumption of the system at given driving conditions and fault levels, and f^0 is the equipment state immediately after service (when all the

fault levels have returned to normal). The rule states that a service task should be performed when the accumulated energy impact (the left-hand-side) since the last service is greater than the service cost (right-hand-side).

The step-by-step procedure adapted from Rossi and Braun for the simplified service scheduler has been summarized below [73].

1. Set the cost of energy (C_u) and the cost of performing the service task to repair the fault (C_s) that is degrading system performance.
2. Initialize a model that estimates the expected power consumption with no performance degradation as a function of the measured driving conditions $P\left(x_t, f^0\right)$.
3. Initialize accumulator variables to zero: $\overline{h}_t$ and H_t .
4. At each decision interval (e.g. every hour), measure the power consumption, driving conditions (e.g. ambient temperature), and run-time.
5. Update the accumulator variables using the following equations:

$$h_t = P\left(x_t, f\right) - P\left(x_t, f^0\right)$$
$$\overline{h}_t = \overline{h}_{t-1} + \frac{\Delta t_{run}}{\tau_h}\left(h_t - \overline{h}_{t-1}\right)$$
$$H_t = H_{t-1} + \overline{h}_t$$
$$\omega = C_s - C_u \cdot H_t.$$

6. Compare ω with 0. If $\omega < 0$, then perform the service task and reset the accumulator variables. Additionally, if comfort violations have occurred, perform the service task.
7. Wait until the next service decision interval and then return to Step 4.

The simplified service decision policy developed by Rossi and Braun was implemented within the simulation framework and example results were generated for systems with different faults imposed.

6.7.2 Modification for Multiple Simultaneous Faults

A significant limitation to the methodology developed by Rossi and Braun is that decisions between different maintenance tasks cannot be directly handled [73]. This is because only the total utility cost impact is considered, and the relative importance of multiple faults is not estimated. Thus, it is impossible to calculate the net benefit of performing different service tasks and selecting the option that provides the maximum benefit at each decision stage. For example, when an air conditioner has relatively minor condenser fouling, yet significant evaporator fouling – the optimal service task is often to change the evaporator air filter only since it is relatively inexpensive and the evaporator fouling likely impacts the system more significantly than the condenser fouling. Using Rossi's method, this service task would be delayed until the utility cost impact became greater than the cost of evaporator fouling service and condenser fouling service. As a result, the system operates at a lower average net efficiency.

In order to improve the original simplified method developed by Rossi and Braun, previously described virtual sensor approaches for automated fault detection and diagnostics (Chapter 2) and fault impact evaluation models (Chapter 3) are used within a modified method to the estimate benefits of performing different service tasks. The inclusion of these measures provides two sources of information that can make deciding between service tasks possible: measured fault levels and isolated fault impacts. The remainder of this Chapter will describe an algorithm for deciding between service tasks step-by-step and will present simulated comparisons between the different maintenance methods for multiple fault scenarios. In addition to handling multiple

fault service decisions, the operating cost function was modified to include equipment cost impacts that account for the effects of increased equipment run-time on replacement costs.

The classification rule of Rossi and Braun has been modified to inform the service action taken at any point in time, t,

$$\int_{t_{0,i}}^{t}\left[C_u\cdot\bar{h}_t\left(x_t,f_t,a_i\right)+C_e\cdot\bar{g}_t\left(x_t,f_t,a_i\right)\right]dt\gtrless_{\omega_0}^{\omega_1}C_s\left(a_i\right)\quad\forall a_i\in A \tag{6.14}$$

where $t_{0,i}$ is the time since the i^{th} component was previously serviced, C_u is the cost per unit time of electricity consumption, C_e is the time averaged equipment replacement costs assuming the system has a finite number of run-time hours, $C_s\left(a_i\right)$ is the service cost required to perform service task a_i to repair the i^{th} component, the net accumulated energy consumption benefit to perform service task a_i is given by

$$\bar{h}_t\left(x_t,f_t,a_i\right)=\bar{h}_{t-1}\left(x_{t-1},f_{t-1},a_i\right)+\frac{\Delta t_{run}}{\tau_h}\left[h_t\left(x_t,f_t,a_i\right)-\bar{h}_{t-1}\left(x_{t-1},f_{t-1},a_i\right)\right] \tag{6.15}$$

and the net accumulated equipment runtime saving to perform service task a_i is given by

$$\bar{g}_t\left(x_t,f_t,a_i\right)=\bar{g}_{t-1}\left(x_{t-1},f_t,a_i\right)+\frac{\Delta t_{run}}{\tau_g}\left[g_t\left(x_t,f_t,a_i\right)-\bar{g}_{t-1}\left(x_{t-1},f_t,a_i\right)\right] \tag{6.16}$$

where x_t represents the external driving conditions that affect system performance, f_t is a vector containing each fault level, and τ_h, τ_g are the time constants of low pass filters used to reduce the diurnal fluctuations on system performance and ensure $\bar{h}$ and $\bar{g}$ are increasing functions.

The estimated power consumption savings for performing service task a_i at any time are calculated using

$$h\left(x_t,f_t,a_i\right)=P\left(x_t,a_0\left(f_t\right)\right)-P\left(x_t,a_i\left(f_t\right)\right) \tag{6.17}$$

where $P(x_t,\cdot)$ is a function that estimates the power consumption of the system at given driving conditions and fault levels, a_0 is the "do nothing" service task which has a functional form given by

$$a_0(f_t) \rightarrow f_t \tag{6.18}$$

which states the "do nothing" service task has no impact on the fault levels of the system and the i^{th} service task repairs selected faults

$$a_i(f_t) \rightarrow f_t^i. \tag{6.19}$$

Conceptually, this means when action a_i is applied to the system, the fault levels affected by the service task are returned to their normal values (as if the faults did not exist). The result of Equation (6.17) is the difference between power consumption for the current fault levels and the power consumption for the system if the service task was performed on the system.

In a similar manner, the runtime savings for performing service task a_i is given by

$$h(x_t, f_t, a_i) = \Delta t(x_t, a_0(f_t)) - \Delta t(x_t, a_i(f_t)) \tag{6.20}$$

where $\Delta t(x_t,\cdot)$ is a function that estimates the run-time requirement of the system at given driving conditions and fault levels. Equations (6.17) and (6.20) can be evaluated using the models described in Chapter 3.

Despite the rather complex mathematical formulation, the decision rule described by Equation (6.14) has a relatively straightforward explanation. The rule states that for all service tasks that can be applied to the system at each decision interval, a service task should be performed when the accumulated energy and equipment cost impacts (the left-hand-side) is greater than the service cost required to perform the task (right-hand-side). In order to account

for discounted service costs when performing multiple service tasks at the same time, unions between multiple service actions should be included in the set of available service actions, A.

The step-by-step procedure for implementing a simplified service scheduler that considers multiple simultaneous faults has been summarized below.

1. Set the cost of energy (C_u), the equipment cost (C_e), and define a function for determining costs of performing different service tasks to repair the faults ($C_s(a_i)$) that is degrading system performance.
2. Initialize a model that estimates the expected power consumption *with* performance degradation as a function of the measured driving conditions, $P(x_t, f_t, a_i)$, and the expected run-time for the given driving conditions, $\Delta t(x_t, f_t, a_i)$.
3. Initialize accumulator variables to zero: $H_i(t)$ and $\bar{h}_i(t)$.
4. At each decision interval (e.g. every hour), measure the power consumption, driving conditions (e.g. ambient temperature), fault levels, and run-time.
5. For all service actions, update the accumulator variables using the following equations:

$$
\begin{aligned}
h_{i,t} &= P\left(x_t, a_0(f_t)\right) - P\left(x_t, a_i(f_t)\right) \\
g_{i,t} &= \Delta t\left(x_t, a_0(f_t)\right) - \Delta t\left(x_t, a_i(f_t)\right) \\
\bar{h}_{i,t} &= \bar{h}_{i,t-1} + \frac{\Delta t_{run}}{\tau_h}\left(h_{i,t} - \bar{h}_{i,t-1}\right) \\
\bar{g}_{i,t} &= \bar{g}_{i,t-1} + \frac{\Delta t_{run}}{\tau_g}\left(g_{i,t} - \bar{g}_{i,t-1}\right) \\
H_{i,t} &= H_{i,t-1} + \bar{h}_{i,t} \\
G_{i,t} &= G_{i,t-1} + \bar{g}_{i,t} \\
\omega_i &= C_s(a_i) - C_u \cdot H_{i,t} - C_e \cdot G_{i,t}.
\end{aligned}
$$

6. Compare ω_i with 0. If $\omega_i < 0$, then perform the service task and reset the accumulator variables. Additionally, if a service task is performed, reset other accumulator variables that include the service task. If comfort violations have occurred, perform the service task.
7. Wait until the next service decision interval and then return to Step 4.

The modified service decision rule that considers multiple service actions was implemented in the simulation framework and evaluated for different combinations of faults. Because choosing combinations of fault rates is somewhat arbitrary due to the lack of reliable fault prevalence data, rates for refrigerant charge leakage, condenser fouling, and evaporator fouling were sampled from random distributions [9]. To consider a relatively wide range of combinations of fault rates, uniformly distributed random samples were chosen. For refrigerant leakage fault rates, a uniformly distributed random sample between 0% to 20% leakage per year was selected for each trial. For condenser fouling fault rates, a uniformly distributed random sample between 0% to 20% per 5000 hours of condenser fan runtime was selected for each trial. For evaporator fouling fault rates, a uniformly distributed random sample between 0% to 20% per 5000 hours of evaporator fan runtime was selected for each trial. A distribution of 200 samples was randomly selected from the distributions described and is shown in Figure 6.27.

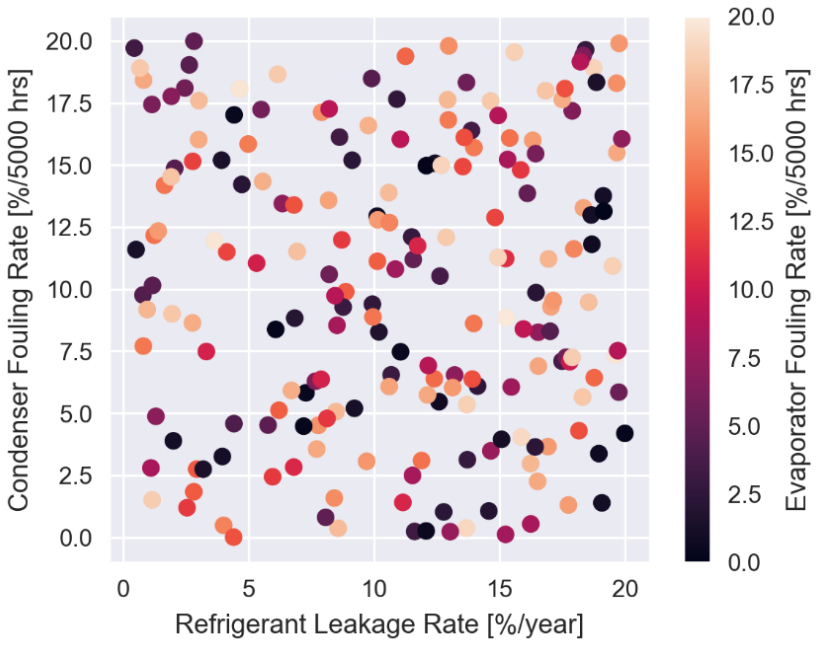

Figure 6.27. Distribution of randomly sampled fault rates used to evaluate the performance of different service decision strategies for multiple fault scenarios.

After sampling the fault rates randomly, the optimal service decision policy and optimal lifetime operating costs for locations in Miami, Atlanta, and Chicago were determined for scenarios with each combination of faults. Next, the simulations at each location and fault combination were repeated using the different service policies previously described. The result of this process was a distribution of lifetime operating costs for each location and service policy studied. Finally, the lifetime operating cost of each policy was compared with the corresponding optimal operating cost.

The resulting distributions of additional lifetime operating costs relative to optimal costs using the original (unmodified) service scheduler developed by Rossi and Braun for the random sample of fault combinations are shown in Figure 6.28. In this policy, the total accumulated energy impact is estimated and used to calculate the additional utility cost due to faults. At each decision interval, the net utility costs are compared to the costs of performing the three service

tasks: adding refrigerant, cleaning the condenser coil, and changing the evaporator coil filter. Because all three service tasks are considered, this requires greater accumulated utility cost impact before service is performed. The additional lifetime operating costs relative to the optimal costs in Miami, Atlanta, and Chicago were 6.9%, 7.4%, and 8.7% respectively. Because Miami has the longest cooling season, this location has higher normal electrical energy consumption for cooling than the other locations. This makes preventative maintenance more cost effective since larger energy consumption savings are possible. The additional costs in Atlanta and Chicago are greater than Miami due to lower cooling load requirements. These locations also have a winter season, where no cooling is required which makes the service policy less effective. This is because the policy is not able to quantify the impact of charge leakage during the winter months – which makes doing service late in cooling season possible and creates a lag in the spring before accumulated impacts become greater than the cost of service. These are two behaviors the optimal policy can avoid since the optimization horizon is over the entire equipment life.

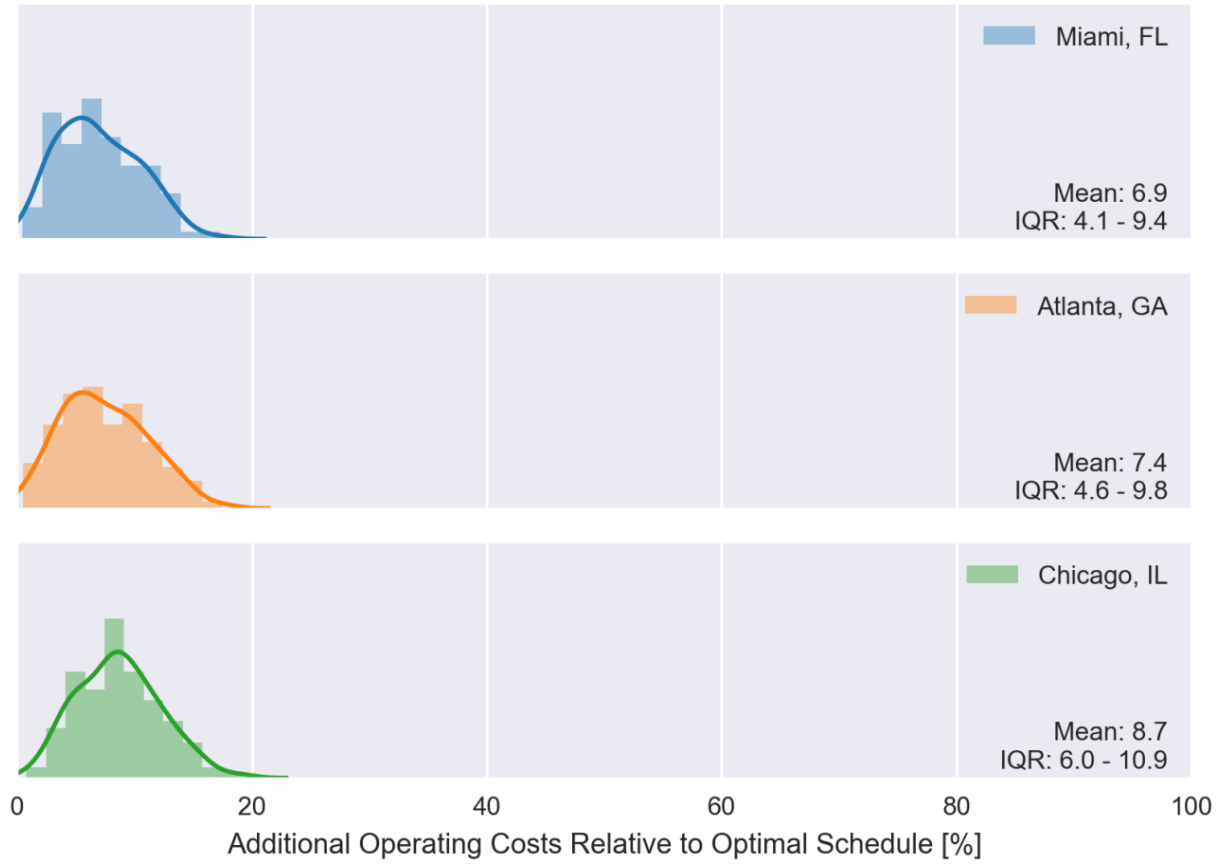

Figure 6.28. Additional lifetime operating costs relative to optimal lifetime costs when the simplified service decision methodology developed by Rossi and Braun is used to service multiple faults in three different locations.

The distributions of additional lifetime operating costs using the modified service scheduler that considers multiple service tasks for the random sample of fault combinations is shown in Figure 6.29. In this policy, the total accumulated energy impacts for different faults are estimated and used to calculate the additional utility cost consumed. At each decision interval, the net utility costs are compared to the costs of performing different combinations of service tasks. When the cost of one of the combinations of service tasks becomes less than the accumulated utility costs for the corresponding faults, service is performed. The additional lifetime operating costs relative to the optimal costs in Miami, Atlanta, and Chicago were 3.7%, 5.5%, and 5.7% respectively. Compared to the simplified service scheduler developed by Rossi

and Braun, additional operating costs savings are possible using the methodology that isolates the impacts of different faults. The remaining costs above the optimal operating costs are caused by suboptimal scheduling around the winter season and not considering end of life payback intervals.

Figure 6.29. Additional lifetime operating costs relative to optimal lifetime costs when the modified service decision methodology that considers multiple service tasks is used to service multiple faults in three different locations.

The results collected from the annual and biennial service policies for each location are shown in Figure 6.30 and Figure 6.33. These results show that annual service policies generally increase operating costs for the faults considered. This is because the cost for service is significantly greater than typical utility cost savings that are possible on a year over year basis.

For systems in locations with higher cooling loads, annual service has lower lifetime operating costs. Biennial service policies reduce operating costs compared to annual service policies. However, they still have relatively high costs when compared with the policies that consider accumulated operating costs estimates. It should also be noted that for systems that have low fault rates, annual and biennial service costs add significant operating costs.

Figure 6.30. Additional lifetime operating costs relative to optimal lifetime costs when annual service schedules are used to service multiple faults in three different locations.

Figure 6.31. Additional lifetime operating costs relative to optimal lifetime costs when biennial service schedules are used to service multiple faults in three different locations.

The additional costs of performing emergency service policies for the random sample of fault rates are shown in Figure 6.33. For systems with higher cooling loads, these policies tend to result in higher lifetime energy consumption. This occurs because some faults, like condenser fouling have small effects on cooling capacity yet significant impacts on efficiency. It should also be noted that the emergency service policies generally have the largest inner-quartile ranges of all the policies considered. These results occur since some fault cases have low fault rates. If faults evolve slowly over time (or if a system never gets a fault) emergency service policies are optimal.

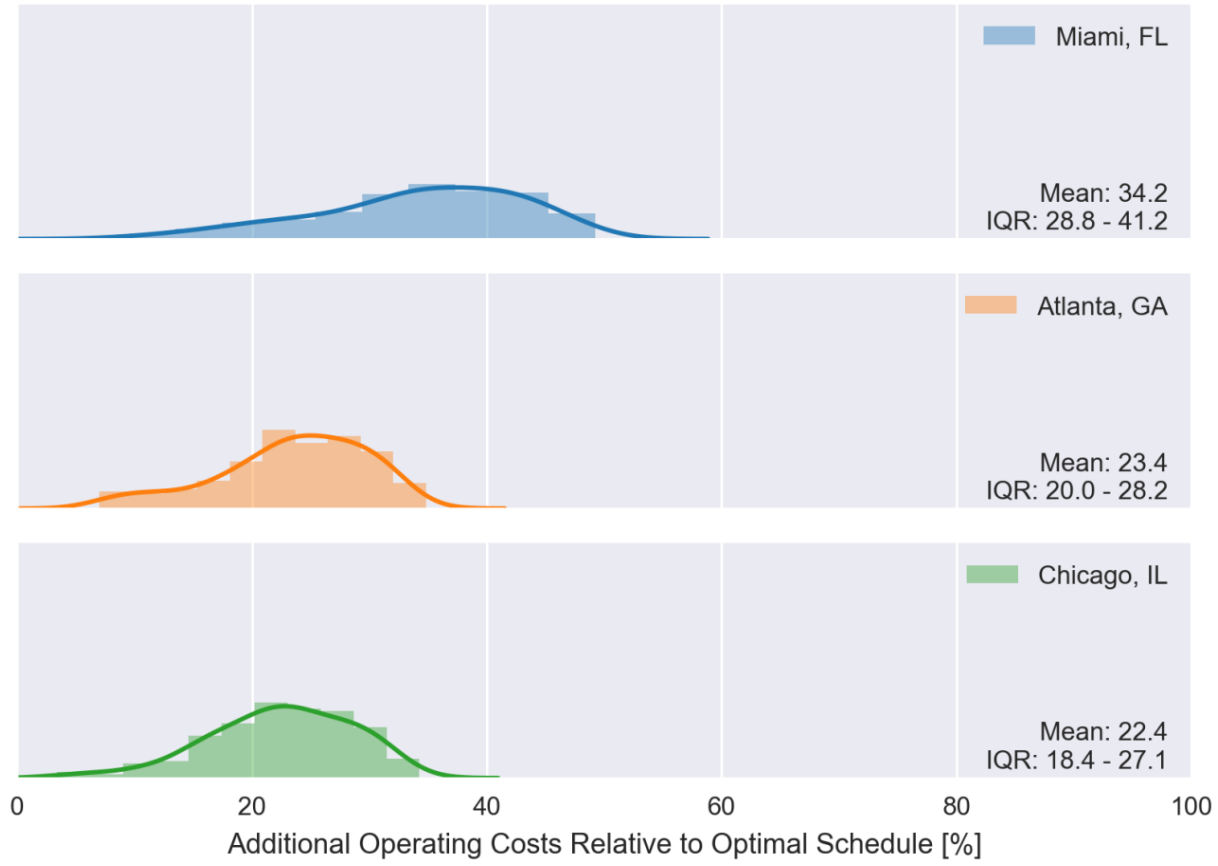

Figure 6.32. Additional lifetime operating costs relative to optimal lifetime costs when emergency service schedules are used to service multiple faults in three different locations.

The additional costs of performing service policies based on cooling capacity degradation thresholds for the random sample of fault rates are shown in Figure 6.33. For systems with smaller cooling loads, these policies tend to result in higher lifetime energy consumption since service is performed more often than the optimal schedules. On the other hand, faults that have insignificant impacts on cooling capacity are ignored for longer periods of time. These two factors contribute to larger inner quartile ranges than some of the other strategies. In general, service decision policies that consider only cooling capacity degradation are not good choices due to these variations in results.

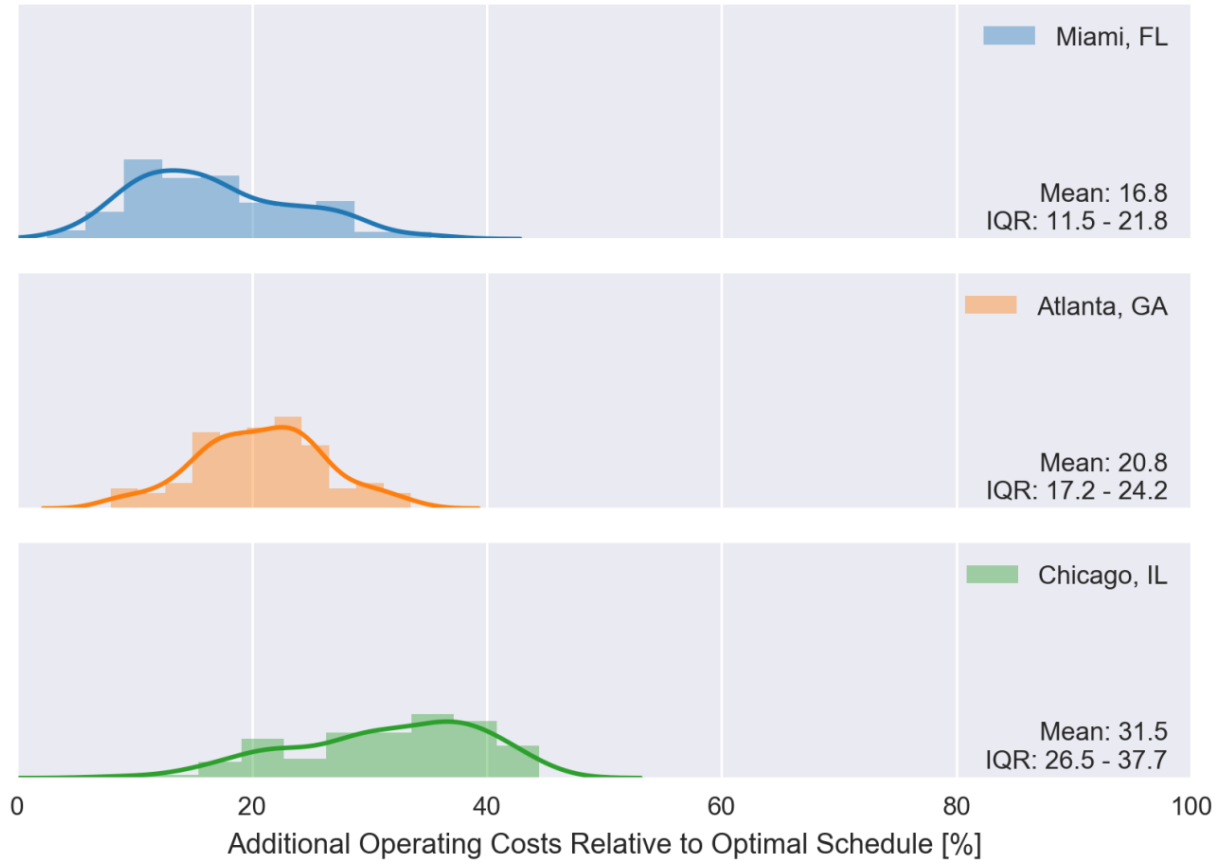

Figure 6.33. Additional lifetime operating costs relative to optimal lifetime costs when service is performed when capacity is degraded by 10% considering multiple faults in three different locations.

The additional costs of performing service policies based on COP degradation and increases in power consumption for the random sample of fault rates are shown in Figure 6.34 and Figure 6.35. These policies tend to decrease operating costs more than capacity-based policies. This is because COP and power consumption capture the relationship between operating costs and efficiency or power consumption more effectively. It should also be noted that the policy based on energy consumption has lower average operating costs than the COP based policy. This is because the energy consumption-based policy characterizes run-time impacts based the faults impact on sensible heat ratio.

Figure 6.34. Additional lifetime operating costs relative to optimal lifetime costs when service is performed when COP is degraded by 10% considering multiple faults in three different locations.

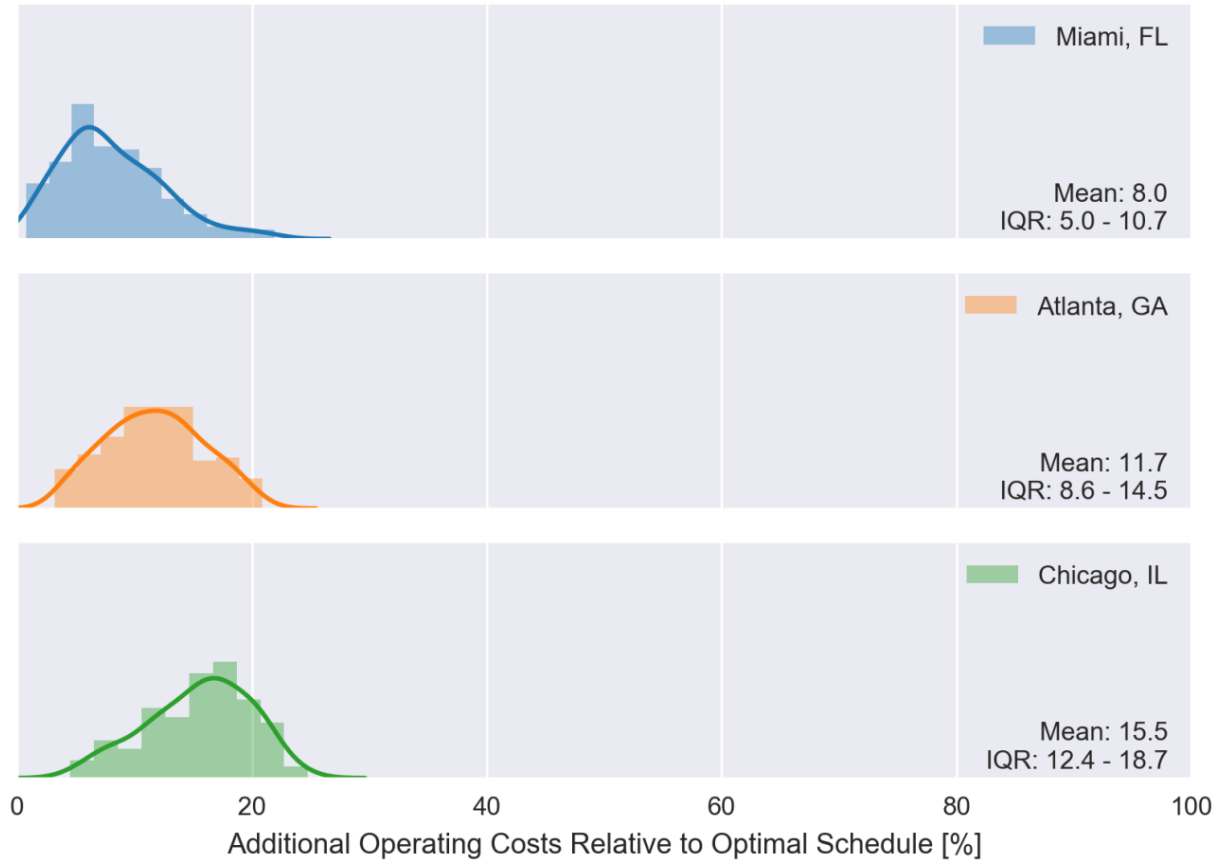

Figure 6.35. Additional lifetime operating costs relative to optimal lifetime costs when service is performed when energy consumption is increased by 10% considering multiple faults in three different locations.

A summary comparing the inner-quartile range and mean additional lifetime operating costs for all the results is included in Table 6.4. In general, policies that do not consider the condition of the system while determining when to do service (periodic service schedules) have the highest lifetime operating costs. Policies that consider the state of the equipment tend to have lower lifetime operating costs, though the metric used to determine when to do service has a significant impact on actual costs. Even in the extreme case, emergency service policies may have lower operating costs than periodic service policies since service costs are saved for systems have minimal faults. Utilizing more information when determining when to do service generally reduces operating costs. These results were observed using the simplified service scheduler

proposed by Rossi and Braun and with the modified service scheduler that considers different service tasks. These strategies tended to have the lowest lifetime operating costs.

Table 6.4. Summary of additional lifetime operating costs relative to optimal lifetime costs determined using dynamic programming for similar buildings in different locations. Inner-quartile ranges and means of 200 randomly sampled fault scenarios are compared.

	Miami, FL			Atlanta, GA			Chicago, IL		
Service Policy	25%	Mean	75%	25%	Mean	75%	25%	Mean	75%
Annual	30.0	31.5	33.0	41.7	44.6	47.2	61.6	65.2	68.6
Biennial	17.1	18.3	19.7	29.0	30.6	32.6	36.2	38.8	40.9
Emergency	28.8	34.2	41.2	20.0	23.4	28.2	18.4	22.4	27.1
CBM – Capacity 10% Threshold	11.5	16.8	21.8	17.2	20.8	24.2	17.7	21.0	25.1
CBM – COP 10% Threshold	6.2	10.4	13.8	12.8	15.3	18.1	15.7	20.2	25.8
CBM – Energy 10% Threshold	5.0	8.0	10.7	8.6	11.7	14.5	12.4	15.5	18.7
Simplified Suboptimal (Rossi and Braun [73])	4.1	6.9	9.4	4.6	7.4	9.8	6.0	8.7	10.6
Modified, Multi-task	1.4	3.7	4.9	3.1	5.5	7.3	3.5	5.7	7.4

Even using the service decision strategies that consider accumulated impacts, the lifetime operating costs were still appreciably higher than the optimal service schedule costs. This occurs for two reasons: 1) the policies do not handle scheduling service around winter seasons and 2) the policies do not optimize service well towards the end of equipment life. It is believed that these deficiencies could be corrected in future work. One possible solution to these problems would be to adapt the service schedulers to use a future optimization horizon. This would require a model for expected utility cost savings during the prediction horizon, as well as model for how the faults would evolve. Since the loads and operation throughout the year and life of HVAC equipment is largely cyclical, it may be possible to learn this model using trended data from past performance.

A second improvement to the models that may help to avoid thermal comfort violations would be to pre-schedule service using a model for the peak cooling loads and the capacity degradation measurements. Rather than allowing the system to cause thermal comfort violations due to insufficient capacity, an algorithm that estimates the peak cooling load over a prediction horizon could be used to calculate the minimum capacity needed. Using virtual sensor measurements, a service schedule could determine if the system will be able to meet all future loads and schedule service if it cannot.

6.8 Summary and Conclusions

Several different types of maintenance strategies have been implemented and compared using a simulation framework that models the interaction between building cooling loads and equipment performance while faults evolve over time. As a common benchmark, dynamic programming was used to find solutions to an optimal service scheduling problem that was formulated to minimize lifetime operating costs by performing service tasks when they are most cost effective during the equipment life. For each simulation scenario considered, the optimal service scheduling problem was solved, and an optimal service policy function was used to calculated optimal operating costs.

Simple, fixed interval service policies were compared with the optimal policies for different fault rates. These comparisons showed that periodic maintenance policies can often lead to significant increases in operating costs, especially if faults grow slowly over time. Reactive maintenance strategies that perform maintenance only when comfort is violated were also implemented and simulated using the framework. Policies that base their decisions on the condition of equipment tended to have less lifetime operating costs.

7. CONCLUSION AND RECOMMENDATIONS

Three processes involved in HVAC system monitoring have been studied using simulated and experimental data: automated fault detection and diagnostics, fault impact evaluation, and service decision-making. Automated fault detection and diagnostics aims to reduce system operating costs by identify performance problems sooner. Automated diagnostics also attempts to provide additional information to service technicians which may reduce time required for troubleshooting and reduce misdiagnoses. Fault impact evaluation can be used to estimate the performance impacts of faults on system performance. This information can be used by building owners or service technician to determine which repairs, if any, have economic incentives. These considerations lead directly into service decision-making where long-term utility costs and comfort impacts of faults can be balanced with preventative maintenance costs. This thesis described work on automated fault detection, diagnostics, impact evaluation, and comparisons between service decision-making policies for packaged air conditioning system for commercial buildings.

Three main research objectives were considered in this work related to detecting and management of air conditioner performance degradation. First, a low-cost approach for automatically detecting and diagnosing degradation faults was implemented and evaluated for multiple direct-expansion (DX) air-conditioning units in a laboratory setting. This included systems with both fixed-speed compressors and variable-speed compressors, as well as systems with fixed orifice expansion valve (FXO) and thermostatic expansion valves (TXV). Furthermore, systems with traditional round-tube plate-fin condenser coils and systems with microchannel condenser coils were evaluated experimentally in this work. The automated fault detection and diagnostics (FDD) approach used virtual sensors to identify and isolate different

faults that cause performance degradation. Faults were considered that reduce capacity, efficiency, and/or sensible cooling fractions over time without causing abrupt failures that completely disable equipment operation. Left unnoticed, these faults can lead to significant increases in equipment energy consumption while still maintain thermal comfort in the spaces they serve.

Adoption of automated FDD systems in the HVAC market has been limited due to the high perceived costs of these systems relative to typical DX equipment costs. As a demonstration of what is possible, a working prototype of the virtual sensor based automated FDD system was implemented using a widely available microprocessor and low-cost sensors. The hardware and software design were implemented for less than $120 using off-the-shelf components. Taking advantage of economies of scale, it is believed these costs could be further reduced using mass production. The system was able to estimate different physical quantities of interest, like refrigerant charge level, refrigerant and air mass flow rates, cooling capacity, and power consumption within 10% of the directly measured values over a range of operating conditions. Additionally, the automated FDD system was able to detect 10% deviations in refrigerant charge and airflow rates with a statistical confidence of 90% in experimental lab tests.

The second main objective of this work was to develop and evaluate improved models for estimating the performance impacts of faults on system performance. There was special consideration of systems with multiple faults and the development of simplified models that could be used as part of an AFDD system for assessing the relative performance impacts of each fault. To do this, a previously developed detailed component-based fault impact model was used to generate extensive data sets for implementing and testing simplified models. Additional experimental data was collected and used to test these models when applied to fixed-speed and

variable-speed rooftop units. The methodology developed uses semi-empirical models for estimating the normal performance of an air conditioner without faults and compares these estimates with the outputs of virtual sensors. Using experimental data, it was shown that the semi-empirical models were able to predict the normal cooling capacity, power consumption, and sensible heat ratio of fixed-speed and variable-speed rooftop units within 10% of the measured data.

To isolate the impacts of multiple simultaneous faults, the semi-empirical normal models were extended to include conditionally independent correction factors that considered the sensitivity of faults on system performance. The correction factors were generally simple, low-order polynomial models which had limited empirical parameters that could be determined using linear regression. Using the detailed performance model, the fault impact of different faults could be isolated to within 10% for 99% of the data points considered. This included combinations refrigerant charge faults, condenser fouling faults, and evaporator fouling faults over the typical operating ranges of air-conditioning equipment.

The third and final research objective of this work was related to comparing maintenance and service strategies used for commercial building air conditioners. Depending on how equipment is serviced throughout its life, the money spent for energy consumption and equipment maintenance can be significantly impacted. To understand these relationships, a simulation framework was developed that modeled the relationship between building cooling loads and equipment performance as faults evolve over time. The simulation framework embedded a neural network approximation of the detailed fault impact model used to investigate performance impacts of faults. This model made it possible to predict cooling capacity and energy

consumption in a computationally efficient manner so that optimal maintenance schedules over the life of the equipment could be studied and compare with alternative service strategies.

Several different types of maintenance strategies were implemented and compared using the simulation. As a common benchmark, dynamic programming was used to find solutions to an optimal service scheduling problem that was formulated for this work. For each simulation scenario considered, an optimal service policy function was determined that minimized lifetime operating costs by performing service tasks when they are most cost effective during the equipment life. Simple, fixed interval service policies were compared with the optimal policies for different fault rates. These comparisons showed that periodic maintenance policies can often lead to significant increases in operating costs, especially if faults grow slowly over time. Reactive maintenance strategies that perform maintenance only when comfort is violated were also implemented and simulated using the framework.

Service policies that use the outputs of virtual sensors to make decisions were also considered in this work. Condition-based policies that required service whenever a performance metric decreases beyond a threshold were implemented. Thresholds that were studied considered reductions in cooling capacity, COP, and increases in estimated energy consumption were considered. These policies tended to perform well for faults that evolve slowly over time, though begin to struggle when faults change performance more rapidly. The final group of policies considered in this work were based on estimations of the accumulated energy cost impacts of faults. Simple heuristic rules that compared the accumulated energy cost impacts with the costs to perform service tasks were implemented and performed better than other methodologies when single faults were considered. When multiple faults affect system performance, the policies must isolate the impacts of different faults and compare these impacts

to different service costs. A modified methodology was developed and implemented in this work that handles multiple simultaneous faults and considers the impact of faults on equipment life in an approximate manner. This approach had the best performance of simple strategies when compared to the optimal scheduling benchmark.

This work considered optimal maintenance scheduling of fixed-speed rooftop air conditioners during cooling seasons. Additional work developing methodologies to estimate the impacts and schedule service for variable-speed systems will be required as technology improves. This will require fair and reproducible laboratory equipment testing of variable-speed systems which does not currently have an industry accepted standard methodology. Initial work on the development of a test methodology for fixed-speed and variable-speed equipment that does not require disabling native controls has been included in Appendix A. This methodology could be used to further test automated FDD systems in a representative and controlled environment that more closely simulates actual building applications.

Developmental work for supermarket refrigeration applications should be considered since energy costs per unit cooling capacity are greater than those for air conditioning and systems already have many sensors that are required for automated FDD. These systems must provide cooling every hour of the day throughout the year and maintain increasingly tight tolerances on food temperatures. As energy becomes more expensive to consume, optimization of equipment control, performance, and preventative maintenance will become increasingly important. Much of the ideas and approaches developed in this work may be suitable for these applications, but further research is needed.

Finally, as variable-speed systems with electronic valves and controls become more and more common, additional research into fault tolerant controls and fault impact mitigation should

be investigated. For each electronic control device added to a system, the number of control freedoms increases and may make it possible to reduce the impacts of faults on systems performance. If it is possible to reduce the energy impacts of faults by adapting equipment control, it may be possible to delay when service is needed to reduce average operating costs.

REFERENCES

1. Rossi, T. and J. Braun, *A Statistical, Rule-Based Fault Detection and Diagnostic Method for Vapor Compression Air Conditioners*. HVAC&R Research, 1997. **3**(1): p. 19-37.
2. Kim, W. and J.E. Braun, *Extension of a virtual refrigerant charge sensor*. International Journal of Refrigeration, 2015. **55**: p. 224-235.
3. Kim, W. and J.E. Braun, *Development and evaluation of virtual refrigerant mass flow sensors for fault detection and diagnostics*. International Journal of Refrigeration, 2016. **63**: p. 184-198.
4. Li, H. and J. Braun, *A Methodology for Diagnosing Multiple Simultaneous Faults in Vapor-Compression Air Conditioners*. HVAC&R Research, 2007. **13**(2): p. 369-395.
5. Li, H. and J. Braun, *Virtual Refrigerant Pressure Sensors for Use in Monitoring and Fault Diagnosis of Vapor-Compression Equipment*. HVAC&R Research, 2009. **15**(3): p. 597-616.
6. Li, H. and J.E. Braun, *Decoupling features and virtual sensors for diagnosis of faults in vapor compression air conditioners*. International Journal of Refrigeration, 2007. **30**(3): p. 546-564.
7. Katipamula, S. and M. Brambley, *Review Article: Methods for Fault Detection, Diagnostics, and Prognostics for Building Systems—A Review, Part I*. HVAC&R Research, 2005. **11**(1): p. 3-25.
8. Katipamula, S. and M. Brambley, *Review Article: Methods for Fault Detection, Diagnostics, and Prognostics for Building Systems—A Review, Part II*. HVAC&R Research, 2005. **11**(2): p. 169-187.
9. Yuill, D.P. and J.E. Braun, *A figure of merit for overall performance and value of AFDD tools*. International Journal of Refrigeration, 2017. **74**: p. 651-661.
10. Yuill, D.P. and J.E. Braun, *Evaluating the performance of fault detection and diagnostics protocols applied to air-cooled unitary air-conditioning equipment*. HVAC&R Research, 2013. **19**(7): p. 882-891.
11. Yuill, D.P. and J.E. Braun, *Effect of the distribution of faults and operating conditions on AFDD performance evaluations*. Applied Thermal Engineering, 2016. **106**: p. 1329-1336.

12. Cowan, A., *Review of Recent Commercial Roof Top Unit Field Studies in the Pacific Northwest and California*. 2004, New Buildings Institute.
13. Jacobs, P., *Small HVAC Problems and Potential Savings Reports*. 2003, California Energy Commission.
14. Li, H. and J.E. Braun, *An overall performance index for characterizing the economic impact of faults in direct expansion cooling equipment*. International Journal of Refrigeration, 2007. **30**(2): p. 299-310.
15. Kim, W. and J.E. Braun, *Performance evaluation of a virtual refrigerant charge sensor*. International Journal of Refrigeration, 2013. **36**(3): p. 1130-1141.
16. Li, H. and J. Braun, *Development, Evaluation, and Demonstration of a Virtual Refrigerant Charge Sensor*. HVAC&R Research, 2009. **15**(1): p. 117-136.
17. Breuker, M. and J. Braun, *Evaluating the Performance of a Fault Detection and Diagnostic System for Vapor Compression Equipment*. HVAC&R Research, 1998. **4**(4): p. 401-425.
18. Chen, B. and J.E. Braun, *Simple Rule-based Methods for Fault Detection and Diagnosis Applied to Packaged Air Conditioners*. ASHRAE Transactions, 2001. **107**(1).
19. Siegel, J.A. and C.P. Wray, *An Evaluation of Superheat-Based Refrigerant Charge Diagnostics for Residential Cooling Systems*. ASHRAE Transactions, 2002. **108**(2).
20. Armstrong, P.R., et al., *Detection of Rooftop Cooling Unit Faults Based on Electrical Measurements*. HVAC&R Research, 2006. **12**(1): p. 25.
21. Kim, M., S.H. Yoon, and W.V. Payne, *Cooling Mode Fault Detection and Diagnosis Method for a Residential Heat Pump*. 2008, National Institute of Standards and Technology: Gaithersburg, MD.
22. Li, H. and J.E. Braun, *Decoupling features for diagnosis of reversing and check valve faults in heat pumps*. International Journal of Refrigeration, 2009. **32**(2): p. 316-326.
23. Kim, W. and J.E. Braun, *Evaluation of the impacts of refrigerant charge on air conditioner and heat pump performance*. International Journal of Refrigeration, 2012. **35**(7): p. 1805-1814.
24. Kim, W. and S. Katipamula, *A review of fault detection and diagnostics methods for building systems*. Science and Technology for the Built Environment, 2018. **24**(1): p. 3-21.

25. Kim, M., et al., *Performance of a Residential Heat Pump Operating in the Cooling Mode with Single Faults Imposed.* Applied Thermal Engineering, 2009. **29**(4): p. 9.
26. Wichman, A. and J. Braun, *Fault Detection and Diagnostics for Commercial Coolers and Freezers*. HVAC&R Research, 2009. **15**(1): p. 77-99.
27. Commission, C.E., *Building Energy and Efficiency Standards for Residential and Nonresidential Buildings*. 2013, California Energy Commission.
28. Katipamula, S., K. Gowri, and G. Hernandez, *An open-source automated continuous condition-based maintenance platform for commercial buildings.* Science and Technology for the Built Environment, 2016. **23**(4): p. 546-556.
29. Katipamula, S., et al., *Transactional network: Improving efficiency and enabling grid services for buildings.* Science and Technology for the Built Environment, 2016. **22**(6): p. 643-654.
30. Grace, I.N., D. Datta, and S.A. Tassou, *Sensitivity of Refrigeration System Performance to Charge Levels and Parameters for On-line Leak Detection.* Applied Thermal Engineering, 2005. **25**(4): p. 10.
31. Hjortland, A.L., *Probabilistic Fault Detection and Diagnostics for Packaged Air-Conditioner Outdoor-Air Economizers*, in *Department of Mechanical Engineering*. 2014, Purdue University.
32. Piacentino, A. and M. Talamo, *Innovative thermoeconomic diagnosis of multiple faults in air conditioning units: Methodological improvements and increased reliability of results.* International Journal of Refrigeration, 2013. **36**(8): p. 2343-2365.
33. Yu, D., H. Li, and M. Yang, *A virtual supply airflow rate meter for rooftop air-conditioning units*. Building and Environment, 2011. **46**(6): p. 1292-1302.
34. ANSI/AHRI, *Standard for Performance Rating of Positvie Displacement Refrigerant Compressors and Compressor Units*. 2004, Air-Conditioning, Heating, and Refrigeration Institute.
35. Dabiri, A.E. and C.K. Rice, *A Compresosor Simulation Model with Corrections for the Level of Suction Gas Superheat.* ASHRAE Transactions, 1981. **87**(2).
36. Payne, V. and D. O'Neal, *A Mass Flow Rate Correlation for Refrigerants and Refrigerant Mixtures Flowing Through Short Tubes.* HVAC&R Research, 2004. **10**(1): p. 73-87.

37. Lenger, M.J., A.M. Jacobi, and P.S. Hrnjak, *Superheat Stability of an Evaporator and Thermostatic Expansion Valve*. 1998, Air-Conditioning Research Center: Univeristy of Illinois, Urbana-Champaign.
38. Foundation, P.S., *Python Language Reference, version 3.5*.
39. ANSI/AHRI, *Standard for Performance Rating of Unitary Air-Conditioning & Air-Source Heat Pump Equipment*. 2009, Air-Conditioning, Heating, and Refrigeration Institute.
40. Patil, A., et al. *Development and Evaluation of Automated Virtual Refrigerant Charge Sensor Training Kit*. in *International Refrigeration and Air Conditioning Conference*. 2016. West Lafayette, IN.
41. ANSI/ASHRAE, *Standard Methods for Laboratory Airflow Measurement*. 1993, American Society of Heating, Refrigeration, and Air-Conditioning Engineers.
42. ANSI/ASHRAE, *Standard Method for Temperature Measurement*. 2013, American Society of Heating, Refrigeration, and Air-Conditioning Engineers.
43. ANSI/ASHRAE, *Standard Method for Humidity Measurement*. 2014, American Society of Heating, Refrigeration, and AIr-Conditioning Engineers.
44. Breuker, M. and J. Braun, *Common Faults and Their Impacts for Rooftop Air Conditioners*. HVAC&R Research, 1998. **4**(3): p. 303-318.
45. Harms, T., E. Groll, and J. Braun, *Accurate Charge Inventory Modeling for Unitary Air Conditioners*. HVAC&R Research, 2003. **9**(1): p. 55-78.
46. Shen, B., J.E. Braun, and E.A. Groll, *Improved methodologies for simulating unitary air conditioners at off-design conditions*. International Journal of Refrigeration, 2009. **32**(7): p. 1837-1849.
47. Palmiter, L., et al., *Measured Effect of Airflow and Refrigerant Charge on the Seasonal Performance of an Air-Source Heat Pump using R-410A*. Energy and Buildings, 2011. **43**(7): p. 9.
48. Cheung, H. and J.E. Braun, *Simulation of Fault Impacts for Vapor Compression Systems by Inverse Modeling. Part I: Component Modeling and Validation*. HVAC&R Research, 2013. **19**(7): p. 26.

49. Cheung, H. and J.E. Braun, *Simulation of Fault Impacts for Vapor Compression Systems by Inverse Modeling. Part II: System Modeling and Validation.* HVAC&R Research, 2013. **19**(7): p. 15.
50. Brandemuehl, M.J. and S. Gabel, *Development of a Toolkit for Secondary HVAC System Energy Calculations.* ASHRAE Transactions, 1994. **99**(1).
51. Nyika, S., et al., *Generalized Performance Maps for Single- and Dual-Speed Residential Heat Pumps*. ASHRAE Transactions, 2014. **120**(2): p. 16.
52. Incropera, F.P., et al., *Fundamentals of Heat and Mass Transfer*. 6th ed. ed. 2007, Hoboken, NJ: John Wiley & Sons, Inc.
53. Nyika, S., et al., *Generalized Performance Maps for Variable-Speed Residential Heat Pumps*. ASHRAE Transactions, 2013. **119**(1): p. 8.
54. Cheung, H. and J.E. Braun, *Performance mapping for variable-speed ductless heat pump systems in heating and defrost operation.* HVAC&R Research, 2014. **20**(5): p. 545-558.
55. Cheung, H. and J.E. Braun, *Component-based, gray-box modeling of ductless multi-split heat pump systems*. International Journal of Refrigeration, 2014. **38**: p. 30-45.
56. Cai, J. and J.E. Braun, *Summary of the Variable-Speed RTU Model Trained with Anonymous Manufacturer's Performance Data - DLAT Version*. 2016, Center for High Performance Buildings.
57. ANSI/ASHRAE, *Methods of Testing for Rating Electrically Driven Unitary Air-Conditioning and Heat Pump Equipment*. 2009, American Society of Heating, Refrigeration, and Air-Conditioning Engineers.
58. ANSI/ASHRAE, *Standard Methods for Pressure Measurement*. 2014, American Society of Heating, Refrigeration, and Air-Conditioning Engineers.
59. ANSI/ASHRAE, *Standard Methods for Refrigerant Mass Flow Measurement Using Flowmeters*. 2013, American Society of Heating, Refrigeration, and AIr-Conditioning Engineers.
60. ANSI/ASHRAE, *Standard Methods for Power Measurement*. 2014, American Society of Heating, Refrigeration and Air-Conditioning Engineers.
61. ASHRAE, *Handbook of Fundamentals*. 2017, Atlanta, GA: American Society of Heating, Refrigeration, and Air-Conditioning Engineers.

62. Taylor, S.T. and C.H. Cheng, *Economizer High Limit Devices and Why Enthalpy Economizers Don't Work*, in *ASHRAE Journal*. 2010, American Society of Heating, Refrigeration, and Air-Conditioning Engineers. p. 11.
63. Cheung, H. and J.E. Braun, *An empirical model for simulating the effects of refrigerant charge faults on air conditioner performance.* Science and Technology for the Built Environment, 2016. **23**(5): p. 776-786.
64. Du, Z., P.A. Domanski, and W.V. Payne, *Effect of Common Faults on the Performance of Different Types of Vapor Compression Systems.* Appl Therm Eng, 2016. **98**: p. 61-72.
65. Chollet, F., *Keras*. 2015, Github.
66. Goodfellow, I., Y. Bengio, and A. Courville, *Deep Learning*. 2016, Cambridge, MA: MIT Press.
67. Wilcox, S. and W. Marion, *Users Manual for TMY3 Data Sets*. 2008, National Renewable Energy Laboratory: Golden, CO.
68. Sherif, Y.S. and M.L. Smith, *Optimal Maintenance Models for Systems Subject to Failure - A Review.* Naval Research Logistics, 1981. **28**(1): p. 47-75.
69. Barlow, R.E. and F. Proschan, *Mathematical Theory of Reliability*. Classics in Applied Mathematics. 1996, Philadelphia: Society for Industrial and Applied Mathematics.
70. Barlow, R.E., *Engineering Reliability*. ASA-SIAM Series on Statistics and Applied Probability. 1998, Philadelphia: Society for Industrial and Applied Mathematics.
71. Valdez-Flores, C. and R.M. Feldman, *A Survey of Preventative Maintenance Models for Stochastically Deteriorating Single-Unit Systems.* Naval Research Logistics, 1989. **36**(1): p. 419-446.
72. Wilson, E.J.H., et al., *A parametric study of energy savings from cleaning coils and filters in constant air volume HVAC systems.* HVAC&R Research, 2013. **19**(5): p. 616-626.
73. Rossi, T. and J. Braun, *Minimizing Operating Costs of Vapor Compression Equipment with Optimal Service Scheduling.* HVAC&R Research, 1996. **2**(1): p. 3-25.
74. Kirk, D.E., *Optimal Control Theory: An Introduction*. Dover Books on Electrical Engineering. 2004: Dover Publications.
75. Denardo, E.V., *Dynamic Programming: Models and Applications*. Dover Books on Computer Science. 2003: Dover Publications.

76. Bellman, R., *Dynamic Programming*. Dover Books on Computer Science. 2003: Dover Publications.

77. ANSI/AHRI, *Standard for Performance Rating of Commercial and Industrial Unitary Air-Conditioning and Heat Pump Equipment*. 2011, Air-Conditioning, Heating, and Refrigeration Institute.

78. ANSI/AHRI, *Standard for Performance Rating of Variable Refrigerant Flow (VRF) Mult-Split Air-Conditioning and Heat Pump Equipment*. 2010, Air-Conditioning, Heating, and Refrigeration Institute.

79. ANSI/ASHRAE, *Methods of Testing for Rating Seasonal Efficiency of Unitary Air Conditioners and Heat Pumps*. 2010, American Society for Heating, Refrigeration, and Air-Conditioning Engineers.

80. ANSI/ASHRAE, *Standard Methods of Measurement of Flow of Liquids in Pipes Using Orifice Flowmeters*. 1989, American Society of Heating, Refrigeration, Air-Conditioning Engineers.

81. ANSI/ASHRAE, *Method of Test for Measurement of Flow of Gas*. 2006, American Society of Heating, Refrigeration, and AIr-Conditioning Engineers.

82. ANSI/ASHRAE, *Standard Methods for Volatile-Refrigerant Mass Flow Measurements Using Calorimeters*. 2011, American Society of Heating, Refrigeration, and Air-Conditioning Engineers.

83. ANSI/ASHRAE, *Standard Method for Measuring the Proportion of Lubricant in Liquid Refrigerant*. 2015, American Society of Heating, Refrigeration, and Air-Conditioning Engineers.

84. Cremaschi, L. and P.P. Perez Paez, *Development of a Load-based Testing of Light Commercial Unitary HVAC*. 2016, American Society of Heating, Refrigeration, and Air-Conditioning Engineers: Oklahoma State University.

85. Cremaschi, L. and P.P. Perez Paez, *Experimental feasibility study of a new load-based method of testing for light commercial unitary heating, ventilation, and air conditioning (ASHRAE RP-1608)*. Science and Technology for the Built Environment, 2017. **23**(7): p. 1178-1188.

86. CSA, *Performance Standard for Split-System and Single-Packaged Air Conditioners and Heat Pumps*. 2014, Canadian Standards Association.

87. Henderson, H.I. and K. Rengararajan, *A model to predict the latent capacity of air conditioners and heat pumps at part-load conditions with constant fan operation.* ASHRAE Transactions, 1996. **102**(1).

APPENDIX A. LOAD-BASED TESTING METHOD FOR AIR CONDITIONING EQUIPMENT

A.1 Background and Motivation

Evaluating the performance of direct expansion (DX) air conditioning equipment is an important step in the development of novel technologies for improving efficiency and reducing energy consumption in today's and tomorrow's buildings. Qualifying and quantifying the performance benefits of new technologies requires identifying a market need or inefficiency, developing a solution that meets design criteria, and verifying the solution meets expected performance. During the solution development and verification process, an engineer has many tools at his or her disposal – from detailed computation models and energy simulation programs to component level test apparatuses to full-scale system level equipment testing environments. However, as air conditioning equipment becomes more advanced, these tools must also advance so that they may accurately and fairly measure or predict the performance of these new systems.

Two energy-saving technologies that have been introduced in many new DX air conditioning systems relatively recently are electronically commutated motors (ECM) for fans and inverter-driven compressors. These technologies can significantly reduce seasonal energy consumption of residential and commercial air conditioning systems by continuously modulating system cooling or heating capacity to balance building loads. This reduces equipment cycling and helps improve part-load efficiencies of equipment. To obtain maximum performance benefits, these systems are coupled with more advanced controls in comparison to those traditionally found in fixed-speed systems. This coupling between variable-speed motors and compressors with more advanced controls is integral to achieving performance improvements of the system.

For this reason, it is interesting to note that existing commercial and residential DX equipment testing standards and methodologies require the embedded control installed on variable-speed equipment to be disabled during testing [39, 77, 78]. Instead, the fan and compressor speeds of variable-speed systems are usually fixed to prescribed levels and steady-state performance of the system is measured. Rather than letting the embedded controls of the variable-speed systems modulate their delivered capacity to maintain space conditions, the performance of these systems is measured in the same manner as their fixed-speed counterparts. Because of this, performance of these systems is measured and rated based on operation that is very different than what would be observed in the field.

This paper presents a new load-based approach for testing DX air conditioning equipment that allows embedded controls to respond normally to an environment that simulates a typical building. The method is valid for different types of systems, including fixed-speed, variable-speed, and ductless, providing a consistent approach for comparing their performance in a realistic manner. The methodology was applied to the testing of packaged rooftop unit (RTU) air conditioners for three different case studies: fixed-speed on/off control, two-speed control, and variable-speed control. In order illustrate the benefits of load-based testing, the same RTU control modes were also tested using existing equipment ratings procedures and existing seasonal performance ratings were determined based on data obtained from the two different testing methodologies applied to the three control modes.

A.2 Review of Existing Methods for Characterizing Performance of RTUs

ASHRAE has worked to develop and update psychrometric chamber testing methodology standards for DX equipment for determining the cooling and heating capacities and efficiencies in a systematic and reproducible manner [57, 79]. Together with ASHRAE Standard 41 series, a

standardized way to install, measure, and operate different types of DX equipment in psychrometric chambers has been created [41-43, 58-60, 80-83].

To measure and rate the performance of direct expansion air conditioning equipment, a consistent set of tests are applied to equipment so that performance of different systems can be compared at common rating points. Predominately for small- to mid-size DX unitary air conditioning equipment (often called rooftop units, RTUs), a well-known and accepted testing and ratings standard (AHRI 210/240) is used to measure the expected total cooling capacity, sensible cooling capacity, and total power consumption of a system at several operating conditions [39]. Depending on the functionality of the system (e.g. whether the system has multiple cooling stages, the type of fan motor installed, etc.), a different set of required and optional tests are applied.

Predominately, three figures of merit have been used to rate equipment performance: EER, SEER, and IEER [39, 77, 78]. The EER, or energy efficiency ratio, is the ratio between the cooling capacity delivered by a system in kBtu/hr to total power consumed by the system in kW at a specific operating condition. The seasonal energy efficiency ratio, SEER, is a binned based calculation relating the cooling capacity delivered and power consumed at different ambient conditions with an adjustment for part-load degradation that occurs due to cycling at low loads. The part-load degradation factor can be determined from a specific cycling test or a default value can be employed. The integrated energy efficiency ratio, IEER, is essentially a weighted average EER at different operating conditions.

Cremaschi and Paez performed an experimental study on the development and feasibility of a load-based methodology for testing light commercial air conditioning systems [84, 85]. In this work, field and simulation data were collected to determine typical sensible and latent building

loads for rooftop units with economizers. Using a simple balance point representation of these loads as function of outdoor temperature and occupancy, an RTU with economizer was tested in psychrometric chambers under different typical loads. This was accomplished by disabling the traditional psychrometric chamber control system in the indoor space and instead fixing the amount of sensible and latent heat added to the space. By doing this, the temperature and humidity within the space responded dynamically to the difference between the load in the indoor room and the capacity delivered by the RTU. Additionally, the methodology made it possible to measure the performance impacts of outdoor air economizers under different ambient conditions and ventilation requirements. Despite implementing a load-based control algorithm in the psychrometric chambers, Cremaschi and Paez found issues with repeatability and reproducibility of the testing results noting variations in actual measured loads. Additionally, they posited that it is unlikely that testing results would be reproducible at different laboratories using their proposed methodology [84]. Finally, an additional short-coming of this work was that native control of the RTU was not used during the testing methodology – but rather lab equipment was used to "mimic" the normal controller performance of the RTU [85].

The methodology developed in the current paper differs from the approach of Cremaschi et al. in that it utilizes the existing psychrometric chamber controls and adjusts the temperature and humidity to emulate the response of a building [84, 85]. Some of this development effort has occurred in collaboration with the development of an update to a Canadian Standards Association (CSA) standard for measuring the performance of heat pumps and air conditioners [86].

A.3 Methodology and Implementation in Psychrometric Test Chamber Controller

The test methodology involves using a virtual building model to adjust the indoor room temperature and humidity setpoints for the psychrometric conditioning system in a manner that mimics the response of a real building to the RTU operation. The temperature and humidity responses of the virtual building are calculated based on simple load models along with measurements of the test unit sensible and latent cooling rates. With this approach, the test unit thermostat naturally responds to the dynamic temperature variation to control the equipment capacity in response to a deviation from its setpoint. The following sections provide details of the load models and setpoint updating schemes along with a description of test setup.

A.3.1 Sensible Load Control Derivation

The virtual building sensible cooling load, $\dot{Q}_{load,s}$, is the sum of an internal load component, $\dot{Q}_{int,s}$, and an external load component, $\dot{Q}_{ext,s}$,

$$\dot{Q}_{load,s} = \dot{Q}_{int,s} + \dot{Q}_{ext,s} \tag{A.1}$$

In this representation, the internal gain component can be thought of as the heat gain within the building due to the building occupants, electrical plug-loads, and other internal heat sources. The external gains are modeled as heat transfer driven by temperature differences between the conditioned space and the surrounding environment using an overall heat conductance of the building, UA_s,

$$\dot{Q}_{load,s} = \dot{Q}_{int,s} + UA_s \cdot \left(T_{oa} - T_{za}\right) \tag{A.2}$$

where T_{oa} is the dry-bulb temperature of the ambient air surrounding the building and T_{za} is the dry-bulb temperature of the conditioned space. The conductance can be chosen to consider both conductance through the envelope and the effects of ventilation and infiltration.

The transient response of the temperature within the space is simulated using a simple energy balance with a lumped capacitance,

$$C_s \frac{dT_{za}}{dt} = \dot{Q}_{load,s} - \dot{Q}_{cool,s} \tag{A.3}$$

where C_s is an effective thermal capacitance of the conditioned space and $\dot{Q}_{cool,s}$ is the sensible cooling capacity delivered by the air conditioning equipment. Equation (A.3) shows that for larger sensible load imbalances, the rate of change in indoor room temperature is increased. To simulate the response of a typical building using psychrometric chambers, the temperature of the indoor room where the equipment thermostat is located is controlled to match this response.

In order to obtain a simple updating scheme for the indoor room temperature setpoint, a first-order forward difference formula is applied to estimate the derivative term in Equation (A.3),

$$C_s \frac{T_{za}(t+\Delta t) - T_{za}(t)}{\Delta t} = \dot{Q}_{load,s}(t) - \dot{Q}_{cool,s}(t) \tag{A.4}$$

where $T_{za}(t)$ is the current indoor temperature setpoint at time t and $T_{za}(t+\Delta t)$ is the indoor temperature setpoint at time $t+\Delta t$. A recursive updating equation for the indoor temperature setpoint is obtained by rearranging Equation (A.4) and substituting Equation (A.2) for the sensible load on the space,

$$T_{za}(t+\Delta t) = T_{za}(t) + \frac{\Delta t}{C_s}\left[\dot{Q}_{int,s}(t) + UA_s\left(T_{oa}(t) - T_{za}(t)\right) - \dot{Q}_{cool,s}(t)\right]. \tag{A.5}$$

When implementing this setpoint updating scheme, real-time measurements of the air conditioning equipment sensible cooling capacity, $\dot{Q}_{cool,s}(t)$, are employed. If the psychrometric chamber conditioning system can respond quickly to these setpoint changes, then the indoor room temperature variation provides a realistic dynamic response to the interaction of the test equipment and a typical building load that would be served by that equipment.

It is necessary to determine reasonable values for the sensible internal gain $\dot{Q}_{int,s}$, the sensible overall heat transfer coefficient, UA_s, and the thermal capacitance, C_s, used in the virtual building model that are scaled according to the equipment capacity. A reasonable method for specifying $\dot{Q}_{int,s}$ and UA_s for a given test unit involves the use of the balance point load calculation methodology [61] and a rated sensible cooling capacity of the test unit. In this methodology, the sensible equipment loads at any given outdoor temperature, T_{oa}, can be expressed as,

$$\dot{Q}_{load,s} = SL_D \cdot \frac{T_{oa} - T_{bal}}{T_{OD} - T_{bal,D}} \tag{A.6}$$

where SL_D is the sensible building load at the design point operating condition, T_{OD} is the design point outdoor temperature, $T_{bal,D}$ is the design balance point outdoor temperature (where $\dot{Q}_{load,s} = 0$), and T_{bal} is given by

$$T_{bal} = T_{bal,D} + \left(T_{za} - T_{ID}\right) \tag{A.7}$$

where T_{ID} is the design point indoor temperature and T_{za} is the actual indoor temperature of the space. The sensible design load is then related to the sensible cooling capacity of the test equipment according to

$$SL_D = \frac{SC_D}{1 + f_{os}} \tag{A.8}$$

where SC_D is the equipment sensible cooling capacity at the design point operating condition and f_{os} is a constant used to oversize the system. It can be shown that Equations (A.2) and (A.6) are equivalent when

$$\dot{Q}_{int,s} = \frac{SC_D}{1+f_{os}} \cdot \frac{T_{ID} - T_{bal,D}}{T_{OD} - T_{bal,D}} \tag{A.9}$$

and

$$UA_s = \frac{SC_D}{1+f_{os}} \cdot \frac{1}{T_{OD} - T_{bal,D}}. \tag{A.10}$$

Equations (A.9) and (A.10) provide a straightforward approach for determining values of $\dot{Q}_{int,s}$ and UA_s from the test unit sensible capacity determined from a ratings test and other parameters that are easy to specify and are independent of equipment size. Typical design values for indoor, outdoor, and balance point temperatures are 23.9 °C (75 °F), 35.0 °C (95 °F), and 18.3 °C (65 °F), respectively, whereas a typical oversizing factor is 0.2.

Determining a thermal capacitance for the load model, C_s, is less straightforward due to the dependence on building construction materials and geometry. *Heavy* buildings, or buildings with large thermal mass, will respond very differently than *light* construction, especially with respect to the cycling dynamics of the equipment at low cooling loads. It is difficult to define a typical building capacitance because building materials and geometry depend on several factors such as climatic region, building age, application type, etc. In addition, the use of a single capacitance to represent building dynamics is a major simplification and an effective capacitance that provides reasonable dynamic behavior cannot be calculated from a physical description. As a result, it is necessary to rely on empirical results in defining an approach for specifying a scalable effective building capacity. In this study, a method for specifying C_s was derived from field data and a model describing typical cycling frequencies developed by Henderson and Rengarajan [87]. In their model, the number of on-off cycles of a fixed-speed system is expressed as a function of the part-load ratio, PLR, and the thermostat dead band, ΔT_{DB},

$$N = \frac{\dot{Q}_{cool,s}}{2\Delta T_{DB} C_s} \cdot PLR \cdot (1 - PLR) \tag{A.11}$$

where $\dot{Q}_{cool,s}$ is the sensible cooling capacity of the equipment neglecting cyclic degradation losses, and C_s is the effective thermal capacitance of the space. Based on this model, the maximum number of cycles per hour will occur when the part-load ratio is 50%,

$$N_{max} = \frac{\dot{Q}_{cool,s}}{2\Delta T_{DB} C_s} \cdot \frac{1}{2} \cdot \left(1 - \frac{1}{2}\right) = \frac{\dot{Q}_{cool,s}}{8\Delta T_{DB} C_s}. \tag{A.12}$$

Henderson and Rengarajan used field data and determined that the maximum number of cycles per hour was approximately 3 ($N_{max} = 3$) [87]. Assuming a thermostat dead band $\Delta T_{DB} = 1\,°\text{F}$, then the effective thermal capacitance can be estimated as

$$C_s = \frac{SC_D}{24} \tag{A.13}$$

where the sensible cooling capacity has been replaced by the test unit design cooling capacity, SC_D, employed in Equations (A.9) and (A.10). This provides a scalable approach for specifying C_s in the virtual building load model.

It should be noted that more detailed dynamic sensible load models could be employed to better represent the transient response of a typical building. However, these models come at the cost of increased complexity and require more input parameters that would be difficult to specify in a general and scalable manner. The approach specified in this section is easy to apply and provides reasonable dynamic response for testing equipment with their integrated controls. In addition, the three parameters of the model ($\dot{Q}_{int,s}$, UA_s, C_s) are determined in a systematic manner so that the approach can be readily implemented in different psychrometric chambers with the potential for obtaining reproducible results in different testing laboratories.

A.3.2 Latent Load Control Derivation

It is also important to model the dynamic response of humidity within the conditioned space to properly capture the interaction with equipment controls. An analogous form of the dynamic sensible load model for the dynamic response of the indoor humidity ratio, ω_{za}, is given by

$$C_l \frac{d\omega_{za}}{dt} = \frac{1}{h_{fg}}\left(\dot{Q}_{load,l} - \dot{Q}_{cool,l}\right) \tag{A.14}$$

where $\dot{Q}_{load,l}$ is the latent load in the conditioned space, $\dot{Q}_{cool,l}$ is latent cooling capacity delivered by the air conditioning equipment, C_l is an effective moisture capacitance of the space, and h_{fg} is the enthalpy of vaporization of water. The moisture capacitance has the units of mass (kg) and is somewhat greater than the mass of air in the building due to interactions with moisture absorber materials (e.g., carpet, soft furnishings, paper, etc.) The continuous time derivative of Equation (A.14) is approximated using a first-order forward difference formula and rearranged to give a recursive relation for updating an indoor room humidity ratio setpoint,

$$\omega_{za}\left(t+\Delta t\right) = \omega_{za}\left(t\right) + \frac{\Delta t}{C_l \cdot h_{fg}}\left[\dot{Q}_{load,l}\left(t\right) - \dot{Q}_{cool,l}\left(t\right)\right] \tag{A.15}$$

This latent load model can be expressed in a form that is analogous to the sensible dynamic model by representing the latent load as the sum of an internal gain component, $\dot{Q}_{int,l}$, and an infiltration gain component,

$$\dot{Q}_{load,l}\left(t\right) = \dot{Q}_{int,l} + UA_l \cdot \left(\omega_{oa}\left(t\right) - \omega_{za}\left(t\right)\right) \tag{A.16}$$

where UA_l is an overall "latent" transfer conductance accounting for latent gains driven by differences between indoor and outdoor humidity ratios, ω_{oa}. These gains may be due to ventilation and infiltration.

To complete the latent load analogy, a balance point methodology for latent gains is used to estimate the internal latent gain and the overall latent transfer conductance of the simulated building using design point data. In this representation, the latent load is calculated for a given outdoor humidity ratio, ω_{oa}, using

$$\dot{Q}_{load,l} = LL_D \cdot \frac{\omega_{oa} - \omega_{bal}}{\omega_{oa} - \omega_{bal,D}} \tag{A.17}$$

and

$$\omega_{bal} = \omega_{bal,D} + \left(\omega_{za} - \omega_{ID}\right) \tag{A.18}$$

where ω_{ID} is the design point indoor room humidity ratio, ω_{za} is the indoor room humidity ratio, and $\omega_{bal,D}$ is the design balance point humidity ratio given by

$$\omega_{bal,D} = \omega_{ID} + r_{int}\left(\omega_{OD} - \omega_{ID}\right) \tag{A.19}$$

where r_{int} is the fraction of the design point latent load due to internal latent gains. For many commercial buildings, this fraction may be small (e.g. 0.1 to 0.2), though it may be significant for some building applications (e.g. gymnasiums, auditoriums).

The design latent load is expressed in terms of the design sensible load and a design point sensible heat ratio (SHR_D),

$$LL_D = SL_D \cdot \left(\frac{1}{SHR_D} - 1\right) \tag{A.20}$$

Finally, the internal moisture generation and overall latent heat transfer coefficient may be determined from design information by equating Equations (A.16) and (A.17),

$$\dot{Q}_{int,l} = LL_D \cdot \frac{\omega_{ID} - \omega_{bal,D}}{\omega_{OD} - \omega_{bal,D}} \tag{A.21}$$

and

$$UA_l = LL_D \cdot \frac{1}{\omega_{OD} - \omega_{bal,D}}. \tag{A.22}$$

As with the thermal capacitance, it is difficult to determine an analytical and general approach for estimating a moisture capacitance of a typical building due to the distributed nature of the building materials and a dependence on many building details that would be difficult to specify. As a result, a simple empirical approach is used to specify an effective moisture capacitance by scaling the mass of air, m_{air}, with a multiplier, n_l, that accounts for additional moisture absorbing materials,

$$C_l = n_l \cdot m_{air} \tag{A.23}$$

Furthermore, the mass of air in the virtual building is scaled with the rated cooling capacity of the unit, $\dot{Q}_{rated}$, such that

$$C_l = n_l \cdot k_{scale} \cdot \dot{Q}_{rated} \tag{A.24}$$

In this study, the scaling factor, k_{scale}, was determined assuming a typical rule of thumb for commercial buildings of 1 ton of cooling capacity per 400 ft^2 of floor space (0.095 kW/m^2), a ceiling height of 10 feet (3.04 m), and density of air at standard conditions resulting in a scaling factor of 300 lbm/ton (38.7 kg/kW). In addition, the moisture multiplier, n_l, was assumed to be 5.

A.3.3 Description of Psychrometric Test Facility

The dynamic sensible and latent load control algorithms were implemented in the control systems of a psychrometric chamber test facility. The controller used to control the outdoor room was not modified since tests were conducted at constant ambient conditions. The physical

layout of the rooms, along with how the variable-speed RTU was installed in the chambers is shown in Figure A.0.1.

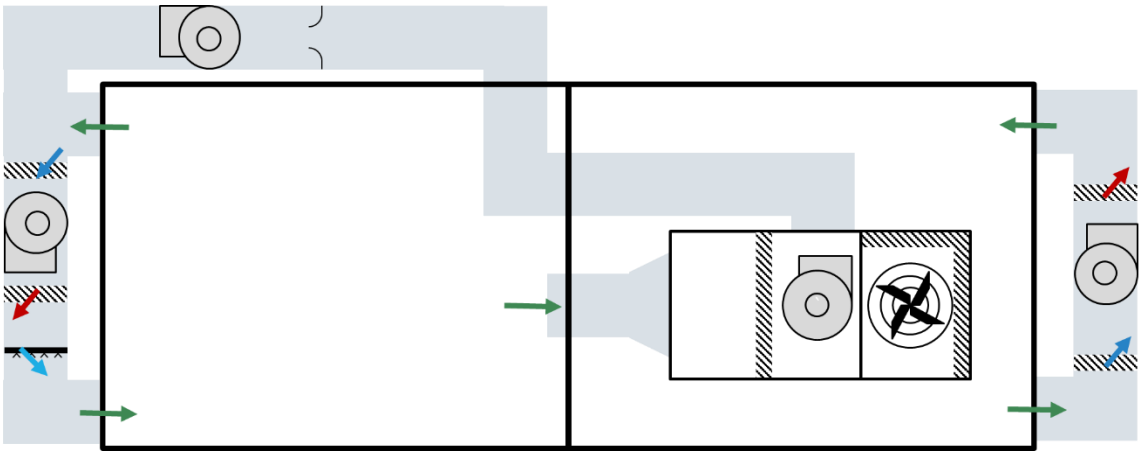

Figure A.0.1. Psychrometric chamber setup for load-based RTU performance testing.

Each room of the psychrometric test facility is conditioned using a variable capacity chiller and electrical reheat that maintain the indoor and outdoor space temperatures at specified setpoints and remove moisture. Additionally, each room uses steam humidifiers to maintain humidity setpoints in each room. Before entering the room re-conditioning air handler, the supply air flow from the RTU passes through a nozzle box air flow measurement station and a booster fan that can control the static pressure at the outlet of the RTU.

Laboratory-grade sensor measurements were used to determine the total and sensible cooling capacity of the system [41-43, 58, 60, 82]. On the airside, thermocouple grids were used at the inlet and outlet of the RTU in accordance with standard measurement recommendations. Humidity was measured using chilled mirror dew point hygrometers and relative humidity sensor probes at the inlet and outlet of the RTU. For verification, total capacity of the system was determined using refrigerant-side measurements as well. The mass flow rate of refrigerant through the system was determined using a Coriolis mass flow meter. To determine system

efficiency, individual power measurements of the compressor, indoor fan, and outdoor fan were used. Additionally, a precision power analyzer was used to measure the total power, voltage, and current delivered to the system during testing.

A.4 Application of Test Methodology to Fixed-Speed and Variable-Speed RTUs

The load-based testing methodology was used to test the performance of two RTUs to compare and characterize the part-load performance of fixed-speed and variable-speed systems. For this comparison, the two 5-ton RTUs tested were virtually identical apart from the type of compressor and control strategy used. The first system was installed with a two-stage, fixed-speed scroll compressor that was controlled using a bang-bang thermostat control strategy. Testing was carried out for this system with both single-stage and two-stage thermostat logic to generate results representative of both single and two-stage equipment. The second RTU has a VFD driven, variable-speed scroll compressor which was controlled to maintain a constant discharge air temperature setpoint. Both systems used the same supply air blower assembly but with different control logic. When the fixed-speed system operated in low-stage, the supply air blower supplied approximately 1200 CFM of air flow. In second-stage mode, the blower supplied approximately 1800 CFM. When the compressor cycled off (for either single-stage or two-stage operation), the air flow was approximately 900 CFM. For the variable-speed RTU, the supply air blower is continuously modulated between a minimum and maximum speed to control a space air temperature in response to a setpoint. In addition, the speed of the compressor is continuously modulated between a minimum and maximum speed to maintain a supply air temperature setpoint. The control strategy of the variable-speed RTU enables nearly continuous capacity modulation of the system under moderate to high cooling loads, though the system

cycles on and off during low cooling load conditions when the compressor and fan operate at their lowest control limits.

The load-based testing parameters used to fix the loads and ambient conditions are listed in Table A.. Ambient conditions were varied between 18.89 °C (66 °F) to 42.22 °C (108 °F) and the design indoor setpoint was fixed at 23.89 °C (75 °F) for all tests. The sensible heat ratio and oversizing factor were set to 0.90 and 0.10 respectively for all tests conducted. Using the sensible heat ratio and oversizing factor, the loads applied to the unit at each ambient condition were determined based on the full-load capacity at the AHRI A_2 test condition [39]. Using an oversizing factor of 0.10 to determine the virtual building load line, the systems lacked sufficient capacity to maintain the load at the highest ambient condition tested. As a result, the indoor temperature for these tests increased significantly until the load in the space matched the cooling capacity delivered by the systems. To maintain consistency between systems at these test conditions, the tests were performed a second time with the indoor and outdoor temperature and humidity held constant at the design conditions while the systems operated at full load.

Table A.1. Load-based testing conditions used for fixed-speed and variable-speed RTU psychrometric chamber testing.

$\mathbf{T_{oa}}$, °C (°F)	$\mathbf{T_{ID}}$, °C (°F)	$\mathbf{SHR_D}$, -	$\mathbf{f_{os}}$, -	**SL**, kW (kBTU/hr)
18.89 (66.0)	23.89 (75.0)	0.90	0.10	0.5 (1.7)
23.89 (75.0)	23.89 (75.0)	0.90	0.10	5.0 (17.1)
27.78 (82.0)	23.89 (75.0)	0.90	0.10	8.5 (29.0)
35.00 (95.0)	23.89 (75.0)	0.90	0.10	15.0 (51.2)
42.22 (108.0)	23.89 (75.0)	0.90	0.10	21.5 (73.4)

The parameters used in the virtual building model throughout experimental testing are shown in Table A. Sensible and latent design performance was determined based on rated system capacity tested at the AHRI design conditions. For this testing, the design load SHR was fixed at 0.90. Because of this, measurements of the outdoor room (or virtual building outdoor humidity

ratio) was not required for the tests. The assumption surrounding a constant SHR may not be accurate for some climatic regions where outdoor air humidity may be correlated with time or year. However, this simplification was made for this work and more complex load models and application types have been left for future work.

Table A.2. Parameters used for sensible and moisture load in virtual building load model during experimental testing.

Sensible Load Model		**Moisture Load Model**	
Parameter	*Value*	*Parameter*	*Value*
UA_s	1.23 kW/°C	UA_l	0.14 kW
$\dot{Q}_{int,s}$	6.69 kW	$\dot{Q}_{int,l}$	0.74 kW
C_s	0.63 kJ/°C	C_l	3300 kg
		k_{scale}	38.7 kg/kW
		n_l	5.0

It should also be noted that the indoor chamber test conditions used for steady-state testing and load-based testing were significantly different than setpoints used in existing equipment rating standards [39]. In the standard, the indoor chamber target drybulb and wetbulb temperatures are 26.7 °C (80 °F) and 19.4 °C (67 °F) respectively. For the testing performed in this work, the target indoor chamber setpoints were 23.9 °C (75 °F) drybulb and 16.1 °C (61 °F). These target setpoints resulted in a cooler more dry indoor condition than those traditionally controlled in equipment standards and lead to lower measured system efficiencies and cooling capacities due to significantly less latent removal.

A.4.1 Fixed-Speed Control Description and Example Results

Single-stage thermostat control logic, shown in Figure A.0.2, was used to test the fixed-speed RTU under ON/OFF control. When the temperature in the indoor room increased to $HB =$ 0.55 °C (1.0 °F) above the room setpoint $SP = 23.9$ °C (75 °F), the compressor was operated

under full load and the indoor fan speed was operated at 100%. When the RTU cooling capacity was greater than the sensible space load and the temperature decreased to $HB = 0.55$ °C (1.0 °F) less than the setpoint, the compressor was cycled off and the indoor fan speed was set to low (50%).

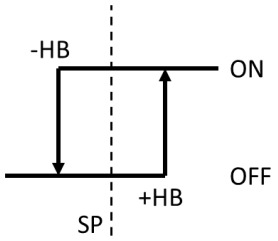

Figure A.0.2. ON/OFF thermostat control logic used during load-based fixed-speed RTU testing.

Single-stage thermostat control logic has been traditionally used to control air conditioning equipment and remains the predominant control used in small commercial building cooling and heating equipment in the US. This logic and the performance of fixed-speed DX equipment becomes very inefficient at small part-loads. Using the load-based psychrometric chamber control methodology, different loads and ambient conditions were controlled while the RTU maintained the indoor setpoint. Example results collected from the fixed-speed RTU operated using ON/OFF control with an applied sensible load equal to approximately 33% of its full-load sensible capacity are shown in Figure A.0.3.

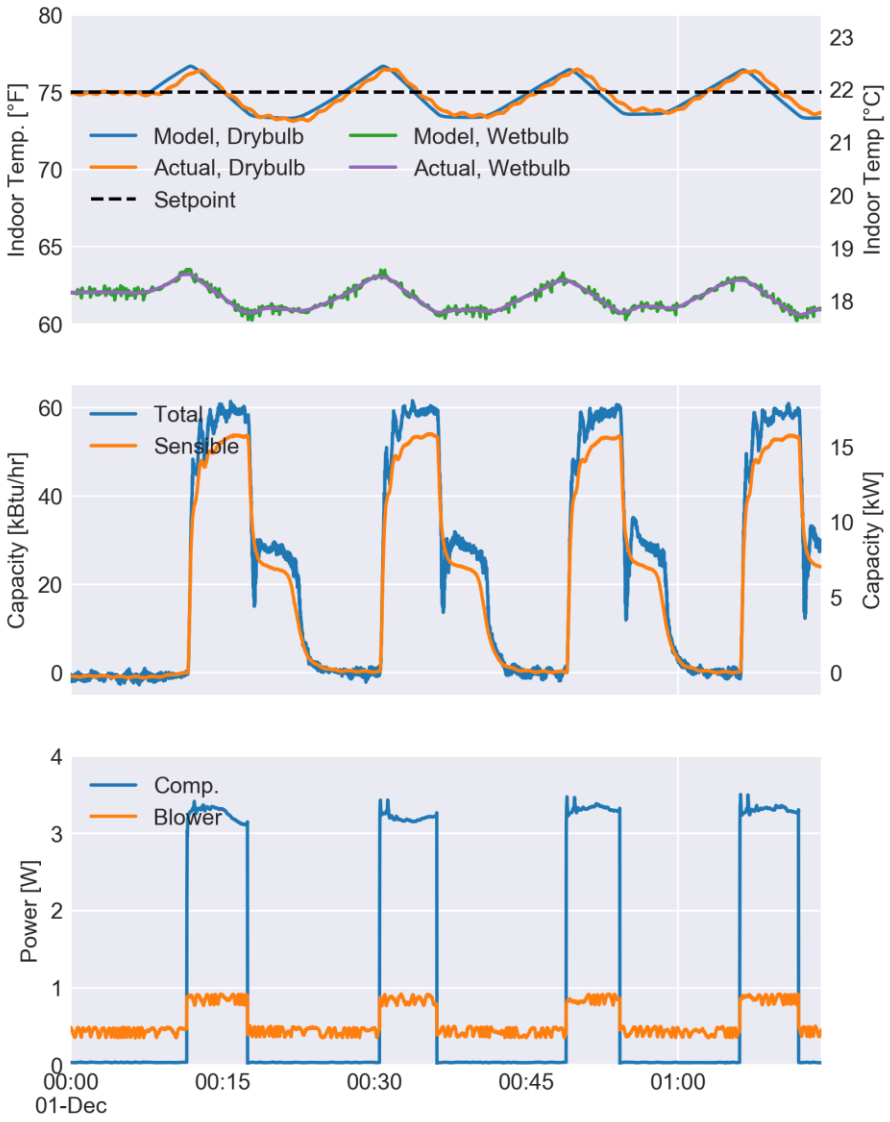

Figure A.0.3. Example load-based test results collected from fixed-speed RTU using single-stage thermostat control with an applied sensible load equal to approximately 33% of full-load sensible capacity.

The indoor temperature plot in Figure A.0.3 demonstrates that the psychrometric chamber control system can closely match the dynamic virtual building load model. Despite a small lag between the actual indoor temperature and the building load model, differences are within 0.06 °C (0.2 °F) throughout the example. The temperature plot also shows the typical response of the indoor space when controlled using a single-stage thermostat. Due to minimum on-times and off-times, the room is often overcooled, and cooling is often delayed. The cooling capacity

and power consumption plots show the natural performance of the system under the applied load without the need for prescribing on-off cycle times. It is important to note that the indoor temperatures shown in Figure A.0.3 were measured using thermocouples, which respond more quickly than a thermostat sensor. It is likely that a thermostat display temperature would not show quite the same range of temperature variation as exists in the space.

The modeled and actual latent responses of the room are also shown in Figure A.0.3 in terms of the indoor room wet bulb. While the RTU cycles on and off, the expected behavior is observed: the indoor room wet bulb decreases while the system provides mechanical cooling and increases while the system is off. Close agreement between the target wet bulb temperature from the virtual building model and the measured wet bulb temperature in the psychrometric chamber can be observed. This indicates that the dynamics of the psychrometric chamber control systems were able to simulate the virtual building model as desired.

It is also interesting to point out that the system delivered significant cooling capacity after the compressor was turned off during each cycle. For the most part this can be attributed to an evaporative cooling effect where water condensed on the surface of the cooling coil re-evaporates. Henderson et al. have researched this phenomenon extensively for fixed-speed air conditioners [87]. To a lesser extent, additional sensible cooling capacity may be attributed to the thermal mass of the cooling coil and refrigerant in the path of the air stream after the compressor is cycled off.

After testing the fixed-speed system using single-stage thermostat control, the system was modified to use two-stage thermostat logic, shown in Figure A.0.4. In this controller, initially the low cooling stage is initiated when temperature increases $HB_1 = 0.55$ °C (1.0 °F) above the setpoint. If further cooling capacity is required, the RTU is operated under full load when the

temperature rises $HB_2 = 0.83$ °C (1.5 °F) above the setpoint. The system uses two fan speeds as well depending on which control mode the system is operated in.

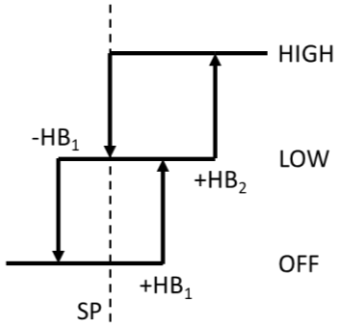

Figure A.0.4 Two-stage thermostat control logic used during load-based fixed-speed RTU testing.

Identical ambient conditions and loads were applied to the two-stage RTU using the load-based psychrometric chamber control methodology. Example results collected from the fixed-speed RTU operated using the two-stage thermostat with an applied sensible load equal to approximately 33% of its full-load sensible capacity are shown in Figure A.0.5. For this load level, the compressor was cycling between off and low stage. In comparison with the single-stage control results, indoor temperature control is improved slightly due to the closer match between equipment sensible cooling capacity and the load when operated using low-stage.

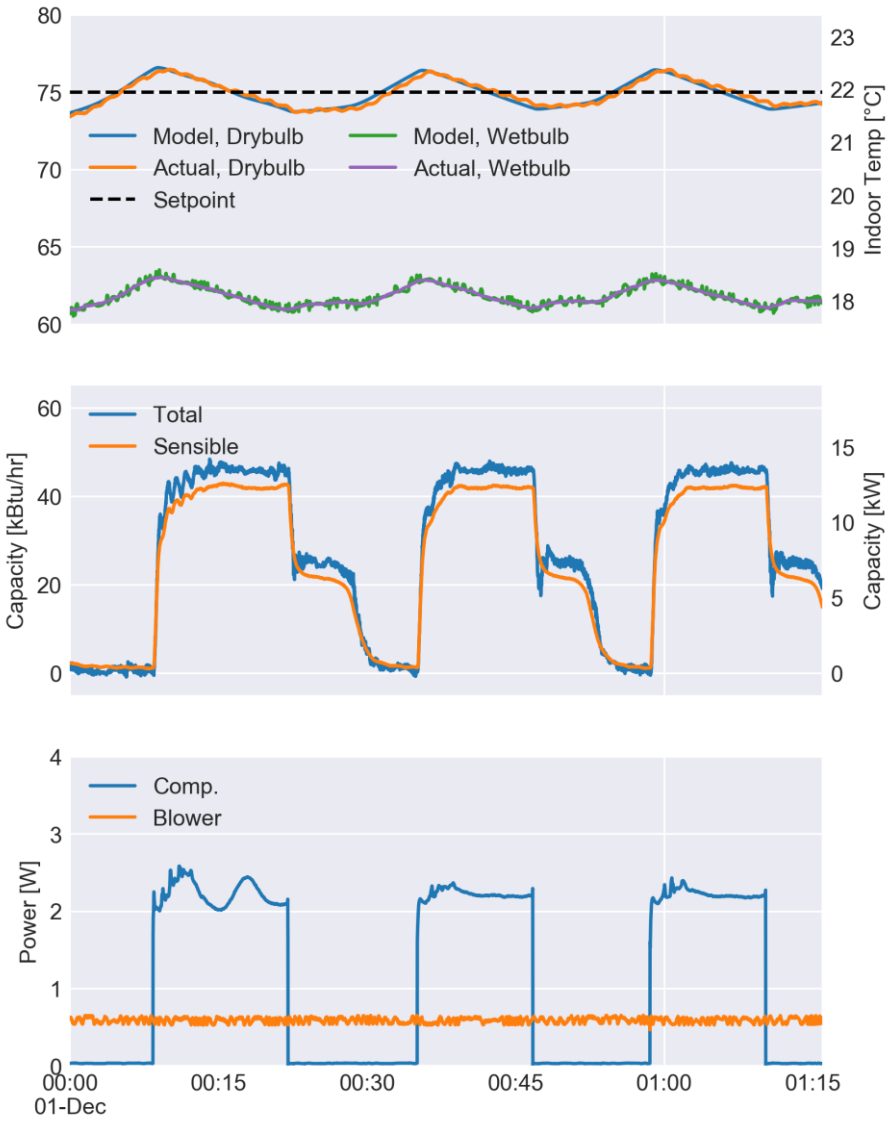

Figure A.0.5. Example load-based test results collected from fixed-speed RTU using two-stage thermostat control with an applied sensible load equal to approximately 33% of full-load sensible capacity.

For these tests, the psychrometric chamber facility was able to follow the indoor room setpoints provided by the virtual building model within approximately 0.17 °C (0.3 °F) after initial transients and room dynamics were controlled. Relatively speaking, these errors were on par with the single-stage test results and affirmed that the psychrometric chamber cooling and heating system was more than capable of simulating the virtual building load dynamics. From a test reproducibility and repeatability standpoint, this is important. It is also interesting to point

out that the performance of the fixed-speed equipment was similar from one ON/OFF cycle to the next. This suggests that within a few cycles, average behavior and performance may be estimated for a given unit and significant test time may not be necessary. It should be noted that this may not be true of all systems, especially those that may implement adaptive control technologies that learn and evolve over time and so repeatability should be studied further.

A.4.2 Variable-Speed Control Description and Example Results

The same load-based testing methodology was used to test the variable-speed RTU. The discharge air temperature in this system is controlled by modulating the compressor speed using a PI controller. While the discharge air setpoint can be adjusted for this unit, all testing performed was done at the default setpoint defined by the manufacturer: 12.8 °C (55 °F). Concurrently, a PI controller was used to control the indoor temperature of the space by adjusting the indoor blower speed.

Example load-based testing results collected from the variable-speed RTU are shown in Figure A.0.6 for the same conditions and loads used for the previous fixed-speed results. In comparison to the fixed-speed results, the improvement in indoor temperature control is noticeable. Fluctuations in indoor humidity also do not occur after the RTU reaches steady-state. Additionally, both compressor and indoor blower power consumption are significantly lower than the fixed-speed systems when they were providing mechanical cooling. Also, the psychrometric chamber control system was able to follow both the dry bulb and wet bulb setpoints determined by the virtual building load model very well. This is unsurprising since the variable-speed controller essentially matches the load, and thus temperature swings in the indoor chamber are much less than for the fixed-speed systems.

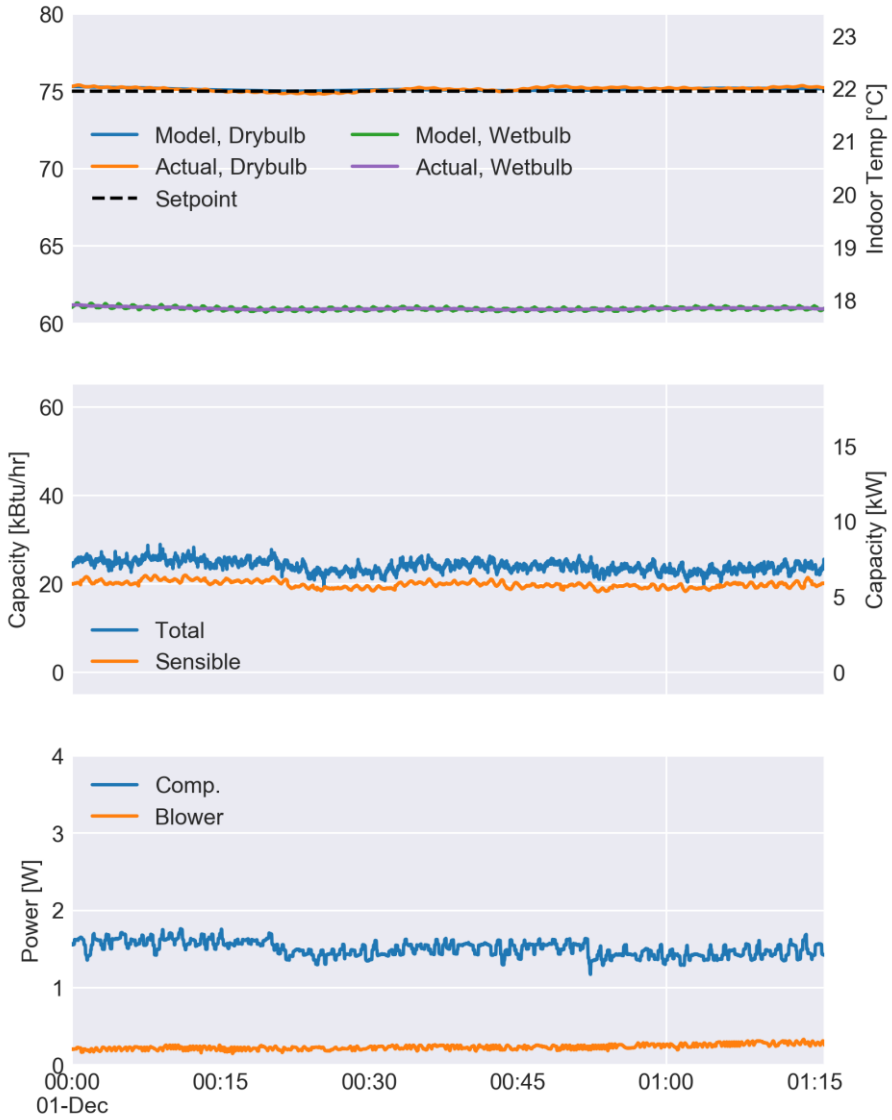

Figure A.0.6. Example load-based test results collected from variable-speed RTU using manufacturer supplied thermostat with an applied sensible load equal to approximately 33% of full-load sensible capacity.

A.4.3 Part-load Comparison of Fixed-Speed and Variable-Speed RTUs

Using the one-stage, fixed-speed RTU control mode, the efficiencies at different ambient conditions were determined from measurements obtained during the load-based testing. Since the unit was cycling on and off for these tests, the efficiencies are based on integrating the cooling provided and electrical consumption over a complete on/off cycle. For comparison purposes, the full-load performance of the RTU was measured using steady-state testing to

calculate system efficiencies as per the methodology described by AHRI 210/240 [39]. The optional dry-coil cyclic and steady-state tests were performed using the fixed-speed RTU to calculate the coefficient of degradation parameter. The combination of steady-state performance results, cycle degradation coefficients, and the load-based model were then used to estimate integrated equipment efficiency as a function of outdoor temperature to enable direct comparisons with load-based test results. A comparison of the load-based results (both measured and estimated from steady-state results) are shown in Figure A.0.7 along with the full-load steady-state efficiencies. Results based on both the cyclic degradation coefficient determined from the measurements and the default value provided by the AHRI 210/240 standard are included.

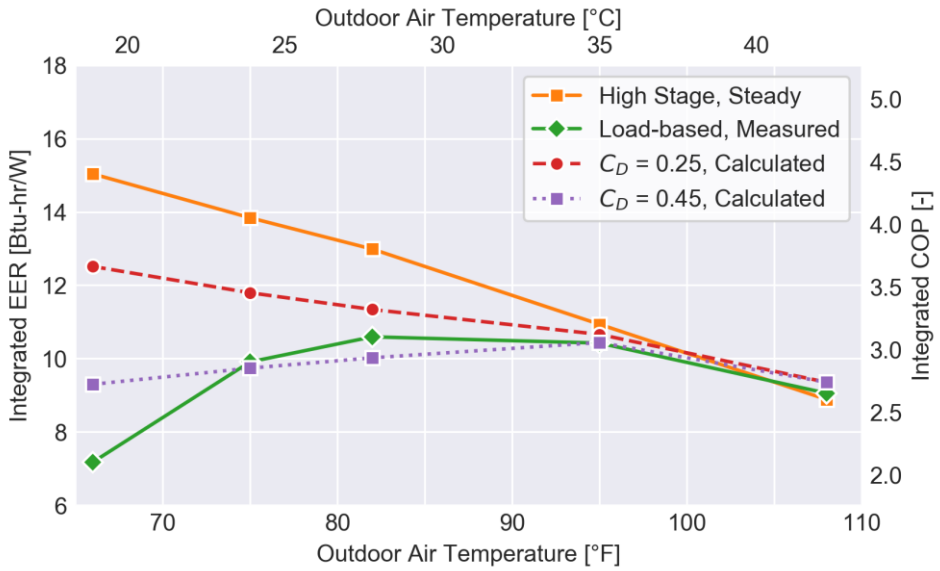

Figure A.0.7. Comparison of integrated efficiency of one-stage, fixed-speed RTU control measured using load-based testing and calculated using methodology described in AHRI Standard 210/240. The two values for the coefficient of degradation used: the default value C_D = 0.25 and the value measured using dry-coil cyclic testing C_D = 0.45.

It is clear from Figure A.0.7 that differences exist between the efficiencies determined from load-based testing and the results calculated based on the AHRI 210/240 methodology. These differences become larger at lower ambient temperatures, or equivalently at lower sensible loads.

The efficiency of the RTU from the load-based testing methodology significantly decreases at lower temperatures due to the more frequent ON/OFF cycles required to maintain the indoor air temperature. Similar behavior is seen for the results determined for the AHRI 210/240 methodology if the measured coefficient of degradation is utilized although the errors become significant at low loads. It is important to note the default value for the coefficient of degradation is too small for this RTU and does not adequately adjust the efficiency for cycling losses. Finally, the differences in efficiencies between the full-load and load-based testing results are dramatic at lower ambient temperatures.

Similar results were obtained for a two-stage RTU over the range of ambient conditions and equivalent building loads and are shown in Figure A.0.8. In this case, the differences between the efficiencies determined from load-based testing and calculated using the methodology described in AHRI Standard 210/240 for two-stage equipment are smaller than for the single-stage control especially when employing the coefficient of degradation determined from tests. Using the default value for the coefficient of degradation does not produce results that follow the same trends as the load-based testing methodology. The steady-state low-stage and high-stage efficiencies monotonically increase with decreasing ambient temperatures in sharp contrast to the load-based testing results where efficiency degrades at lower ambient temperatures and loads due to increase cycling losses.

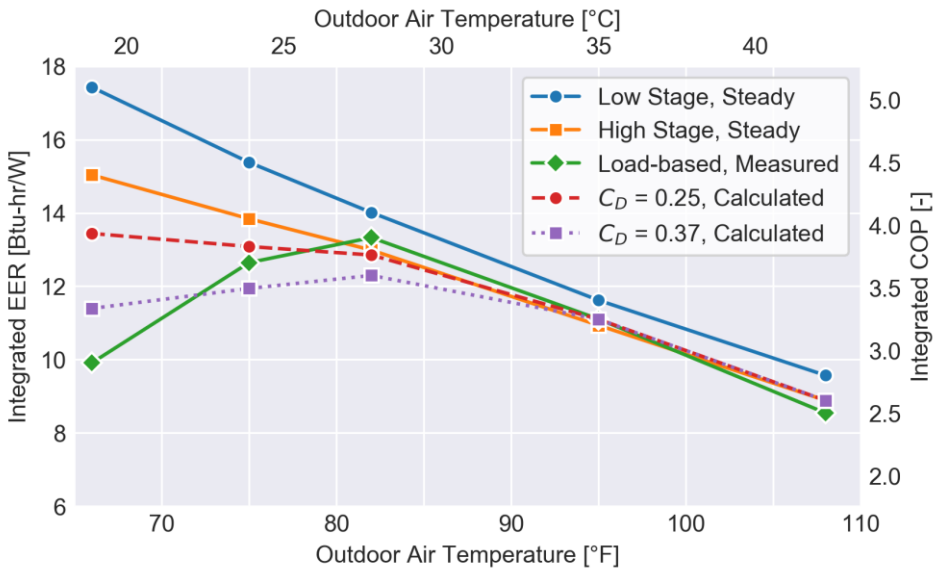

Figure A.0.8. Comparison of integrated efficiency of two-stage, fixed-speed RTU control measured using load-based testing and calculated using methodology described in AHRI Standard 210/240. The two values for the coefficient of degradation used: the default value $C_D = 0.25$ and the value measured using dry-coil cyclic testing $C_D = 0.37$.

Comparisons of integrated system efficiencies determined from load-based testing at different ambient conditions for the three different RTU control modes considered in this study are shown in Figure A.0.9. The results show that significant gains in system efficiencies at part-load are achievable with variable-speed compressors and fans. These results also illustrate the benefit of load-based testing: for nearly identical systems differing only by control modes, the efficiencies can be compared at identical loads and ambient conditions experimentally using psychrometric testing that simulates real-world response. In comparison to the results obtained using steady-state testing, the load-based test results inherently capture inefficiencies due to ON/OFF cycling, or any other losses due to suboptimal control or equipment operation.

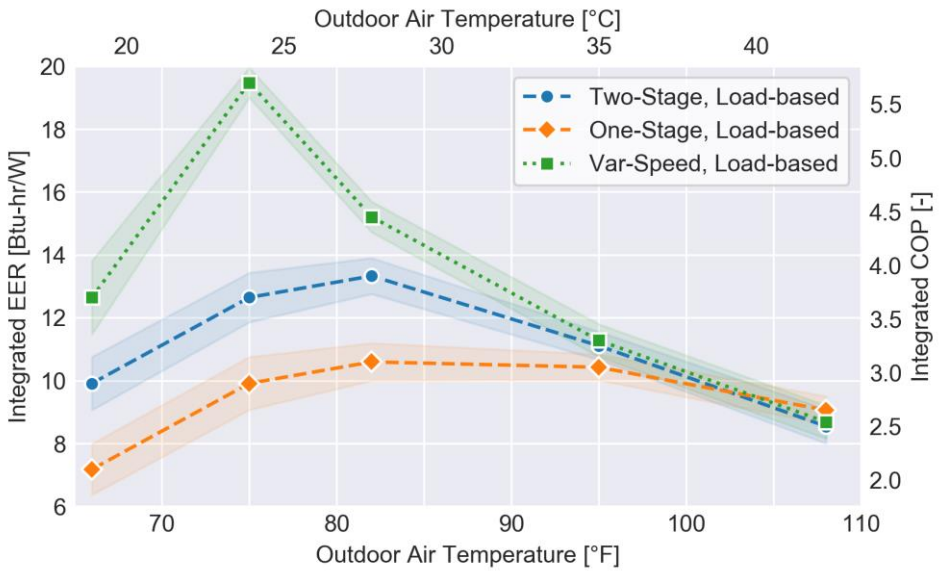

Figure A.0.9. Comparison of fixed-speed and variable-speed RTU efficiencies measured using load-based testing under equivalent virtual building loads. Transparent bands have been added corresponding to 95% uncertainty bounds based on statistical analysis of the experimental data.

Using the data collected from the psychrometric chamber tests, seasonal efficiencies were calculated using the methodology described in AHRI Standard 210/240 and are presented in Table A. The fraction bin hours used for these calculations were copied from the example calculations in AHRI Standard 210/240 and are shown in the Appendix. For load-based testing results, the cyclic degradation coefficient used in the calculation was replaced by measured integrated efficiencies from the psychrometric chamber tests at different operating conditions. For conditions not tested, linear interpolation was used. The results show that significant differences exist between the load-based testing results and the results calculated with cyclic degradation coefficients when default values are employed. The errors are particularly large for the one-stage control but are also significant for two-stage operation. In both cases, the use of the default value leads to overestimated seasonal efficiency for this unit. Determining the degradation coefficient from measurements gives much better results, but differences with the load-based testing results are still significant for the one-stage control. The seasonal efficiencies

for the variable-speed unit determined from load-based testing are significantly better than for the other two control modes. The load-based testing allows a direct comparison of efficiencies that are based on a consistent test method that includes interaction of the controls with the equipment.

Table A.3. Comparison of calculated seasonal efficiencies based on load-based testing and AHRI Standard 210/240 calculation methodology for different RTU control modes.

	Seasonal EER [Btu-hr/W]	Seasonal COP [-]	Percent Difference [%]
One-stage, Default C_D	11.7	3.43	+ 24.2
One-stage, Measured C_D	9.8	2.86	+ 3.6
One-stage, Load-based	9.4	2.76	-
Two-stage, Default C_D	12.9	3.77	+ 7.7
Two-stage, Measured C_D	11.8	3.45	– 1.4
Two-stage, Load-based	11.9	3.50	-
Variable-speed, Default C_D	14.0	4.10	– 9.7
Variable-speed, Measured C_D	14.5	4.24	– 6.5
Variable-speed, Load-based	15.5	4.54	-

It is important to note that the results in Table A seem to imply that the differences in seasonable efficiency between variable-speed, two-stage, and single-stage capacity control are larger when determined using data from load-based testing (15.5 compared to 11.9 and 9.4) than when determined using steady-state testing and a coefficient of degradation for cycling (14.5 compared to 11.8 and 9.8 with measured C_D and 14.0 compared to 12.9 and 11.7 with default $C_D = 0.25$). Current methodologies for calculating seasonal efficiencies may underpredict cyclic degradation inefficiencies of fixed-speed systems and overestimate cyclic losses of variable-speed systems. It is believed that load-based testing provides a more representative evaluation of the benefits of variable-speed equipment.

A.5 Discussion and Recommendations for Future Work

A psychrometric chamber control methodology derived from the dynamic sensible and latent energy balances of a virtual building was developed and used to perform load-based testing of the performance of RTUs with single-stage, two-stage, and variable-speed control technologies. It was demonstrated that this new methodology could be applied to an existing psychrometric chamber control system. The psychrometric room system was able to adjust indoor test chamber conditions to closely track the virtual building model output and provided representative dynamic responses for both the fixed-speed and variable-speed equipment. This enabled performance comparisons of the systems tested in a direct manner under equivalent loads and ambient conditions, including natural cyclic degradation inefficiencies and control methodologies. The embedded control of the equipment was not altered in anyway – reducing potential discrepancies between the measured performance of systems tested in a laboratory and in the field.

Seasonal efficiencies were estimated based on current standards and compared with results determined from load-based test data for the fixed-speed equipment. The current approaches agreed reasonably well with load-based test results if cyclic degradation coefficients were determined from testing, but the standard results were poor using default parameters. In general, it is believed that the load-based testing methodology more accurately and naturally captures the effects of cycling behavior in a systematic manner since the equipment can cycle at its natural frequency rather than a purely time-based test methodology. The load-based methodology also provides a consistent approach for testing units with fixed and variable-speed controls and generally provides a more accurate comparison of the performance of different control technologies since testing is performed with the units integrated controls. In this study, there

was about a 50% improvement in seasonable efficiency when comparing variable-speed with single-stage on/off control of an RTU based on load-based testing.

The approaches presented in this paper represent a first step in the development of improved equipment testing and rating procedures that are needed to properly characterize performance benefits associated with advanced controls. However, there is much additional work that is needed before standards are ready. For example, the virtual building models for sensible and latent loads are rather simplistic and the effect of utilizing more detailed building load models on performance ratings should be considered. It is also critical to demonstrate that load-based testing can provide reproducibility and repeatability of measured equipment performance of the same equipment in different labs. This is certainly important for equipment ratings and should be considered for testing standards. Additionally, the systems installed in this study did not have outdoor air economizers installed. However, most RTU have these installed in the field and performance may be impacted by significant outdoor air fractions entering the systems. It is also important to consider application to residential air conditioners and heat pumps, which have different control requirements that may significantly impact performance, such as defrost cycles. Often, variable-speed equipment requires special thermostats for enabling the full control capability intended by the manufacturer. The dynamic response of the thermostat sensor measurement and how it differs from the true bulk air temperature measurement in psychrometric chambers should be considered.

APPENDIX B. COMPARISONS OF FAULT IMPACT NEURAL NETWORK META-MODEL AND DETAILED MODEL OUTPUTS

This appendix provides further comparisons between the detailed model output and the artificial neural network fault impact meta-model. The trends and accuracy of the meta-model show that the approximation is well-behaved and captures the trends of the detailed model well.

B.1 Comparison of Refrigerant Charge Fault Impact Predictions

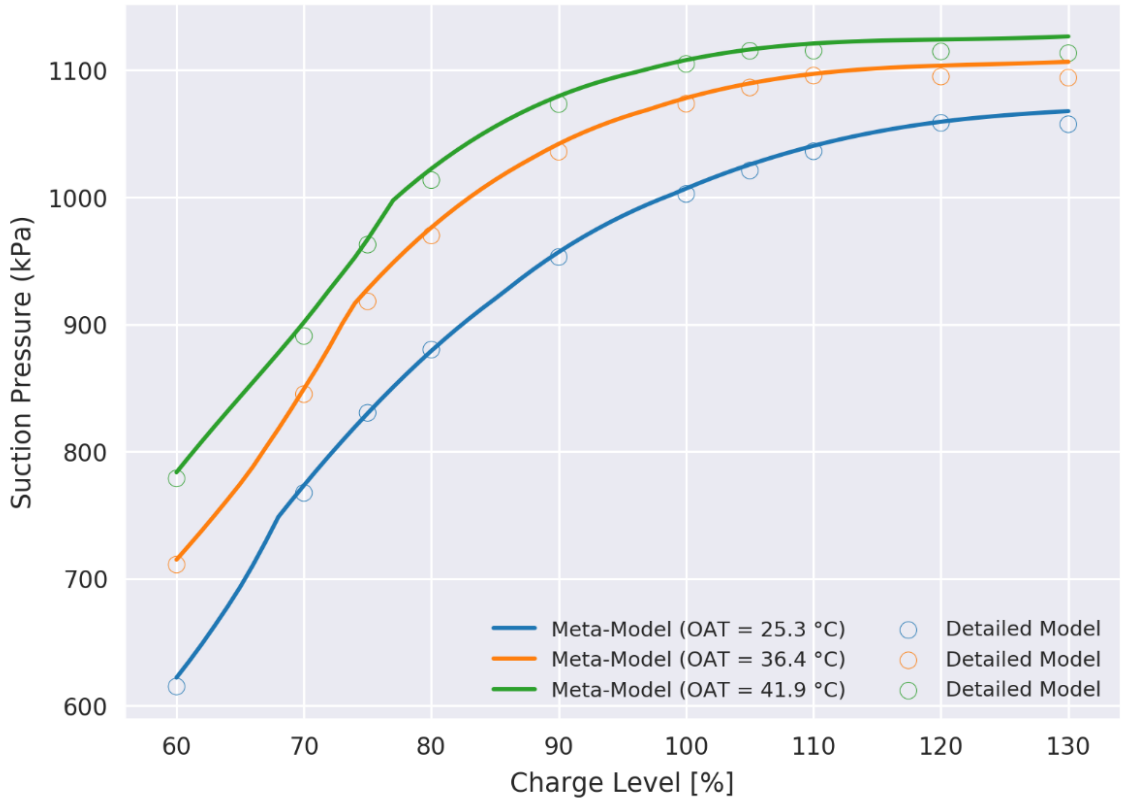

Figure B.1. Comparison of suction pressure trends over a range of refrigerant charge levels predicted by meta-model and detailed model at different ambient conditions for system with fixed orifice expansion valve.

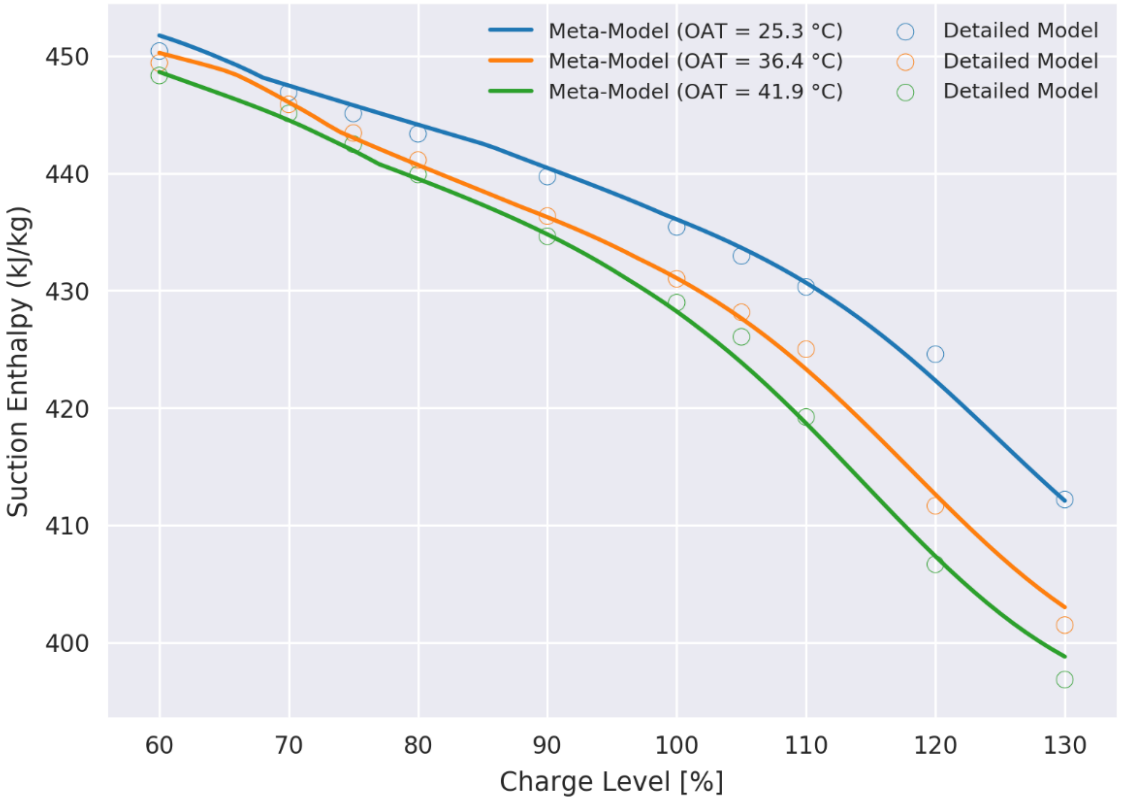

Figure B.2. Comparison of suction enthalpy over a range of refrigerant charge levels predicted by meta-model and detailed model at different ambient conditions for system with fixed orifice expansion valve.

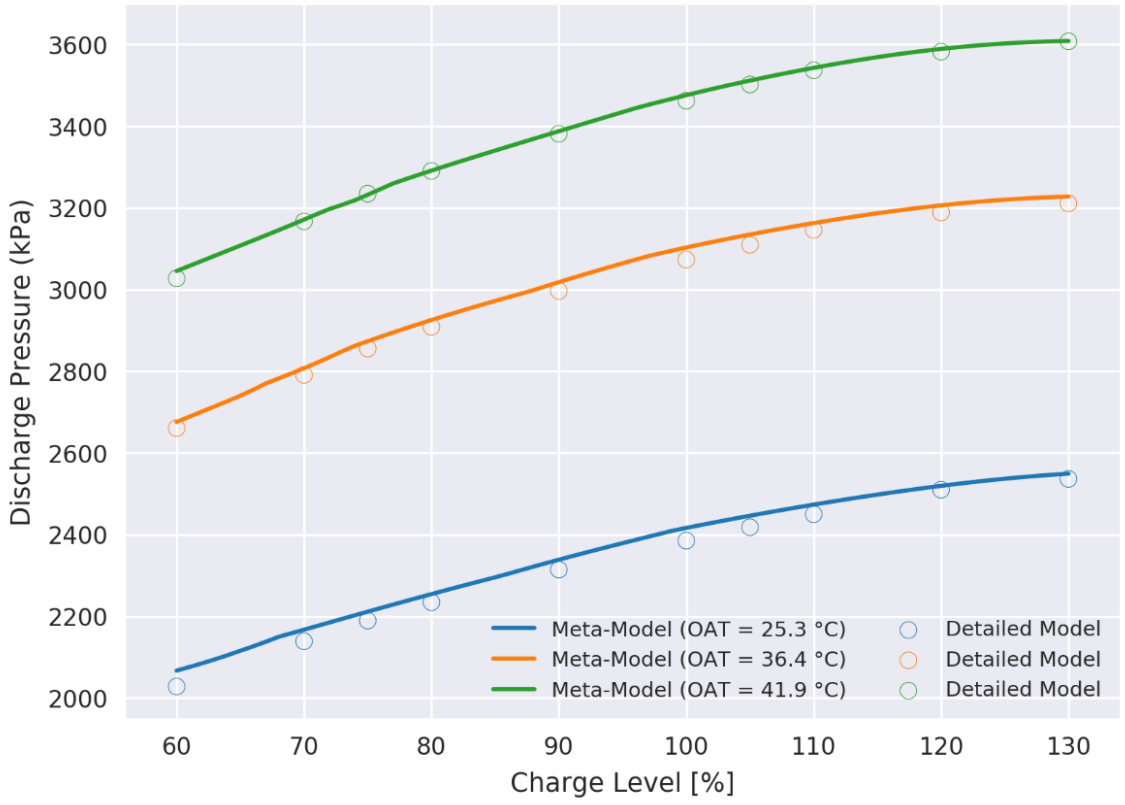

Figure B.3. Comparison of discharge pressure trends over a range of refrigerant charge levels predicted by meta-model and detailed model at different ambient conditions for system with fixed orifice expansion valve.

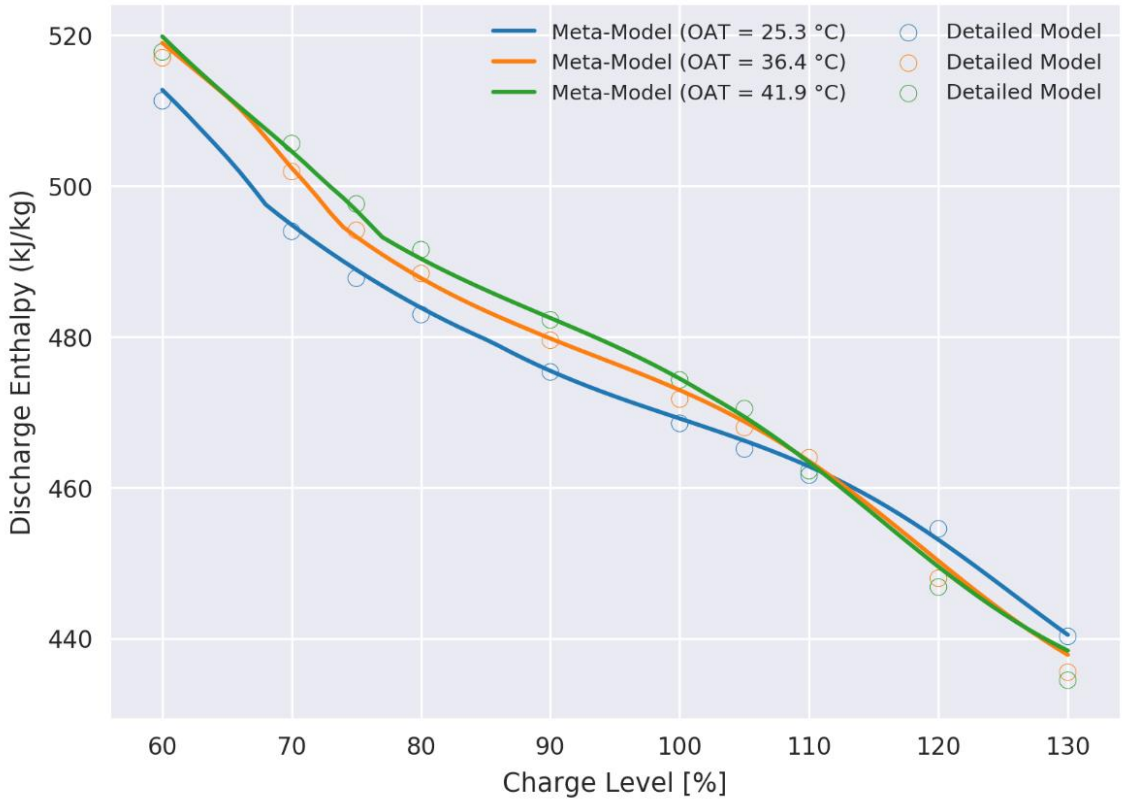

Figure B.4. Comparison of discharge enthalpy trends over a range of refrigerant charge levels predicted by meta-model and detailed model at different ambient conditions for system with fixed orifice expansion valve.

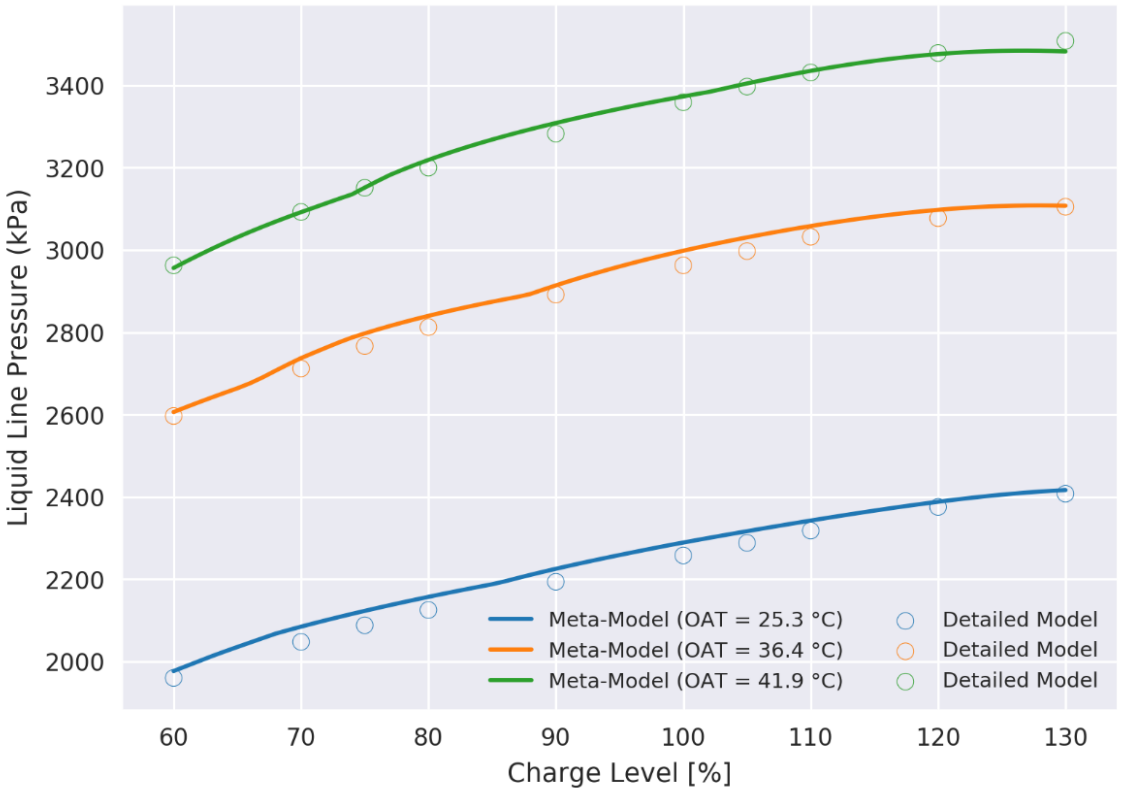

Figure B.5. Comparison of liquid line pressure trends over a range of refrigerant charge levels predicted by meta-model and detailed model at different ambient conditions for system with fixed orifice expansion valve.

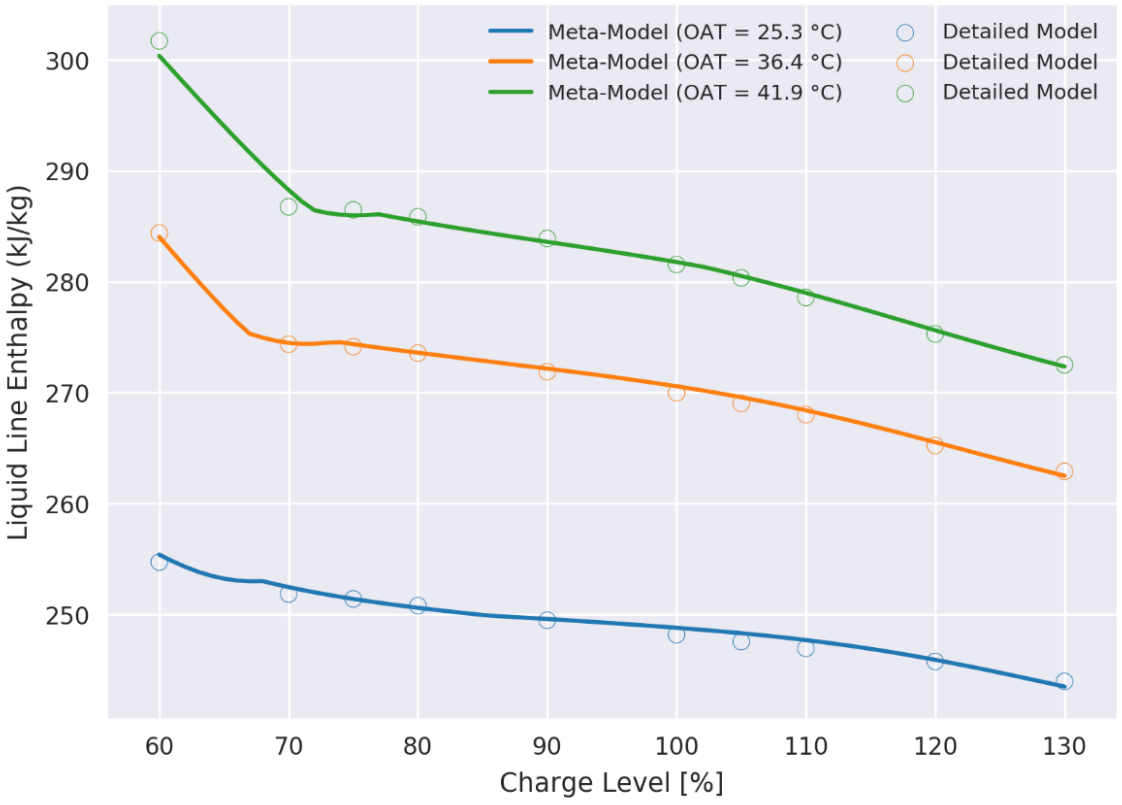

Figure B.6. Comparison of liquid line enthalpy trends over a range of refrigerant charge levels predicted by meta-model and detailed model at different ambient conditions for system with fixed orifice expansion valve.

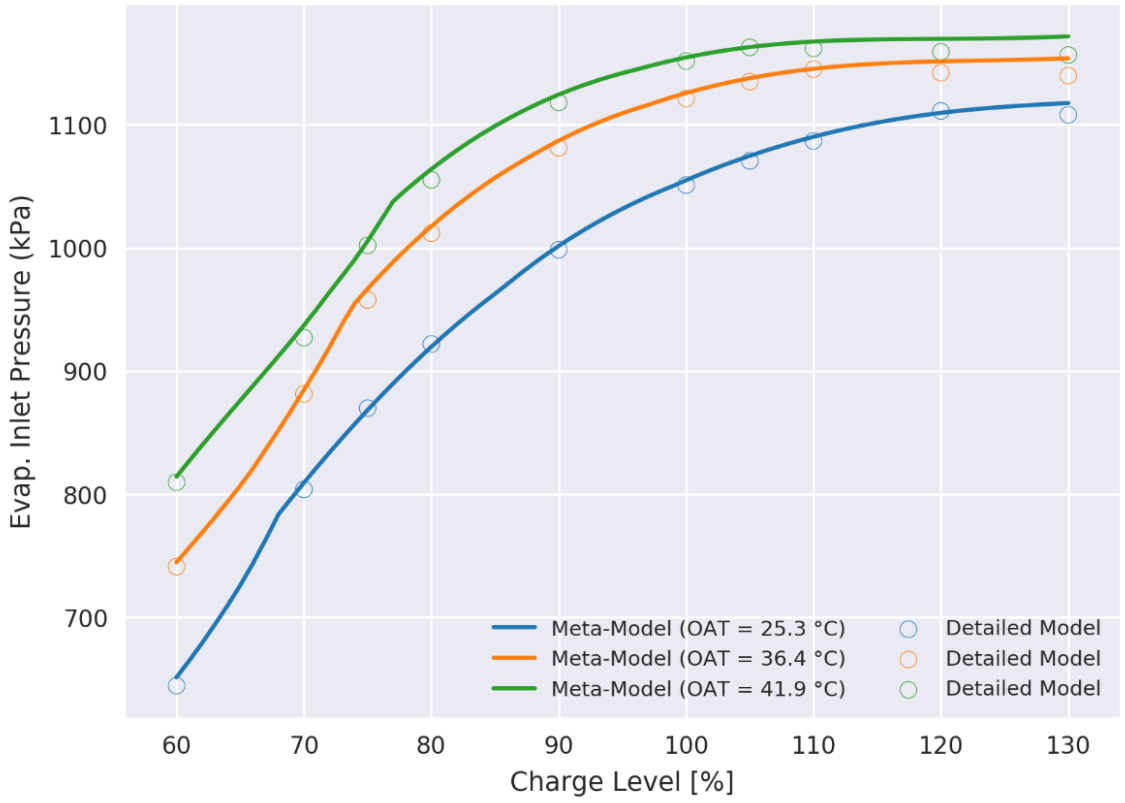

Figure B.7. Comparison of evaporator inlet pressure trends over a range of refrigerant charge levels predicted by meta-model and detailed model at different ambient conditions for system with fixed orifice expansion valve.

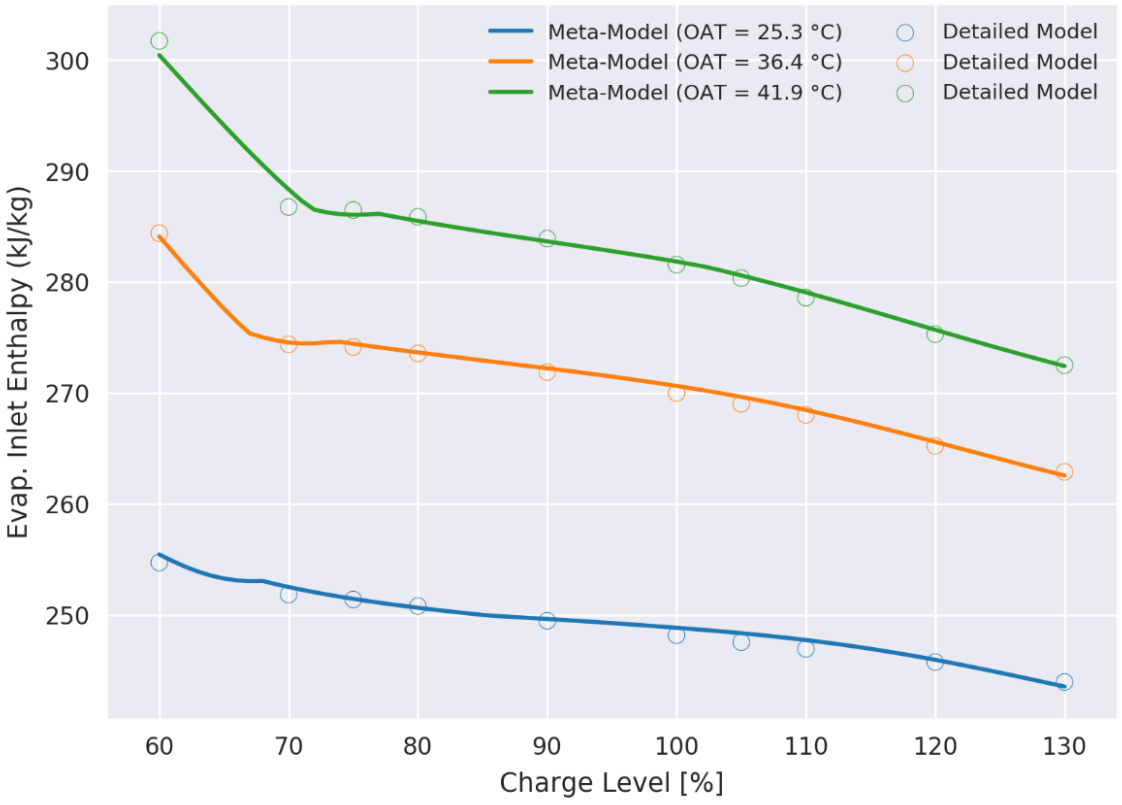

Figure B.8. Comparison of evaporator inlet enthalpy trends over a range of refrigerant charge levels predicted by meta-model and detailed model at different ambient conditions for system with fixed orifice expansion valve.

B.2 Comparison of Condenser Fouling Fault Impact Predictions

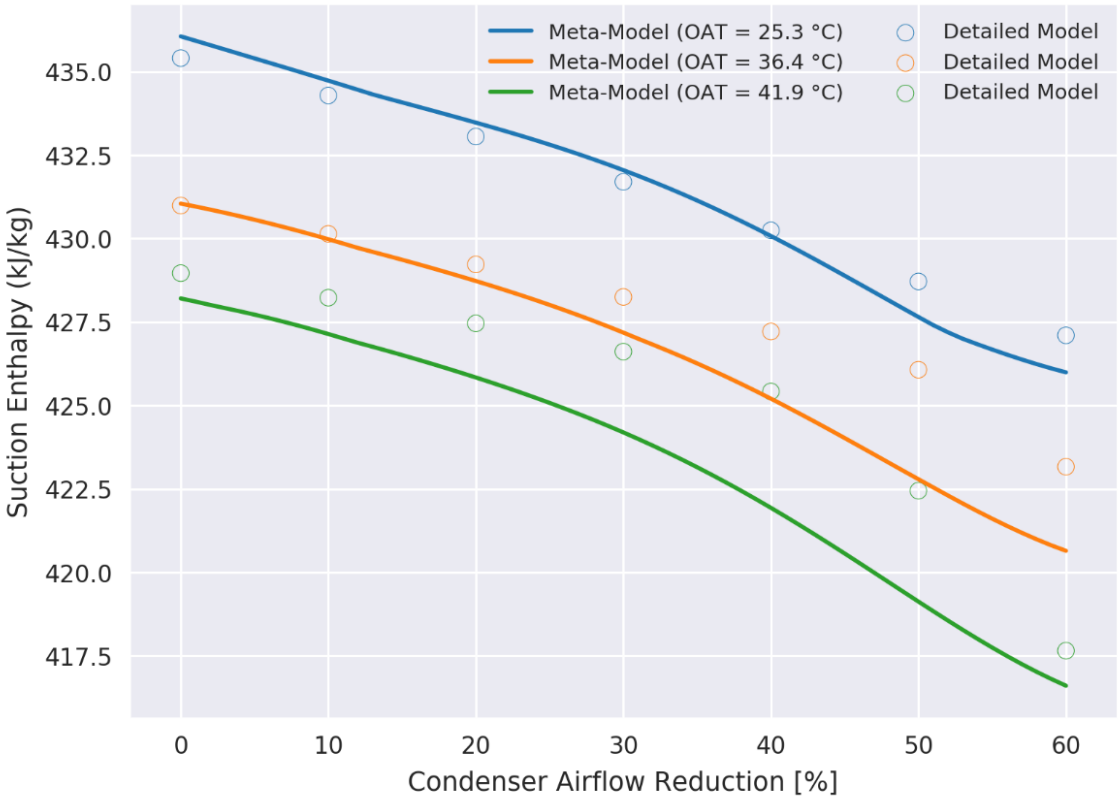

Figure B.9. Comparison of suction pressure trends over a range of condenser air flow reductions predicted by meta-model and detailed model at different ambient conditions for system with fixed orifice expansion valve.

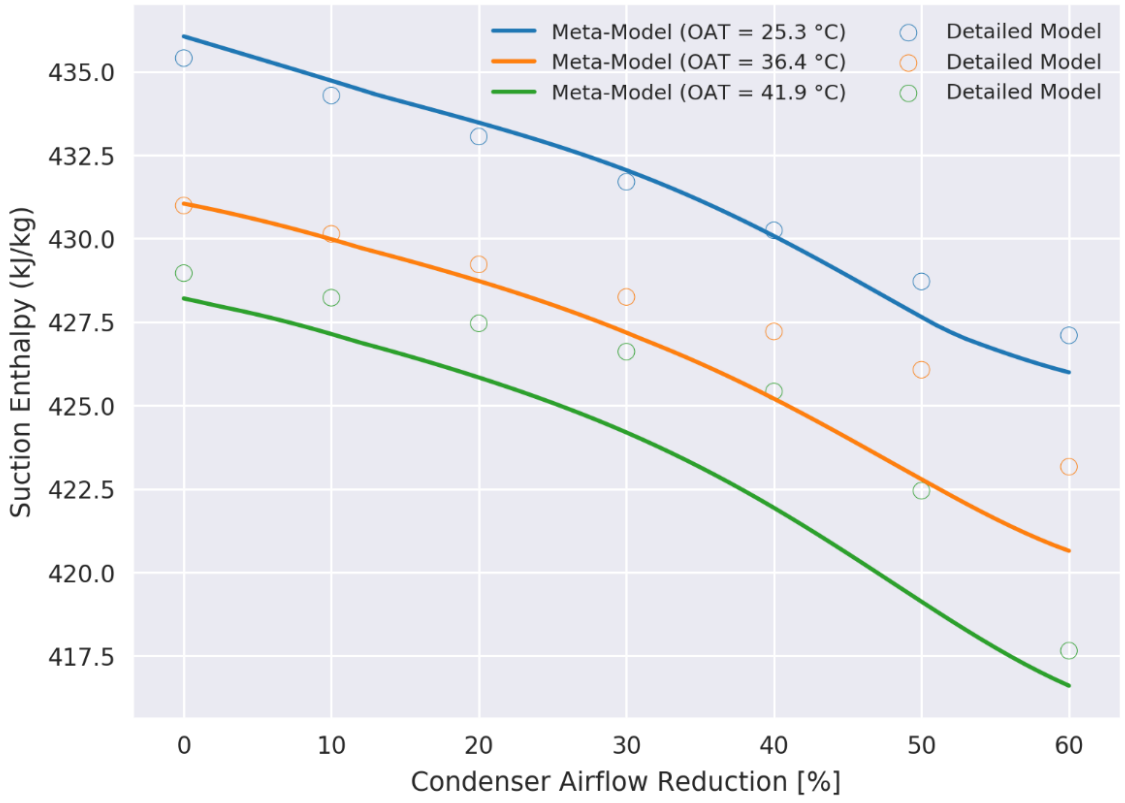

Figure B.10. Comparison of suction enthalpy over a range of condenser air flow reductions predicted by meta-model and detailed model at different ambient conditions for system with fixed orifice expansion valve.

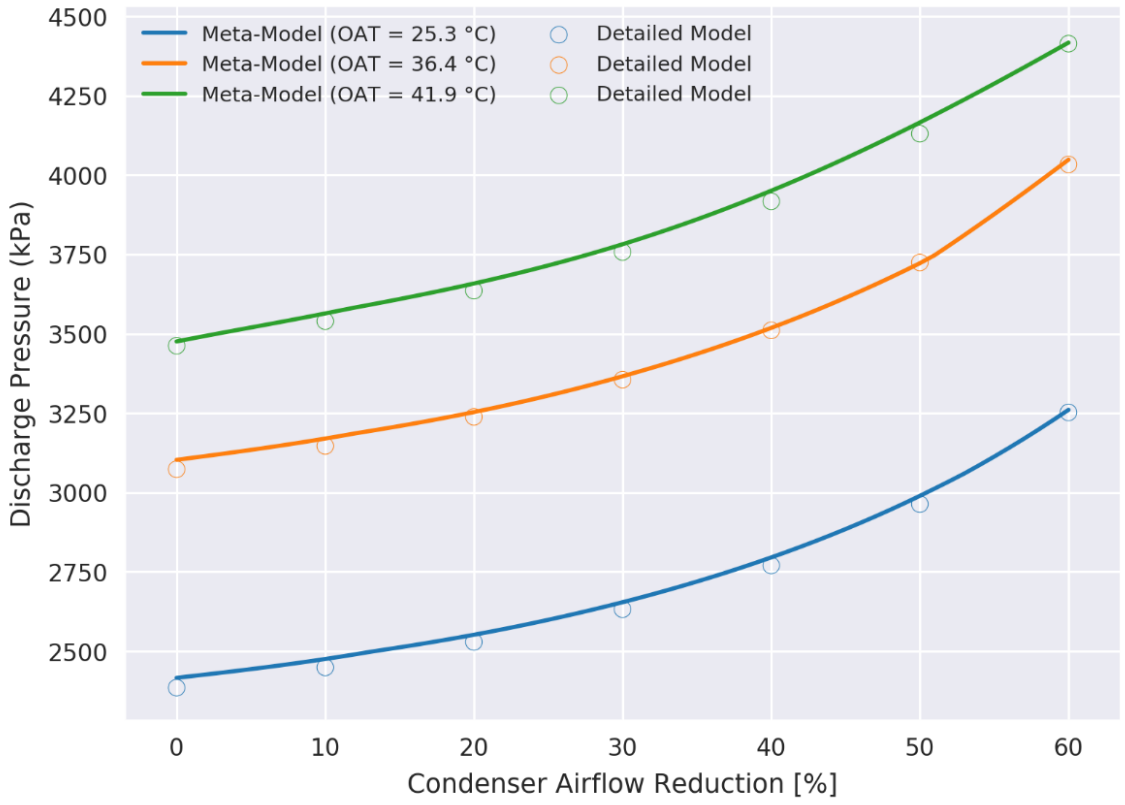

Figure B.11. Comparison of discharge pressure trends over a range of condenser air flow reductions predicted by meta-model and detailed model at different ambient conditions for system with fixed orifice expansion valve.

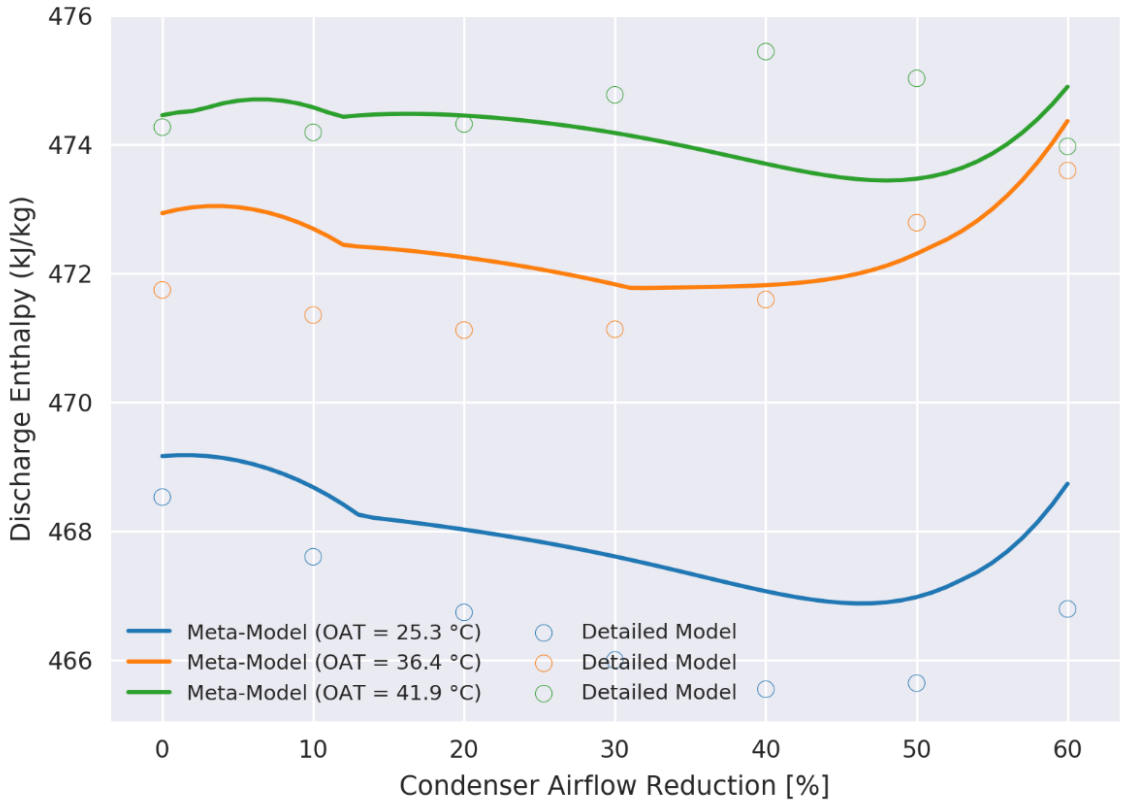

Figure B.12. Comparison of discharge enthalpy trends over a range of condenser air flow reductions predicted by meta-model and detailed model at different ambient conditions for system with fixed orifice expansion valve.

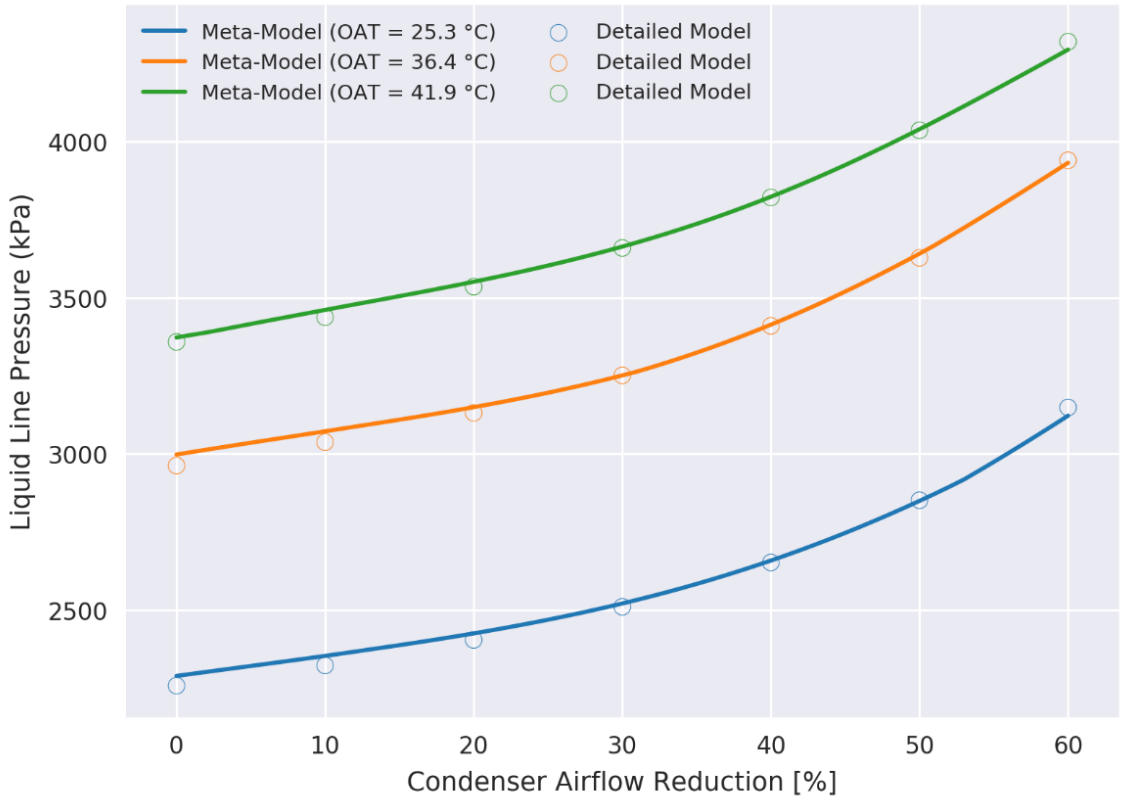

Figure B.13. Comparison of liquid line pressure trends over a range of condenser air flow reductions predicted by meta-model and detailed model at different ambient conditions for system with fixed orifice expansion valve.

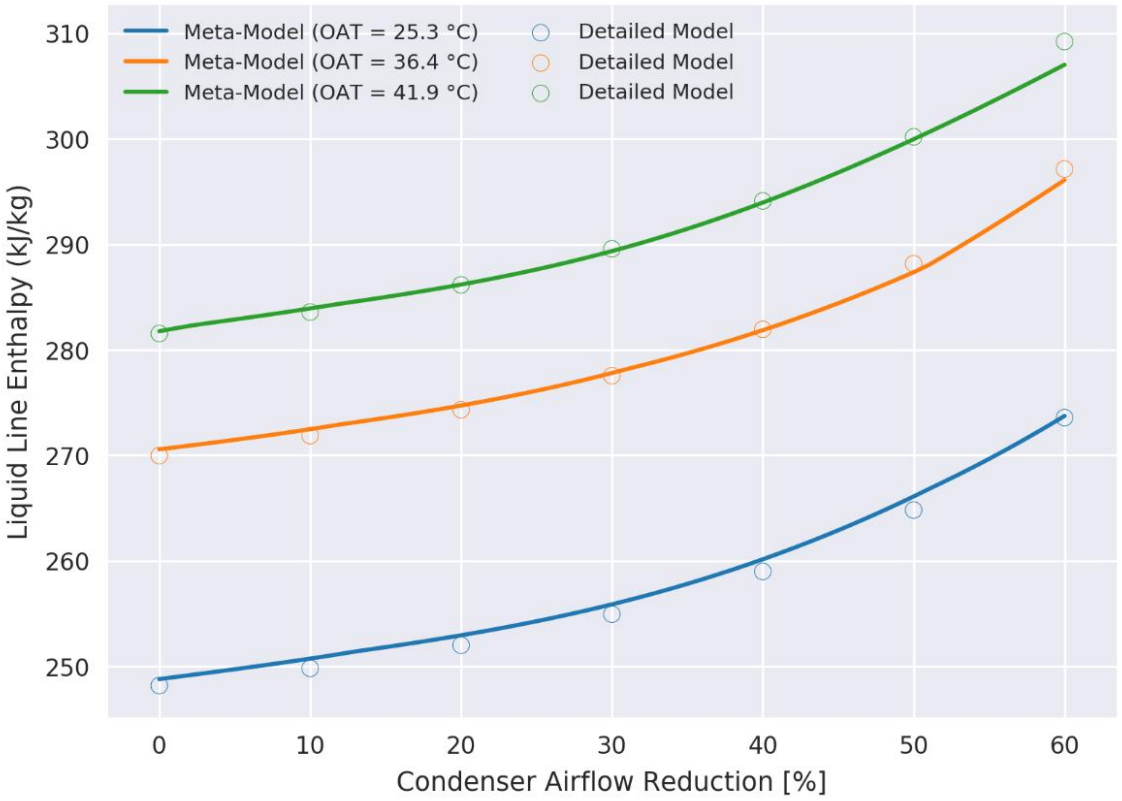

Figure B.14. Comparison of liquid line enthalpy trends over a range of condenser air flow reductions predicted by meta-model and detailed model at different ambient conditions for system with fixed orifice expansion valve.

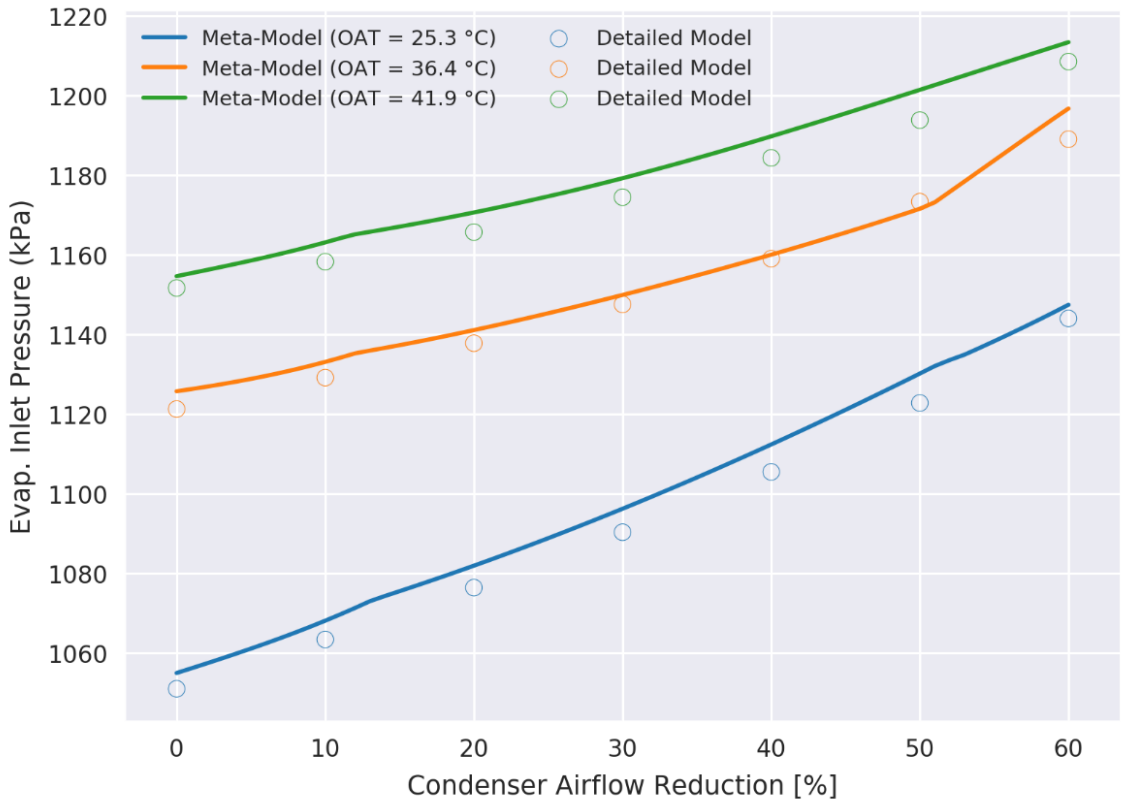

Figure B.15. Comparison of evaporator inlet pressure trends over a range of condenser air flow reductions predicted by meta-model and detailed model at different ambient conditions for system with fixed orifice expansion valve.

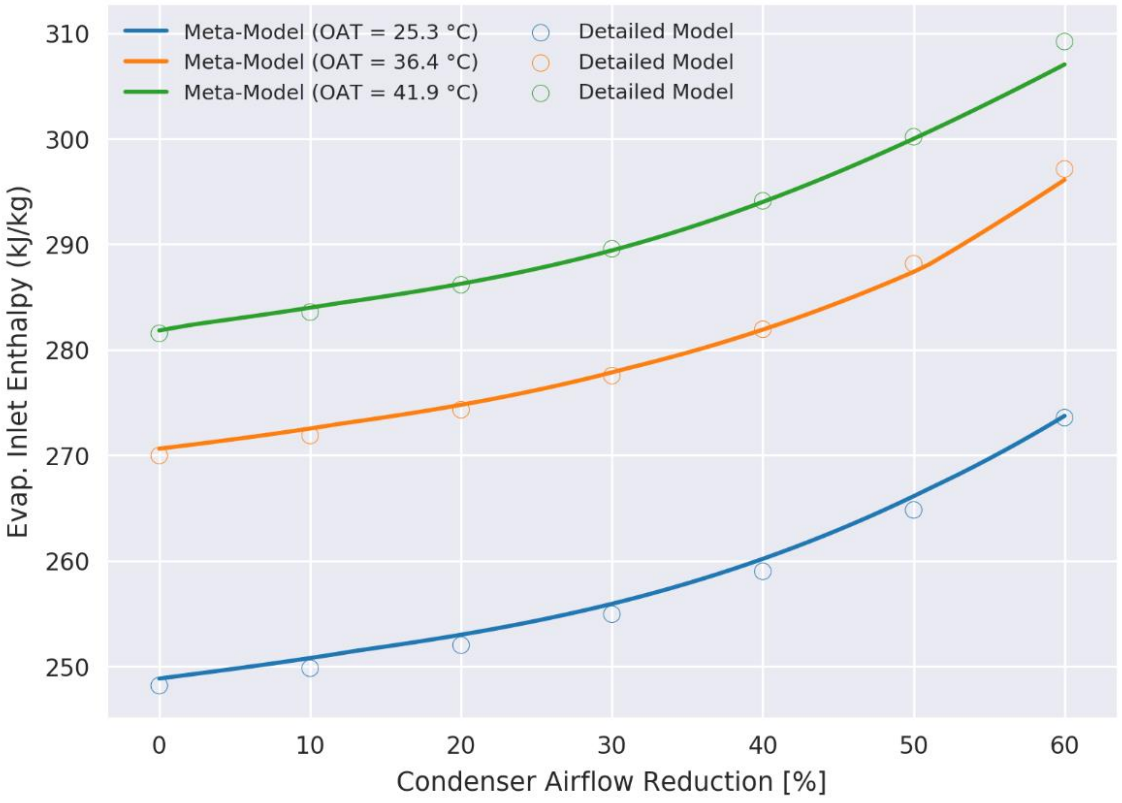

Figure B.16. Comparison of evaporator inlet enthalpy trends over a range of condenser air flow reductions predicted by meta-model and detailed model at different ambient conditions for system with fixed orifice expansion valve.

B.3 Comparison of Evaporator Fouling Fault Impact Predictions

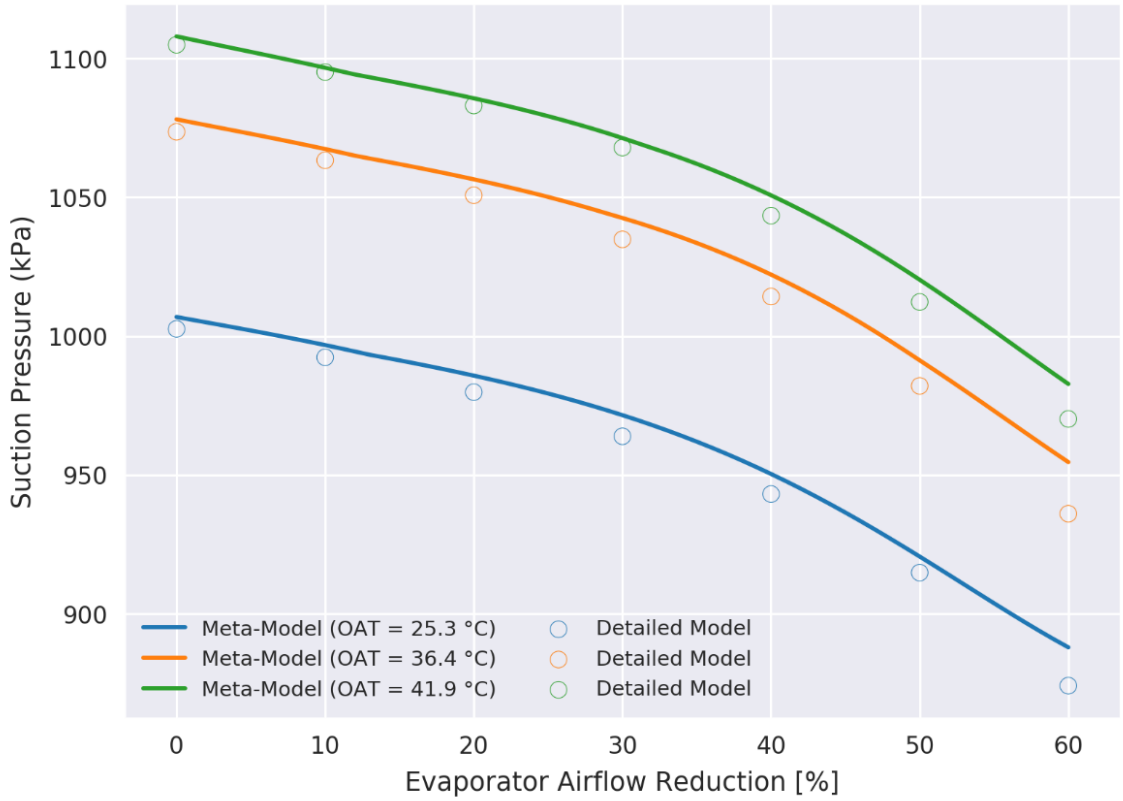

Figure B.17. Comparison of suction pressure trends over a range of evaporator air flow reductions predicted by meta-model and detailed model at different ambient conditions for system with fixed orifice expansion valve.

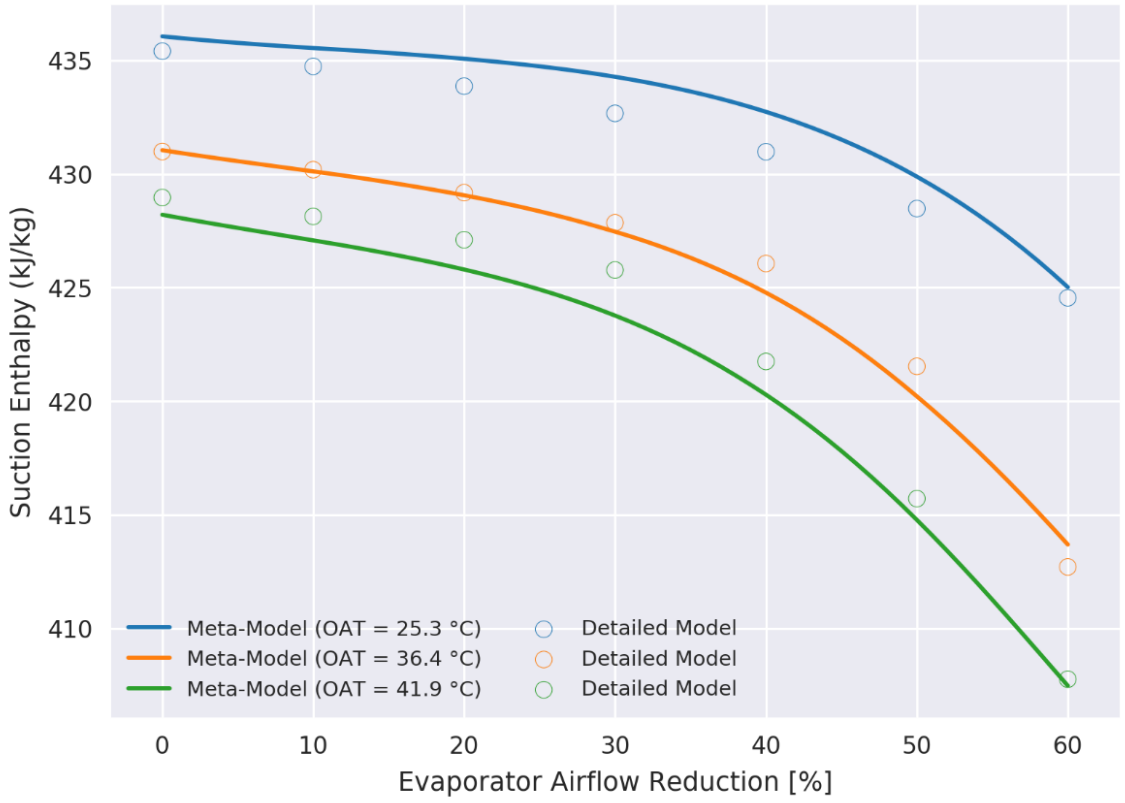

Figure B.18. Comparison of suction enthalpy over a range of evaporator air flow reductions predicted by meta-model and detailed model at different ambient conditions for system with fixed orifice expansion valve.

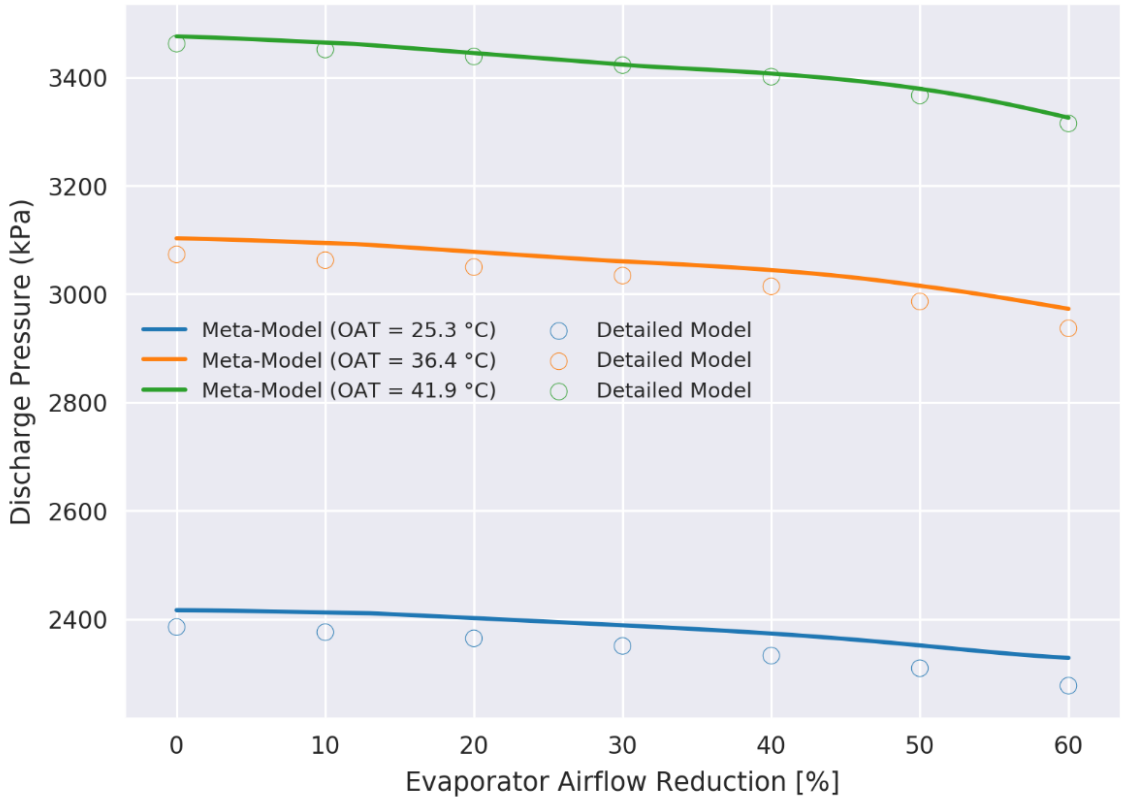

Figure B.19. Comparison of discharge pressure trends over a range of evaporator air flow reductions predicted by meta-model and detailed model at different ambient conditions for system with fixed orifice expansion valve.

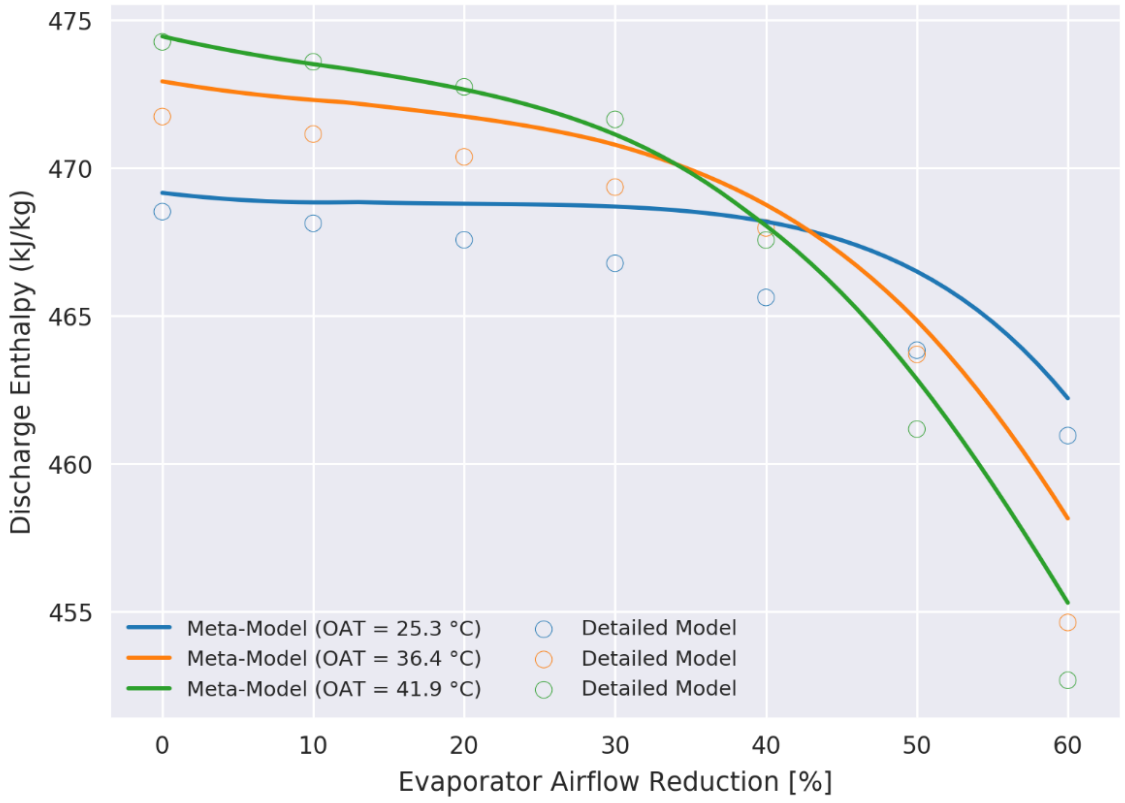

Figure B.20. Comparison of discharge enthalpy trends over a range of evaporator air flow reductions predicted by meta-model and detailed model at different ambient conditions for system with fixed orifice expansion valve.

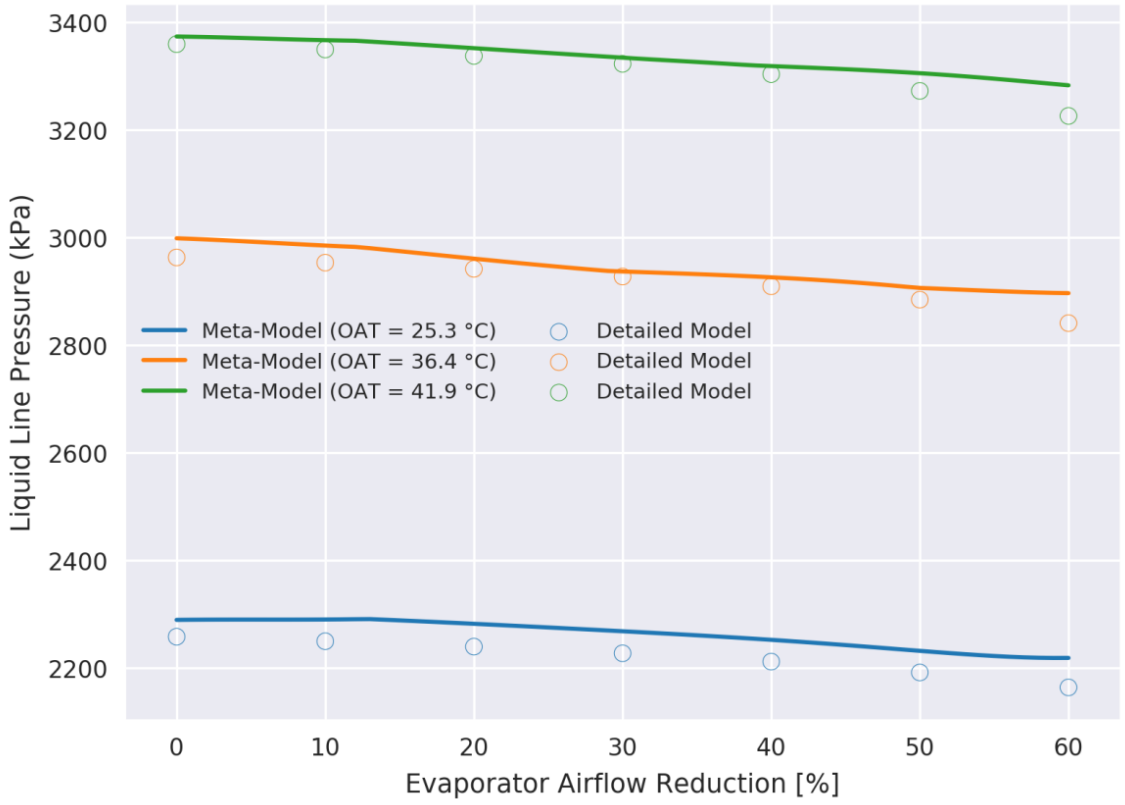

Figure B.21. Comparison of liquid line pressure trends over a range of evaporator air flow reductions predicted by meta-model and detailed model at different ambient conditions for system with fixed orifice expansion valve.

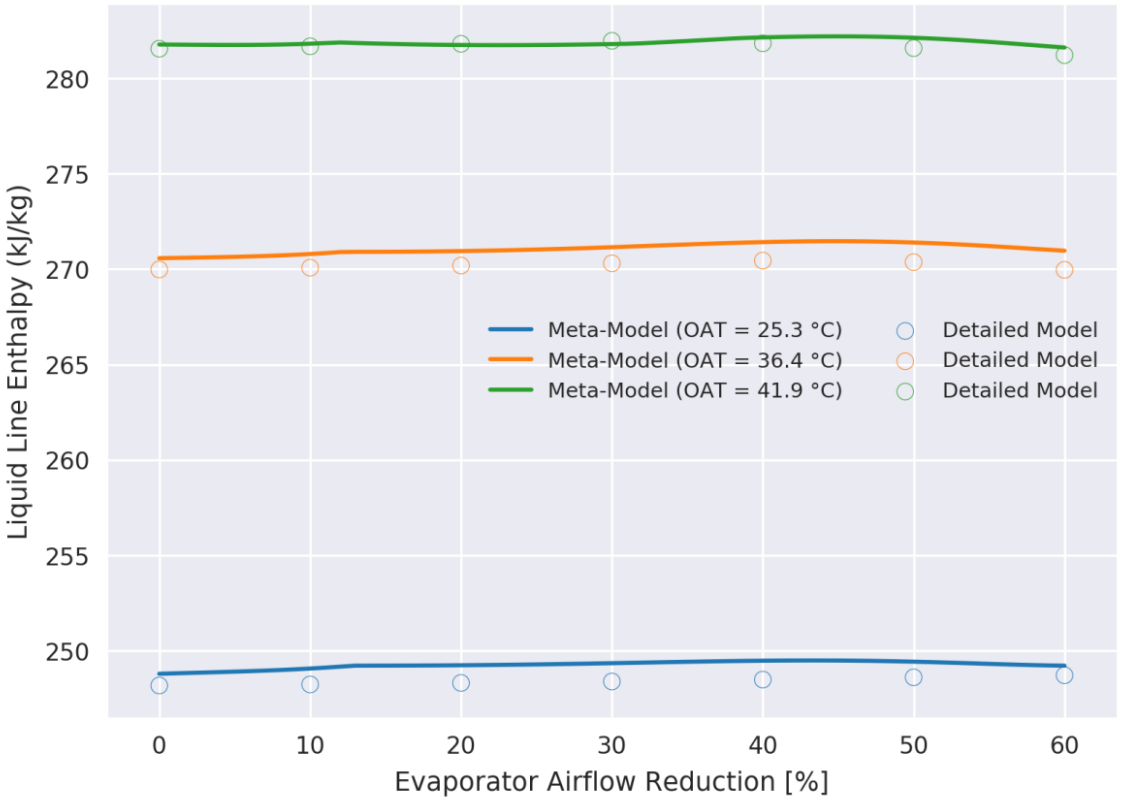

Figure B.22. Comparison of liquid line enthalpy trends over a range of evaporator air flow reductions predicted by meta-model and detailed model at different ambient conditions for system with fixed orifice expansion valve.

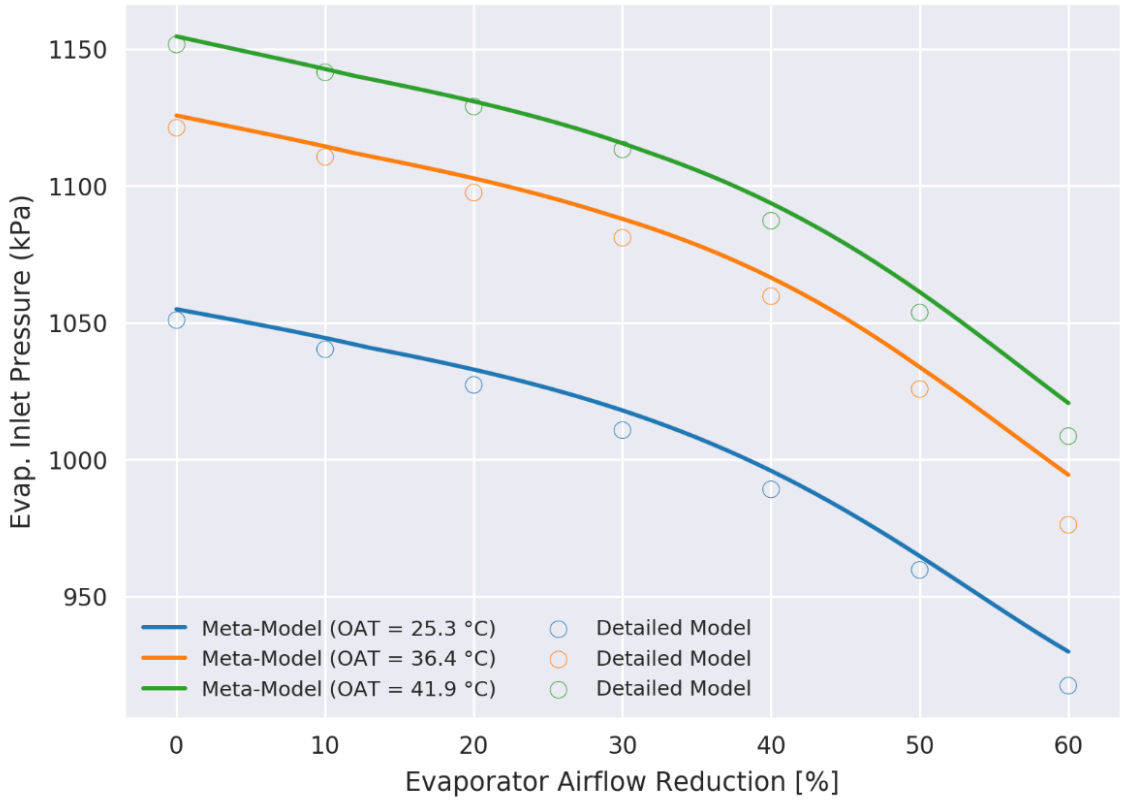

Figure B.23. Comparison of evaporator inlet pressure trends over a range of evaporator air flow reductions predicted by meta-model and detailed model at different ambient conditions for system with fixed orifice expansion valve.

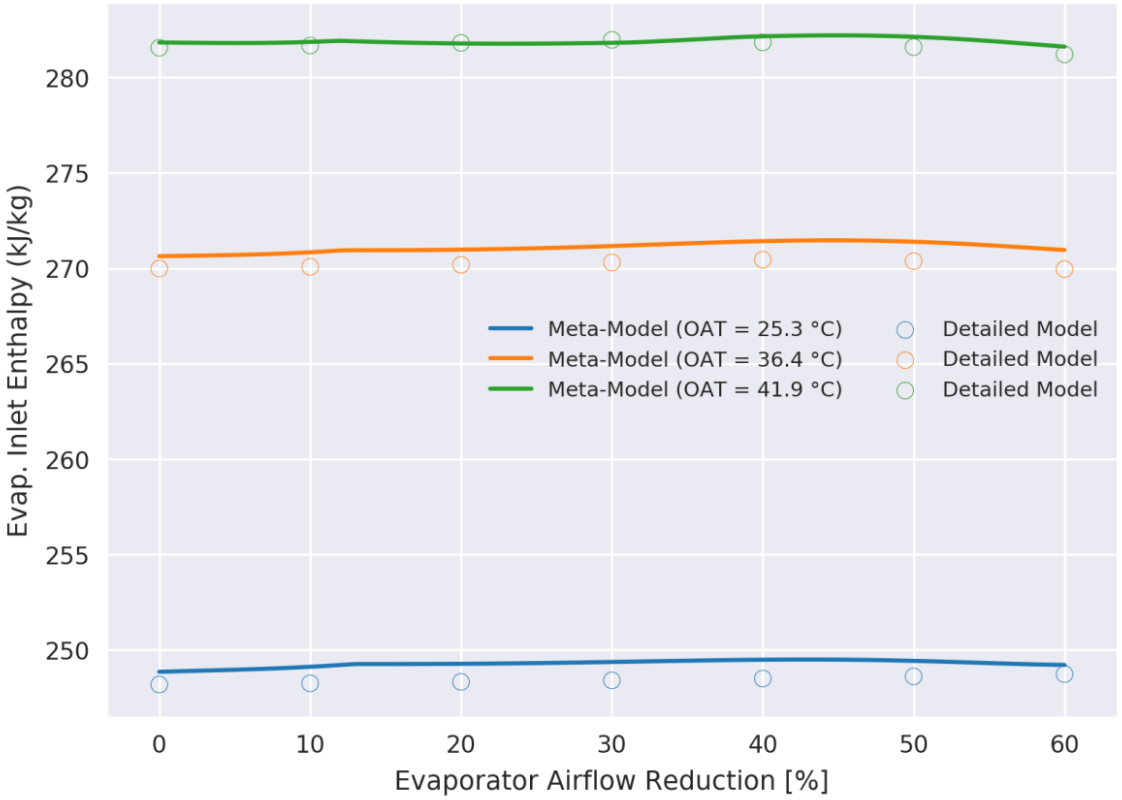

Figure B.24. Comparison of evaporator inlet enthalpy trends over a range of evaporator air flow reductions predicted by meta-model and detailed model at different ambient conditions for system with fixed orifice expansion valve.

VITA

Mr. Andrew Hjortland received his Bachelor of Science degree in Mechanical Engineering from the Milwaukee School of Engineering (MSOE) in the spring of 2011, in which his area of emphasis was on thermodynamics. His senior design project was designing and building solar photovoltaic and solar thermal tests stands for undergraduate engineering education. While at MSOE, he worked as a electrical drafter at the University of Wisconsin – Madison Synchrotron Radiation Center and later as a technical intern at Johnson Controls, Inc.

Upon entering the Mechanical Engineering graduate program at Purdue University in the Fall of 2011, Mr. Hjortland began a project sponsored by the US Department of Energy's Energy Efficient Buildings Hub on rooftop air conditioner fault detection and diagnostics. This work led to his Master's Thesis: "Probabilistic Fault Detection and Diagnostics for Packaged Air Conditioner Outdoor-Air Economizers." After completing his Master of Science degree in Mechanical Engineering, Mr. Hjortland began work which lead to his doctoral dissertation: "Estimating the Impacts of Multiple Faults on Packaged Air Conditioning Equipment Performance." From 2011 to 2012, he was awarded the Laura Winkelman Davidson fellowship by the Purdue University Mechanical Engineering Department. While at Purdue University, he has published several technical journal and conference papers.

PUBLICATIONS

Andrew L. Hjortland & James E. Braun. 2018. "Load-based Testing Methodology for Fixed-Speed and Variable-Speed Unitary Air Conditioning Equipment." *Science and Technology for the Built Environment. In review.*

Andrew L. Hjortland & James E. Braun. 2018. "Artificial Neural Networks for Fast Rooftop Unit Fault Impact Modeling and Simulation." *International Refrigeration and Air Conditioning Conference. In review.*

Andrew L. Hjortland & James E. Braun. 2018. "Comparing Maintenance Strategies for Rooftop Units having Multiple Faults through Simulation." *International Refrigeration and Air Conditioning Conference. In review.*

Donghun Kim, **Andrew L. Hjortland**, Timothy I. Salsbury, Kirk H. Drees, John M. House and James E. Braun. 2018. "Feasibility Study of Extremum Seeking Controller Applied to a Variable-Speed Rooftop Unit." *International Refrigeration and Air Conditioning Conference. In review.*

Andrew L. Hjortland & James E. Braun. 2018. "Efficient Virtual Refrigerant Charge Sensor Tuning using D-Optimal Experimental Design." *In preparation.*

Andrew L. Hjortland, Akash Patil, & James E. Braun. 2018. "Automated Virtual Refrigerant Charge Sensor Training Methodology for RTUs." *In preparation.*

Andrew L. Hjortland & James E. Braun. 2018. "An Economic Impact Model for RTU Stuck Outdoor-Air Damper Faults during Mechanical Cooling Operation." *In preparation.*

Andrew L. Hjortland, James E. Braun & Mikhail Gorbounov. 2016. "The Effectiveness of using Total System Power for Fault Detection in Rooftop Units." *International Refrigeration and Air Conditioning Conference*. Paper 2627

Akash Patil, **Andrew L. Hjortland** & James E. Braun. 2016. "Development and Evaluation of an Automated Virtual Refrigerant Charge Sensor Training Kit." *International Refrigeration and Air Conditioning Conference*. Paper 2629

Andrew L. Hjortland & James E. Braun. 2016. "Development of an Embedded RTU FDD using Open-Source Monitoring and Control Platform" *International Refrigeration and Air Conditioning Conference*. Paper 2379

Jinliang Wang, Mikhail Gorbounov, Murat Yasar, Hayden Reeve, **Andrew L Hjortland**, & James E Braun. 2016. "Lab and Field Evaluation of Fault Detection and Diagnostics for Advanced Roof Top Unit." *International Refrigeration and Air Conditioning Conference*. Paper 2069

Andrew L. Hjortland & James E. Braun. 2016. "Virtual Sensors for Rooftop Unit Air-side Diagnostics." *Science and Technology for the Built Environment*, 22(2): 189–200. http://dx.doi.org/10.1080/23744731.2016.1124715

Andrew L. Hjortland & James E. Braun. 2014. "Development of Economic Impact Models for RTU Economizer Faults." *International Refrigeration and Air Conditioning Conference*. http://docs.lib.purdue.edu/iracc/1437/

Andrew L. Hjortland. May 2014. *Probabilistic Fault Detection and Diagnostics for Packaged Air-Conditioner Outdoor-Air Economizers*. Master's Thesis. Purdue University, School of Mechanical Engineering. Major Professor: James E. Braun.

Andrew L. Hjortland & James E. Braun. 2012. "General Outdoor Air Economizer Fault Detection & Diagnostics Assessment Method." *International Refrigeration and Air Conditioning Conference*. http://docs.lib.purdue.edu/iracc/1249/

Stephanie Drozek, Christopher Damm, Ryan Enot, **Andrew Hjortland**, Brandon Jackson, Benjamin Steffes, & Kevin Rode. 2012. "Design, Installation, and Performance Characterization of a Laboratory-Scale Solar Thermal System for Experiments in Solar Energy Utilization." *ASME 2012 International Mechanical Engineering Congress and Exposition. Vol. 5. Houston, TX.*