

# **GREY BOX MODELLING AND ADVANCED CONTROL SCHEME FOR BUILDING HEATING SYSTEMS**

by

Surinder Jassar

Master of Technology Degree in Instrumentation and Control  
Engineering, Panjab Agricultural University, India, 2001

A dissertation  
presented to Ryerson University  
in partial fulfillment of the  
requirement for the degree of  
Doctor of Philosophy  
in the Program of  
Electrical and Computer Engineering.

Toronto, Ontario, Canada, 2011

© Surinder Jassar, 2011

## **Author's Declaration**

I hereby declare that I am the sole author of this dissertation.

I authorize Ryerson University to lend this dissertation to other institutions or individuals for the purpose of scholarly research.

Signature

I further authorize Ryerson University to reproduce this dissertation by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

Signature

## Instructions on Borrowers

Ryerson University requires the signatures of all persons using or photocopying this dissertation. Please sign below, and give address and date.

[illegible]

## **Abstract**

# **GREY BOX MODELLING AND ADVANCED CONTROL SCHEME FOR BUILDING HEATING SYSTEMS**

**©Surinder Jassar**

**Doctor of Philosophy**

**Electrical and Computer Engineering**

**Ryerson University, Toronto, Canada, 2011.**

This dissertation is aimed at generating new knowledge on Recurrent Neuro-Fuzzy Inference Systems (RenFIS) and to explore its application in building automation.

Inferential sensing is an attractive approach for modeling the behavior of dynamic processes. Inferential sensor based control strategies are applied to optimize the control of residential heating systems and demonstrate significant energy saving and comfort improvement. Despite the rapidly decreasing cost and improving accuracy of most temperature sensors, it is normally impractical to use a lot of sensors to measure the average air temperature because the wiring and instrumentation can be very expensive to install and maintain.

To design a reliable inferential sensor, of fundamental importance is to build a simple and robust dynamic model of the system to be controlled. This dissertation presents the development of an innovative algorithm that is suitable for the robust black-box model. The algorithm is derived from ANFIS (Adaptive Neuro-Fuzzy Inference System) and is referred to as RenFIS.

Like all other modeling techniques, RenFIS performance is sensitive to the training data. In this study, RenFIS is used to model two different heating systems, hot water heating system

and forced warm-air heating system. The training data is collected under different operational conditions. RenFIS gives better performance if trained with the data set representing overall qualities of the whole universe of the experimental data. The robustness analysis is conducted by introducing simulated noise to the training data. Results show that RenFIS is less sensitive than ANFIS to the quality of training data.

The RenFIS based inferential sensor is then applied to design an inferential control algorithm that can improve the operation of residential heating systems. In current practice, the control of heating systems is based on the measurement of air temperature at one point within the building. The inferential control strategy developed through this study allows the control to be based on an estimate of the overall thermal performance. This is achieved through estimating the average room temperature using a RenFIS based inferential sensor and incorporating the estimate with conventional control technology. The performance of this control technology has been investigated through simulation study.

# Acknowledgments

It is a pleasure to thank the many people who made this dissertation possible.

I am heartily thankful to my supervisors, Dr. Zaiyi Liao and Dr. Lian Zhao, for their professional directions, essential comments and kind support. With their enthusiasm, their inspiration, and their great efforts to explain things clearly and simply, they helped to make research fun for me. Throughout my dissertation period they provided sound advice and good ideas. I would have lost without them.

I am also thankful to my examination committee, Dr. Kamran Raahemifar, Dr. Matthew Kyan, and Dr. Hua Ge for their insightful comments on the dissertation. I am very grateful to Dr. Jian-Lei Niu for his valuable comments and for him travelled from Hong Kong to Toronto for my defense.

This work is partially supported by National Sciences and Engineering Research Council of Canada (NSERC) with research project numbers as: 313375-07 and 293237-09, and Ryerson University.

My special thanks go to my husband, Jagwinder, for his incessant love and companionship. Without him, this work would have never completed. I am also grateful to my little daughters, Arvind and Ekroop who always cheer me up. Last but not least, I am thankful to my parents for their endless love and support.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Objectives of the Study . . . . .	2
1.2	Background . . . . .	3
1.2.1	Climate Change: . . . . .	3
1.2.2	Indoor Environment/Thermal Comfort . . . . .	4
1.2.3	Energy . . . . .	5
1.3	Methodology . . . . .	7
1.3.1	Survey . . . . .	7
1.3.2	Modelling . . . . .	8
1.3.3	Experimental Data . . . . .	9
1.3.4	Performance of the Inferential Model . . . . .	9
1.3.5	Simulation Study . . . . .	10
1.4	The Potential for Improving Energy Efficiency and Thermal Comfort . . . .	10
1.5	Modelling Environment . . . . .	12
1.6	Structure of the Dissertation . . . . .	16
<b>2</b>	<b>Principle of Hybrid Neuro-Fuzzy Systems</b>	<b>18</b>
2.1	Introduction . . . . .	18
2.2	Fuzzy Logic Systems . . . . .	18
2.3	Fuzzy Inference System . . . . .	19
2.4	Neural Networks . . . . .	21
2.5	Hybrid Neuro-Fuzzy Systems . . . . .	23
2.5.1	Adaptive Neuro-Fuzzy Inference Systems . . . . .	25
2.5.2	ANFIS Architecture . . . . .	26
2.6	Neuro-Fuzzy System Identification . . . . .	28
2.6.1	Structure Identification . . . . .	30
2.6.2	Parameter Identification . . . . .	42
2.7	The Feedback Hybrid Neuro-Fuzzy System . . . . .	45
2.8	Concluding Remarks . . . . .	46
<b>3</b>	<b>Development of the RenFIS based Inferential Model</b>	<b>48</b>
3.1	Introduction . . . . .	48

3.2	Neuro-Fuzzy Modelling for Average Air Temperature Estimation . . . . .	51
3.2.1	Input and Output Variable Selection . . . . .	52
3.2.2	Experimental Data . . . . .	53
3.2.3	Space Partitioning . . . . .	55
3.2.4	Number and Type of MFs . . . . .	55
3.2.5	Learning of the Model Parameters . . . . .	60
3.3	Model Validation . . . . .	67
3.3.1	Performance Indicators . . . . .	67
3.3.2	Results . . . . .	67
3.4	Concluding Remarks . . . . .	73
<b>4</b>	<b>Impact of Data Quality on Model Performance</b>	<b>77</b>
4.1	Introduction . . . . .	77
4.2	Experimental Methodology . . . . .	78
4.2.1	Experimental Factors . . . . .	79
4.2.2	Experimental Design . . . . .	80
4.3	Experimental Results . . . . .	82
4.3.1	Experiment 1 Results: Errors in the Test Data . . . . .	82
4.3.2	Experiment 2 Results: Errors in the Training Data . . . . .	83
4.4	TANE Algorithm for Noisy Data Detection . . . . .	83
4.4.1	Functional Dependencies . . . . .	83
4.4.2	TANE Algorithm . . . . .	85
4.4.3	Improvement in Estimation Accuracy . . . . .	86
4.5	Training Data: Sensitivity and Robustness Analysis . . . . .	86
4.5.1	Sensitivity Analysis . . . . .	87
4.5.2	Robustness Analysis . . . . .	93
4.6	Concluding Remarks . . . . .	99
<b>5</b>	<b>Control of Residential Forced Warm-Air Heating Systems</b>	<b>104</b>
5.1	Control of Building Heating Systems . . . . .	104
5.2	Control of Forced Warm-Air Heating Systems . . . . .	106
5.3	Improving Energy Efficiency and Thermal Comfort . . . . .	109
5.3.1	Energy Efficiency and Thermal Comfort: High Efficiency Heating Systems . . . . .	109
5.3.2	Energy Efficiency and Thermal Comfort: Control of Heating Systems . . . . .	111
5.4	Survey . . . . .	115
5.4.1	Sample Population Distribution . . . . .	115
5.4.2	Survey Findings and Analysis . . . . .	117
5.5	Inferential Sensor Technology . . . . .	120
5.6	Concluding Remarks . . . . .	121



<b>6</b>	<b>Advanced Control Scheme for Building Heating Systems</b>	<b>122</b>
6.1	Introduction . . . . .	122
6.2	The Inferential Control Scheme . . . . .	123
6.3	The Performance of the Inferential Control Scheme . . . . .	124
6.3.1	The Simulated Building . . . . .	124
6.3.2	Performance Evaluation Indices . . . . .	125
6.3.3	Analysis of Different Control Strategies . . . . .	128
6.4	Performance of Different Control Strategies . . . . .	129
6.4.1	The Performance of the Proposed Control Strategy in a Heating System with a Condensing and ON-OFF Control Furnace . . . . .	130
6.4.2	The Performance of the Proposed Control Strategy in a Heating Systems with a Non-Condensing and ON-OFF Control Furnace . . . . .	134
6.4.3	The Performance of the Proposed Control Strategy in a Heating Systems with a Condensing Furnace, Controlled Blower Speed and Combustion Level . . . . .	136
6.4.4	The Performance of the Proposed Control Strategy in Heating Systems with a Non-Condensing Furnace, Controlled Blower Speed and Combustion Level . . . . .	139
6.5	Concluding Remarks . . . . .	142
<b>7</b>	<b>Conclusion and Future Work</b>	<b>146</b>
7.1	Conclusion of the Dissertation Work . . . . .	146
7.2	Future Work . . . . .	149
	<b>Bibliography</b>	<b>151</b>
	<b>A List of Publications</b>	<b>168</b>
	<b>B Nomenclature</b>	<b>170</b>
	<b>C Abbreviations</b>	<b>173</b>
	<b>D Survey Questionnaire</b>	<b>175</b>

# List of Tables

1.1	ASHRAE-55 Standard for Thermal Comfort. . . . .	5
3.1	Model Architectures and Training Parameters. . . . .	66
3.2	Statistical Analysis between Training and Testing Data Sets. . . . .	66
3.3	Hot-Water Heating System Results. . . . .	71
3.4	Forced Warm-Air Heating System Results. . . . .	75
4.1	Experimental Design. . . . .	81
4.2	Four Combinations of the Variables for each Fraction Error Level in Experiment 1. . . . .	81
4.3	Randomly assigned Percentage Increase (+) or Decrease (-) for a given Amount-Error Level in Experiment 1. . . . .	82
4.4	Approximate Functional Dependencies Detected using the TANE Algorithm. . . . .	85
4.5	Comparison of Results. . . . .	86
4.6	Sensitivity Analysis for Hot-Water Heating System. . . . .	101
4.7	Sensitivity Analysis for Forced Warm-Air Heating System. . . . .	101
4.8	Robustness Analysis for Hot-Water Heating System. . . . .	102
4.9	Robustness Analysis for Forced Warm-Air Heating System. . . . .	102
5.1	Residential Heating and Cooling Systems (ASHRAE 2007). . . . .	105
5.2	Residential Heating and Cooling Energy Efficient Techniques. . . . .	114
5.3	Summary of Survey Sample. . . . .	116
5.4	Satisfaction with Room Temperature for Male and Female Respondents. . . . .	117
5.5	Percentage of Different Types of Houses Equipped with Programmable Thermostats. . . . .	118
6.1	Occupancy Schedule. . . . .	125
6.2	ASHRAE Thermal Sensation Scale. . . . .	127
6.3	HDH for Toronto for the Year of 2008. . . . .	129
6.4	Occupancy Schedule for Single-zone. . . . .	129
6.5	Heating Energy Consumption in MJ for a Heating System with a Condensing Furnace and ON-OFF Control. . . . .	133
6.6	Heating Energy Consumption in MJ for a Heating System with a Non-Condensing Furnace and ON-OFF Control. . . . .	136

6.7	Heating Energy Consumption in MJ for a Heating System with a Condensing Furnace, Controlled Blower Speed and Combustion Level. . . . .	139
6.8	Heating Energy Consumption in MJ for a Heating System with a Non-Condensing Furnace, Controlled Blower Speed and Combustion Level. . . . .	141

# List of Figures

1.1	Energy Use by Residential Sector. . . . .	6
1.2	Energy Use by Commercial Sector. . . . .	6
1.3	A Flowchart of the Research Methodology. . . . .	8
1.4	Block Diagram Representation of Single Zone Control. . . . .	11
1.5	Block Diagram Representation of Zoned Control Scheme. . . . .	11
1.6	Block Diagram Representation of Inferential Model Based Control Scheme. . . . .	13
2.1	Fuzzy Inference System. . . . .	20
2.2	ANFIS Architecture. . . . .	26
2.3	Grid Partitioning. . . . .	32
2.4	Initial MFs for $x_1$ , $x_2$ and $y$ . . . . .	33
2.5	Hybrid Learning Algorithm: One Epoch. . . . .	45
2.6	The Feedback System for Neuro-Fuzzy Model. . . . .	46
3.1	Real System vs. Soft Sensor Model. . . . .	49
3.2	Inferential Model Showing Input and Output Variables. . . . .	52
3.3	Experimental Data for Hot-Water Heating System (February 2000). . . . .	54
3.4	Experimental Data for Forced Warm-Air Heating System (January 2008). . . . .	56
3.5	ANFIS-GRID Model Structure. . . . .	57
3.6	ANFIS-SUB Model Structure. . . . .	58
3.7	Training and Testing Errors Obtained by the Neuro-fuzzy Model using Different Combinations of Number of MFs for Hot-Water Heating System. . . . .	59
3.8	Training and Testing Errors Obtained by the Neuro-fuzzy Model using Different Types of MFs for Hot-Water Heating System. . . . .	59
3.9	Training and Testing Errors Obtained by the Neuro-fuzzy Model using Different Number of MFs for Forced Warm-Air Heating System. . . . .	61
3.10	Training and Testing Errors Obtained by the Neuro-fuzzy Model using Different Types of MFs for Forced Warm-Air Heating System. . . . .	61
3.11	Training and Testing Errors Obtained by the Neuro-Fuzzy Model for Different Step Sizes for Hot-Water Heating System. . . . .	62
3.12	Training and Testing Errors Obtained by the Neuro-Fuzzy Model for Different Step Size Increase Rates for Hot-Water Heating System. . . . .	63

3.13	Training and Testing Errors Obtained by the Neuro-fuzzy Model for Different Step Size Decrease Rates for Hot-Water Heating System. . . . .	63
3.14	Training and Testing Errors Obtained by the Neuro-Fuzzy Model for Different Step Sizes for Forced Warm-Air Heating System. . . . .	64
3.15	Training and Testing Errors Obtained by the Neuro-Fuzzy Model for Different Step Size Increase Rates for Forced Warm-Air Heating System. . . . .	65
3.16	Training and Testing Errors Obtained by the Neuro-fuzzy Model for Different Step Size Decrease Rates for Forced Warm-Air Heating System. . . . .	65
3.17	Training Data (day 21 to day 27: February 2000). . . . .	68
3.18	Training Data and ANFIS-SUB Output after Training the Model. . . . .	69
3.19	Comparison of Experimental Output and ANFIS Output. . . . .	70
3.20	Average Air Temperature Estimation using ANFIS-GRID. . . . .	70
3.21	Training Results. . . . .	72
3.22	Training Error. . . . .	72
3.23	Testing Results without Feedback. . . . .	73
3.24	Testing Results with Feedback System. . . . .	74
3.25	Testing Error with Feedback System. . . . .	74
4.1	RMSE ( $^{\circ}\text{C}$ ) Values as Error Level in the Test Data Varies. . . . .	83
4.2	RMSE ( $^{\circ}\text{C}$ ) Values as Error Level in the Training Data Varies. . . . .	84
4.3	Training Data Selection. . . . .	87
4.4	Flow Chart for Training and Testing Process. . . . .	88
4.5	Schematic Diagram for Sensitivity Analysis. . . . .	90
4.6	Comparison of Measured and Estimated Temperatures for the Models Trained with TD1 for Hot-Water Heating System. . . . .	91
4.7	Comparison of Measured and Estimated Temperatures for the Models Trained with TD2 for Hot-Water Heating System. . . . .	92
4.8	Comparison of Measured and Estimated Temperatures for the Models Trained with TD3 for Hot-Water Heating System. . . . .	92
4.9	Comparison of Measured and Estimated Temperatures for the Models Trained with TD1 for Forced Warm-Air Heating System. . . . .	93
4.10	Comparison of Measured and Estimated Temperatures for the Models Trained with TD2 for Forced Warm-Air Heating System. . . . .	94
4.11	Comparison of Measured and Estimated Temperatures for the Models Trained with TD3 for Forced Warm-Air Heating System. . . . .	94
4.12	Schematic Diagram for Robustness Analysis. . . . .	95
4.13	Noisy Experimental Data for Hot-Water Heating System. . . . .	96
4.14	Noisy Experimental Data for Forced Warm-Air Heating System. . . . .	97
4.15	Comparison of Measured and Estimated Temperatures for the Models Trained with $T\hat{D}1$ for Hot-Water Heating System. . . . .	97
4.16	Comparison of Measured and Estimated Temperatures for the Models Trained with $T\hat{D}2$ for Hot-Water Heating System. . . . .	98

4.17	Comparison of Measured and Estimated Temperatures for the Models Trained with $T\acute{D}3$ for Hot-Water Heating System. . . . .	98
4.18	Comparison of Measured and Estimated Temperatures for the Models Trained with $T\acute{D}1$ for Forced Warm-Air Heating System. . . . .	99
4.19	Comparison of Measured and Estimated Temperatures for the Models Trained with $T\acute{D}2$ for Forced Warm-Air Heating System. . . . .	100
4.20	Comparison of Measured and Estimated Temperatures for the Models Trained with $T\acute{D}3$ for Forced Warm-Air Heating System. . . . .	100
5.1	Block Diagram for Thermostat Controlled Heating System. . . . .	107
5.2	Single Zone Control System. . . . .	108
5.3	Zoned Control System. . . . .	109
5.4	Controllability Over Room Temperature in the Dwellings with and without Programmable Thermostats. . . . .	119
5.5	Satisfaction Rating in the Dwellings with and without Programmable Thermostats. . . . .	120
6.1	Furnace Control Scheme using Average Air Temperature Estimator. . . . .	124
6.2	Desired Temperature Schedule for Multi-zone. . . . .	130
6.3	Desired Temperature Schedule for Single-zone. . . . .	131
6.4	Daily Average Air Temperature Profile in a Heating System with a Condensing Furnace and ON-OFF Control Logic for High Heating Load. . . . .	132
6.5	Daily Average Air Temperature Profile in a Heating System with a Condensing Furnace and ON-OFF Control Logic for Medium Heating Load. . . . .	132
6.6	Daily Average Air Temperature Profile in a Heating System with a Condensing Furnace and ON-OFF Control Logic for Low Heating Load. . . . .	133
6.7	Daily Average Air Temperature Profile in a Heating System with a Non-Condensing Furnace and ON-OFF Control Logic for High Heating Load. . . . .	134
6.8	Daily Average Air Temperature Profile in a Heating System with a Non-Condensing Furnace and ON-OFF Control Logic for Medium Heating Load. . . . .	135
6.9	Daily Average Air Temperature Profile in a Heating System with a Non-Condensing Furnace and ON-OFF Control Logic for Low Heating Load. . . . .	135
6.10	Daily Average Air Temperature Profile in a Heating System with a Condensing Furnace, Controlled Blower Speed and Combustion Level for High Heating Load. . . . .	137
6.11	Daily Average Air Temperature Profile in a Heating System with a Condensing Furnace, Controlled Blower Speed and Combustion Level for Medium Heating Load. . . . .	137
6.12	Daily Average Air Temperature Profile in a Heating System with a Condensing Furnace, Controlled Blower Speed and Combustion Level for Low Heating Load. . . . .	138
6.13	Daily Average Air Temperature Profile in a Heating System with a Non-Condensing Furnace, Controlled Blower Speed and Combustion Level for High Heating Load. . . . .	140

6.14	Daily Average Air Temperature Profile in a Heating System with a Non-Condensing Furnace, Controlled Blower Speed and Combustion Level for Medium Heating Load. . . . .	140
6.15	Daily Average Air Temperature Profile in a Heating System with a Non-Condensing Furnace, Controlled Blower Speed and Combustion Level for Low Heating Load. . . . .	141
6.16	Daily Proportion of Time the Temperature in not Maintained within Desirable Range for High Heating Load Conditions. . . . .	143
6.17	Daily Proportion of Time the Temperature in not Maintained within Desirable Range for Medium Heating Load Conditions. . . . .	143
6.18	Daily Proportion of Time the Temperature in not Maintained within Desirable Range for Low Heating Load Conditions. . . . .	144

# Chapter 1

## Introduction

This dissertation is concerned with the design and development of recurrent hybrid neuro-fuzzy based modelling environment and training algorithm for the inferential models. Further investigated is the application of the inferential model in the residential space heating systems.

Among the various methods employed in the design and modelling of inferential models, Adaptive Neuro-Fuzzy Inference System (ANFIS) is very effective and thus popular in a broad range of applications. However, limited by the feed-forward structure, it becomes very ineffective when used for the dynamic systems. To resolve this problem, a new algorithm, Recurrent Neuro-Fuzzy Inference System (RenFIS), is designed by integrating a feedback loop to the conventional ANFIS. RenFIS is then used to build an inferential model that estimates the average room temperature in multiple zone heating systems, as it is theoretically easy but practically expensive to measure this variable.

When RenFIS and the related methods are used to estimate/predict the variables of a dynamic system, there are two primary tasks to be completed: (1) structure identification and (2) parameter adjustment. The dissertation presents a method for determining the initial structure and the learning algorithm parameters for optimal adjustment through a training process. Experimental data obtained from two different types of heating systems: (1) Hot-Water Heating system (2) Forced Warm-Air Heating System, are used to validate the RenFIS algorithm and the training method. Sensitivity and robustness studies have been



carried out to investigate the issue of how the training data should be selected.

## 1.1 Objectives of the Study

Due to technical and/or economic difficulties, many process variables are difficult or impossible to measure directly using physical sensors. Soft sensors are designed to compute the value of such variables based on an inferential model of the targeted system and the value of the relevant variables that are easy to measure.

For instance, a robust method to measure the average air temperature representative of the temperature in all the zones, allows for optimal control of the furnaces in the residential heating system, saving energy and improving indoor environmental quality. Conventionally, the furnaces in a multi-zone heating system are controlled to maintain the desired temperature based on the temperature sensed at one point of the building. Considering the fact that all heating systems are built to produce a desired indoor thermal environment under the changing cold climate, it is reasonable to consider what benefit it could bring to the overall performance if the furnace in a multi-zone heating system is controlled according to a representative measurement of the room temperature in the whole building. As shown in later part of the dissertation, this is an appropriate approach to optimize the operation of the furnaces in multi-zone residential heating systems.

The objectives of this research are:

- To investigate the current control practices for space heating systems in Canada. Determine the potential for saving energy and improving indoor environment quality through incorporating an inferential control strategy to the conventional control methods.
- To develop an efficient and effective modelling environment for the inferential model. The possible modelling environments are: (1) Mathematical Modelling, (2) Model-free approaches such as: (a) Fuzzy Systems (b) Neural Networks (c) Hybrid Neuro-Fuzzy Systems.

- To develop a simple learning algorithm for training the inferential model structure. The space heating system is a dynamic system. The feed-forward approaches which are simple to train cannot effectively cope with dynamic environment, and the feed-back approaches which can model the dynamic systems are difficult to train.
- To design and develop an inferential control scheme for improving the performance of the buildings including thermal comfort and energy efficiency. The inferential control scheme can control the operation of the heating equipment (furnace) based on the average air temperature representative of the temperatures in all the zones.
- To test the performance of the inferential control scheme for a forced warm-air heating system in a simulation environment. A typical residential building is simulated. The performance of the control scheme is compared with the conventional and zoned control methods.

## 1.2 Background

### 1.2.1 Climate Change:

To address the issue of accumulation of the greenhouse gases (GHG) in the atmosphere, the United Nations Environment Programme and the World Metrological Organization in 1988 formed the Intergovernmental Panel on Climate Change (IPCC) tasked with providing the most up-to-date assessment of human effects on climate and human health. The most recent summary for policy makers, the IPCC fourth assessment report, notes that the continued accumulation of GHGs in the atmosphere is in fact leading to noticeable climate change. In terms of globally-averaged surface air temperature, 2005 and 1998 were the two warmest years in the instrumental temperature record and 1995-2007 is the warmest twelve year period (since 1850). The 100-year linear trend (1906-2005) of  $0.74$  [ $0.56$  to  $0.92$ ]  $^{\circ}\text{C}$  is larger than the corresponding trend of  $0.6$  [ $0.4$  to  $0.8$ ]  $^{\circ}\text{C}$  (1901-2000). The linear warming trend over the 50 years from 1956 to 2005 ( $0.13$  [ $0.10$  to  $0.16$ ]  $^{\circ}\text{C}$  per decade) is nearly twice that for the 100 years from 1906 to 2005 [1].

Carbon dioxide,  $CO_2$ , is the most important anthropogenic GHG. In the building sector,  $CO_2$  emissions are approximately 60% of all GHG emissions. Its annual emissions have grown between 1970 and 2004 by about 80%, from 21 to 38 gigatonnes (Gt), and all over represented 77% of total anthropogenic GHG emissions in 2004. The rate of growth of  $CO_2$  emissions was much higher during the 10-year period of 1995-2004 (0.92 GtCO<sub>2</sub> per year) than during the previous period of 1970-1994 (0.43 GtCO<sub>2</sub> per year). Global increase in  $CO_2$  concentrations are due primarily to fossil fuel use.

GHG emissions from the building sector have been increasing at almost 2 percent per year since 1990, and  $CO_2$  emissions from residential and commercial buildings are expected to continue to increase at a rate of 1.4 percent annually through 2025. These emissions come principally from the generation and transmission of electricity used in the buildings, which account for 71 percent of the total. Based on the energy usage, opportunities to reduce GHG emissions appear to be most significant for space heating, air conditioning, lighting, and water heating (United Nations Statistics Division).

In a nut-shell, with only 0.5% of world's population, Canada produces 2.3% of global carbon emissions. A similar pattern is also exhibited in many of the European countries. Buildings emissions are approximately 25% of the total global  $CO_2$  emissions from all sectors [2].

### **1.2.2 Indoor Environment/Thermal Comfort**

American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE) define thermal comfort for one person, as that condition of the mind which expresses satisfaction with the thermal environment. Fanger [3] pointed out if a group of people is subjected to the same room climate, it will not be possible, due to biological variance, to satisfy everyone at the same time; one must then aim at creating an optimal thermal comfort for the group, i.e., a condition in which highest possible percentage of the group is in thermal comfort.

Statistics Canada reports show that Canadian people spend approximately 65.9% of their time in their own houses [4]. Effort should be made to create a space of relaxation and

console. Thermal discomfort develops stress for the occupants, affecting their productivity. A study by Federspiel investigated thermal comfort through complaints in a survey of 23,500 occupants conducted in the commercial buildings [5]. The results found most complaints to be related to temperature extremes, either being too hot or being too cold. In fact, 96.5% occurred at temperature outside of the range 21.1°C-23.8°C. Therefore, comfortable temperature conditions were defined as being within this range. In addition, ASHRAE defines comfortable conditions as a temperature not changing more than 1.8°C in thirty minutes. The ASHRAE standard 55 also specifies that a comfort environment must be maintained 80% of the time and required temperature range for summer and winter seasons is given in Table 1.1 [6].

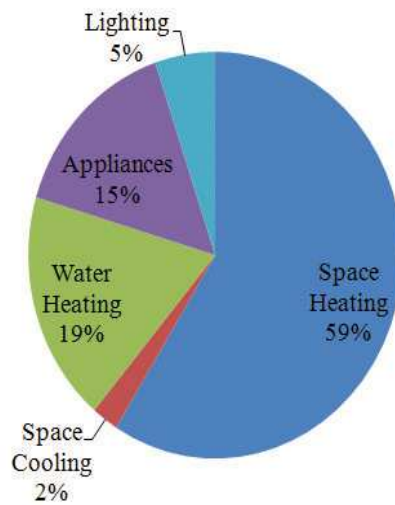
**Table 1.1:** ASHRAE-55 Standard for Thermal Comfort.

Conditions	Relative Humidity	Operating Temperature (°C)
Summer	If 30% then	24.5-28
	If 60% then	23-25.5
Winter	If 30% then	20.5-25.5
	If 60% then	20-24

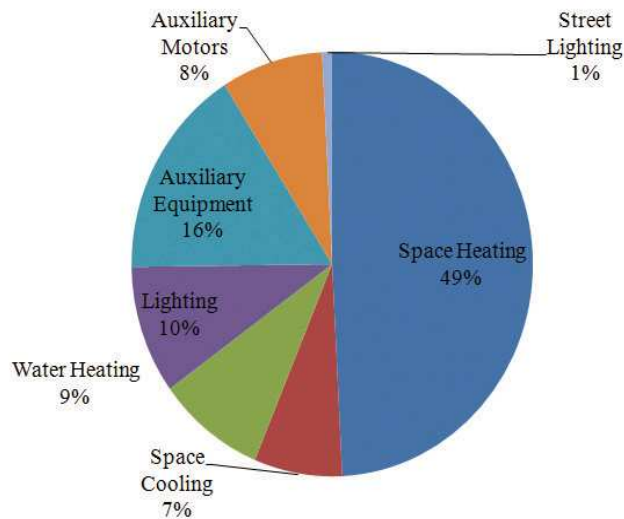
### 1.2.3 Energy

Heating systems consume a substantial proportion of the total energy used for all buildings operations. With growing environmental and energy consciousness, consuming energy more efficiently is a priority. As in Canada approximately one third of its energy use is due to the operation of the buildings, any improvement in building energy usage will contribute to the global energy conservation effort. Fig. 1.1 and Fig. 1.2 show that energy use for space heating is 59% and 50% for residential and commercial buildings respectively [7].

Together with the increasing concerns of poor indoor environment quality to the health conditions, exploring ways in improving the energy efficiency as well as improving comfort in the buildings are important. The energy efficiency of a heating system depends on a lot of factors, such as the thermal performance of the building envelope, the energy efficiency of the heating and distribution systems, the performance of control systems. The effort made



**Figure 1.1:** Energy Use by Residential Sector.



**Figure 1.2:** Energy Use by Commercial Sector.

to improve the energy efficiency of heating systems has been overwhelmingly focused on:

- Improving thermal insulation of the building envelope and air tightness.
- Improving efficiency of furnaces, i.e. replacing non-condensing furnaces by condensing furnaces or using 2-stage furnaces.

- Utilizing solar energy and other renewable energy.
- Using self-programmable thermostats with temperature set back capabilities.

Despite the enormous effort made over the last decades to improve the energy efficiency of these heating systems, a huge potential for further energy saving still persists. However, not enough effort has been put into improving the control of furnaces, as a result, the overall energy efficiency of heating systems is compromised due to the operation of these heating devices.

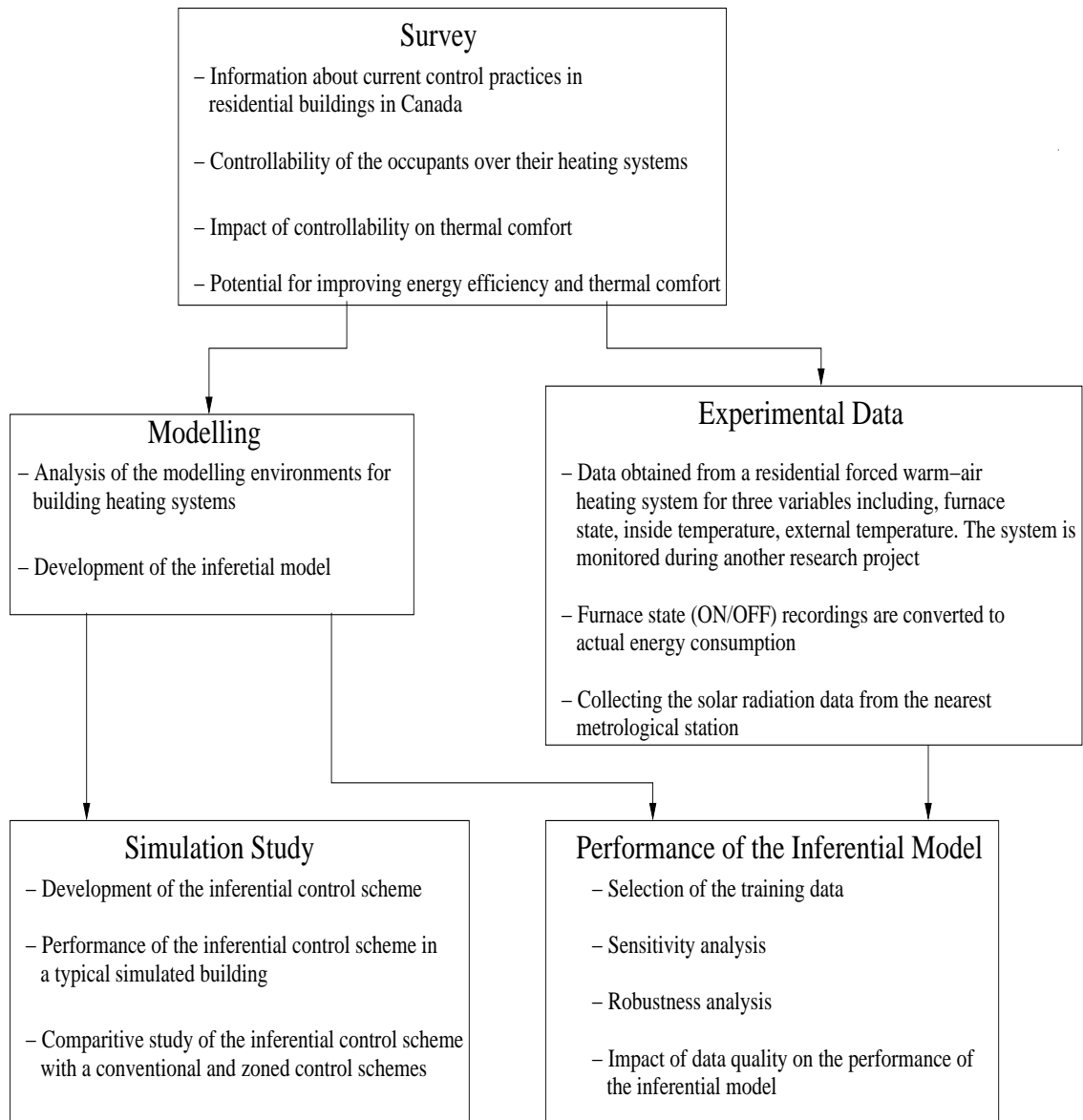
This dissertation presents a method for controlling the operation of the furnace. Inferential sensor models estimate the representative air temperature in the whole building. Based on the output of the estimator the blower speed and the combustion level are controlled, giving an efficient operation of the furnace.

## 1.3 Methodology

Fig. 1.3 illustrates the activities involved in the completion of this research work. Each text box shows a major research task and the specific output. The arrowed lines show the relationship between the activities. They are briefly explained as follows.

### 1.3.1 Survey

A survey on the operation of forced warm-air heating systems in residential buildings in Ontario, Canada was conducted by the author. The survey was aimed at finding the problems associated with the current control of forced warm-air heating systems to identify the potential for improvements. The survey was started with collecting preliminary responses and then an online system was used. The survey results indicated a problem with the control of the conventional heating systems. The survey results are analyzed to check the potential of correlation between the comfort level and the occupant's access to the control over room temperature. Respondents with higher heating system controllability tend to feel more thermally comfortable. In multi-zone heating systems, the survey results show that 36% of the



**Figure 1.3:** A Flowchart of the Research Methodology.

respondents experienced overheating in the winter.

### 1.3.2 Modelling

Inferential sensing is an attractive approach for modelling the behavior of the dynamic processes that require robust control. An inferential sensor model has been developed in this thesis for estimating the average air temperature of multiple-zone buildings, which is

used to optimize control of the heating systems. The inferential sensor model trained with the available short-term measurements can be used to predict the long-term performance of building heating systems.

To design a reliable inferential sensor, of fundamental importance is to build a dynamic model of the system to be controlled that is simple and robust. The methods used for such simplified dynamic model can be categorized into two groups: simplified physical model and black-box model based on advanced mathematical algorithms. This dissertation designs an innovative algorithm that is suitable for the development of a robust black-box model. The algorithm is derived from ANFIS and is referred to as RenFIS. The approach is especially useful for large complex and vague systems, which cannot be defined or represented efficiently by simple methods. Thus, the approach is ideally suited for investigating the complex control problems of building heating systems.

### **1.3.3 Experimental Data**

For training and validation of the inferential model, it is extremely important to collect experimental data under actual building conditions. The experimental data is obtained from two types of heating systems: (1) hot-water heating system and (2) forced warm-air heating system, monitored under different research projects. Hot-water heating system data is collected in a laboratory heating system and forced warm-air heating system data is collected in a residential building. The data collected include, indoor air temperature, exterior temperature, energy used by the heating system, and solar radiation.

### **1.3.4 Performance of the Inferential Model**

An experimental study is conducted to analyze the impact of data quality on the estimation ability of inferential model. The errors in the training data as well as in the testing data reduce the predictive accuracy of the model.

Sensitivity analysis aims at validating the selection of training data. ANFIS and RenFIS, both are sensitive to the selection of training data. Training data sets collected under different



conditions are used to check the effect on the predictive accuracy of the models.

### **1.3.5 Simulation Study**

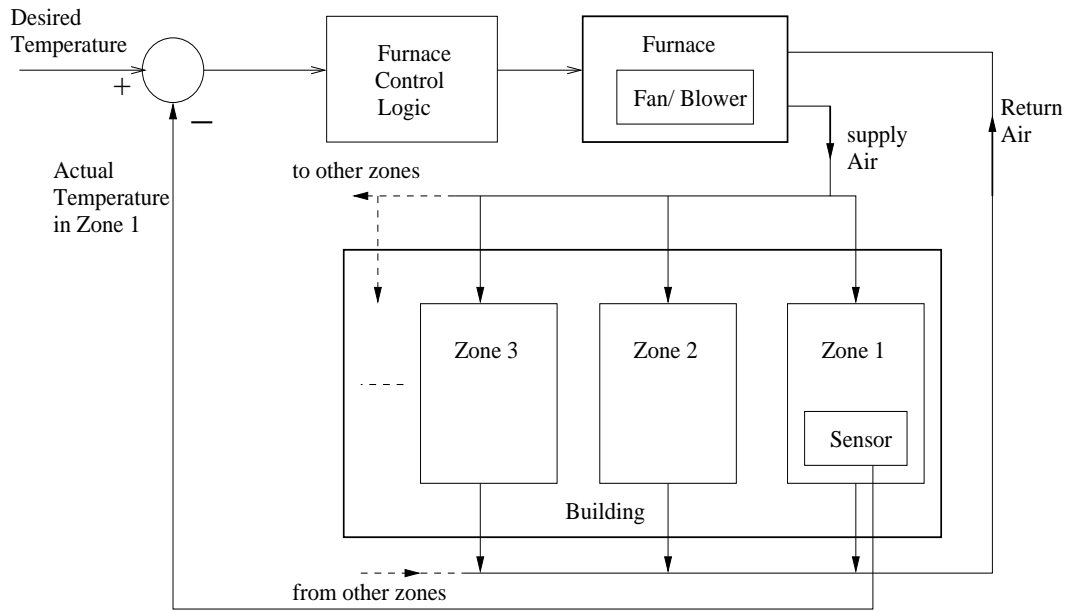
The inferential model is used in the development of an inferential control scheme for optimizing the control of the heating systems. The performance of the scheme is analyzed for a simulated typical residential building. The performance is compared with conventional and zoned control methods to check their potential for improving thermal comfort and energy efficiency.

## **1.4 The Potential for Improving Energy Efficiency and Thermal Comfort**

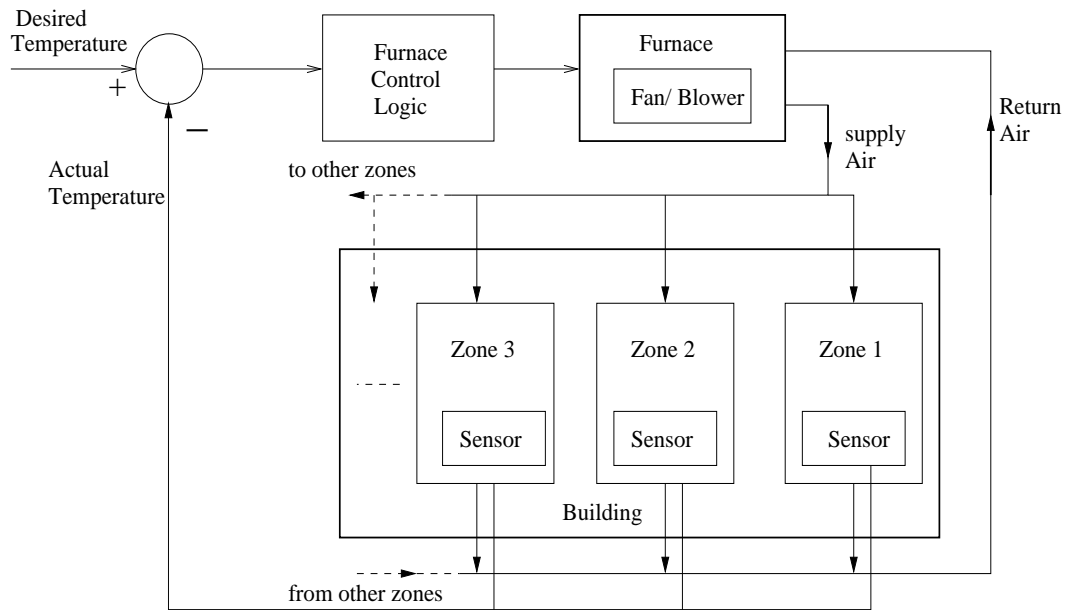
A conventional residential forced warm-air heating system comprises a furnace in which fuel is burned to generate warm-air, the distribution system which supplies the warm-air to the space being conditioned and returns the indoor air back to the furnace, and the air registers that diffuse warm-air to the occupied indoor spaces. In current practice, the control of such heating systems is very problematic. The major equipment, or the furnace, is controlled by a control system that only measures the air temperature at one place in the building, as shown in Fig. 1.4. Although the control system allows flexible scheduling of temperature set-point to accommodate different occupancy schemes, the building is basically treated as a single zone system. Consequently, some rooms in the building are either overheated or unnecessarily heated while not occupied, while some rooms are under-heated due to insufficient heating capacity that result from overheating in other rooms.

Zoned control methods are employed to overcome the above said problem. As shown in Fig. 1.5, the furnace is controlled by a control system that measures the temperature in each zone individually and various additional mechanisms are used in the heating capacity delivered to different zones. This method is expensive and invasive to retrofit.

Similarly for water heating systems, overall energy efficiency of the heating systems is compromised due to improper operation of the boilers. Boilers in the heating systems are



**Figure 1.4:** Block Diagram Representation of Single Zone Control.



**Figure 1.5:** Block Diagram Representation of Zoned Control Scheme.

often considerably oversized in order to provide a substantial margin of capacity [8]. This is because the boilers in heating systems are normally controlled to maintain the supply water temperature, rather than the building air temperature, at a certain set-point, which

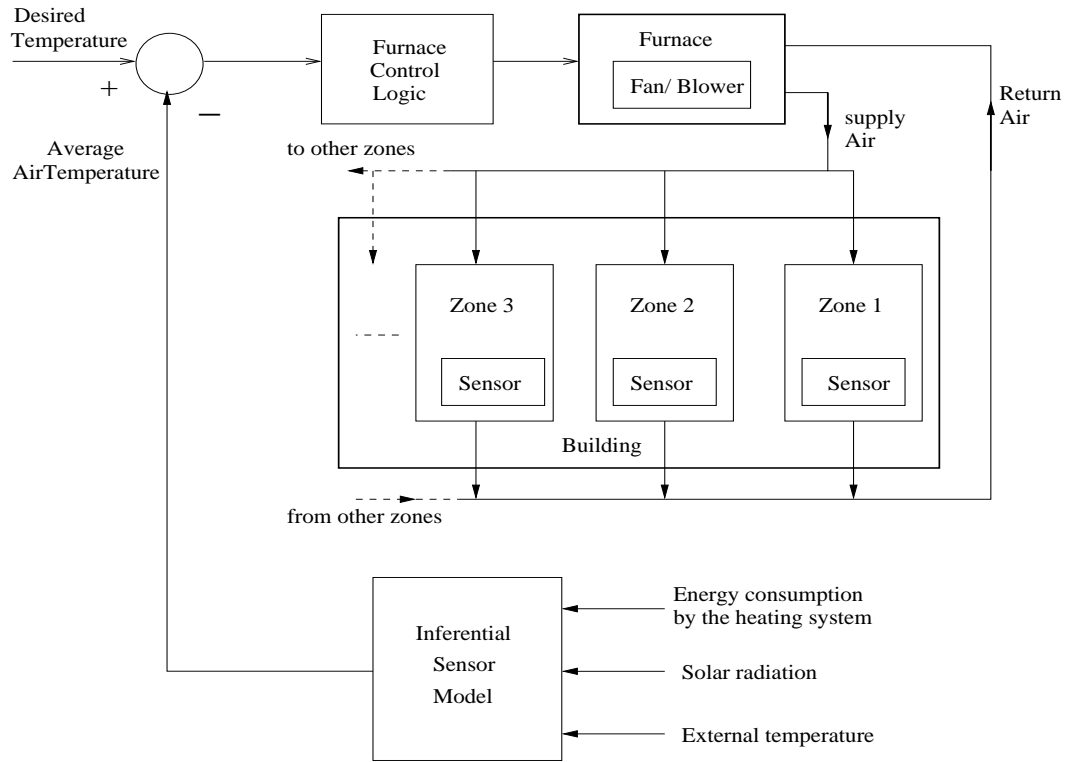
may be a predefined schedule with variable values at different periods of the day. Liao and Dexter [9] have implemented a feedback loop to detect the overall level of the thermal comfort in the building and transferred this information to the water temperature scheduler that can determine the optimal set-point of the water temperature.

The survey conducted by the author shows that approximately 20% of the respondents feel dissatisfied with room temperature in the winter. The results indicate a problem with the control of the conventional heating and cooling systems. There exists a correlation between the comfort level and the occupant's access to the control over room temperature. Respondents with higher heating system controllability tend to feel more thermally comfortable. In multi-zone heating systems, the survey results show that 36% and 27% of the respondents experienced overheating in the winter and overcooling in the summer. In the same house there are places with temperature higher or lower than the desired level. This is because the thermostat only sense the temperature of one zone in which it is installed. A more representative measurement of average air temperature in the building is essential for optimal control of the relevant equipment as the output of a physical sensor is unlikely to be representative for the room being measured.

As it is not practical to directly measure the average temperature in the residential buildings, soft sensing is applicable. Fig. 1.6 presents a block diagram representation of an inferential sensor model based control method for controlling the operation of the furnace. The output of the inferential sensor can be used to control the speed of the blower and the level of the combustion in the furnace for the optimal efficiency.

## 1.5 Modelling Environment

Soft sensing allows process parameters, difficult to measure, to be inferred from other easily made measurements. Inferential model is an important module of a soft sensor, which represents non-linear dynamic behavior of the system under control. The performance of the soft sensor depends on the structure of the inferential model and on how it is configured, e.g. the values of the relevant parameters. Section 1.5 discuss the possible modelling environments



**Figure 1.6:** Block Diagram Representation of Inferential Model Based Control Scheme.

for inferential models.

Different mathematical techniques have been adapted to develop the dynamic models for use in soft sensors. These techniques include:

- Physical model [10].
- Grey-box modelling [11,12].
- Neural networks [13].
- Fuzzy logic systems [14].
- Hybrid neuro-fuzzy systems [15].

### Why Neuro-Fuzzy Systems?

A simplified physical model of multiple-zone space heating system was developed and incorporated with conventional boiler control system to design an Inferential Control Scheme [10].

Physical sensor and the inferential control scheme are difficult to commission, as these intend to estimate the time-dependent value of the average air temperature in the building.

A conventional control method defines a process with a real-world orientation and tries to build a mathematical model to represent that process. The resultant model could not be necessarily approximated to a linear model. Although some linear model can be made, this linear model might only represent some extreme conditions of the nonlinear model. A fuzzy control method can perform much better than conventional control devices under extreme conditions. Verbruggen and Bruijn performed a comparison between conventional and fuzzy control. The authors stated that the advanced conventional control methods only partly satisfy the demands of highly non-linear system behavior. Furthermore, it stated the contribution of fuzzy control for non-linear dynamic systems is as of the availability of a qualitative operator, so that design knowledge can be easily implemented [16].

“Model-free” control methods do not require implicit modelling, which eliminates the system characterization and identification processes, and thus there is no need to allocate time and resources to determine an adequate model of the system under control and to evaluate its validity. Conventional control also has some model-free approaches, such as nonlinear adaptive control and Proportional Integration and Derivative (PID) control [17]. Fuzzy control provides another model-free approach. The conventional control plants depend on precise models. Fuzzy control uses complicated structures and evaluation standards from conventional control and it uses heuristic knowledge to deal with the control system behavior.

The use of fuzzy-logic based control (FLBC) is motivated by its ability to capture the qualitative control strategies and implement highly flexible control behavior. Fuzzy logic and neural networks are the mathematical techniques that allow for model-free control design. They can be used to model dynamic systems. They share the ability to improve the intelligence of systems working in uncertain, imprecise, and noisy environments. Fuzzy systems and neural networks estimate sampled functions and behave as associative memories. Both have an advantage over traditional statistical estimation and adaptive control approaches to function estimation. They estimate a function without requiring a mathematical descrip-

tion of how the output functionally depends on the input, instead they can learn from the historical data recorded during the operation of the system.

Because fuzzy control algorithms emulate human control strategy, their principles are easy to understand. For conventional control theory, more and more advanced mathematical tools are required to model the non-linear dynamic system behavior, which are not as easy to understand as the fuzzy control methods are. Therefore, practical engineers who are the front-line of designing consumer products tend to use the approaches that are simple and easy to understand and implement. Fuzzy control is just such an approach. From a practical point of view, the development cost is the most important criterion for a successful product for modern technology. Wang [18] pointed out that fuzzy control is easy to understand, and the time to learn the approach is short; that is, the “software cost” is low. Also, because fuzzy control is simple to implement, the “hardware cost” is also low. Additionally, there are software tools (for instance: fuzzy logic tool available for designing fuzzy controllers). Thus, fuzzy control is an approach that has a high performance/cost ratio.

Lin and Lee [19] established that neural networks have a large number of highly interconnected processing elements, which demonstrate the ability to learn and generalize from training patterns or data. Fuzzy systems estimate functions with fuzzy samples, while neural networks use numerical data samples.

### **Why ANFIS?**

ANFIS, which contains fuzzy IF-THEN rules, has the ability to deal with nonlinear control problems [20]. A typical ANFIS network is built through three major steps. The first is to construct a set of fuzzy IF-THEN rules with corresponding membership functions to generate the I/O pairs. The next step is to build fuzzy IF-THEN rules and fuzzy inference systems. The last is to create the structures and learning rules of these adaptive networks.

Slotine and Li [21] pointed out that nonlinearities frequently cause undesirable behavior of the control system. However, most real life control problems are of the non-linear nature. System modelling is the part of the design needs major change to model nonlinear

control systems. Basically, system modelling is based on conventional tools (e.g. differential equations). Sometimes, we can solve differential equations by using mathematical tools. However, most of the time we cannot properly formulate the necessary differential equations to represent the nonlinear control systems. So, for an ill-defined and uncertain sophisticated nonlinear system, the differential equation representation is not suitable. Some nonlinear control problems include linguistic variables, for instance human descriptive or vague intuitive thinking, so such problems become more complicated and difficult to solve. At this time, the conventional control techniques are only good for simpler problems, and cannot be applied to complex and dynamic control problems. A more robust control theory is required to solve the problems of the more complicated nonlinear control systems.

This research has concentrated on integrated neuro-fuzzy technology as a modelling environment to cope with these problems of physical models, neural networks and fuzzy logic approaches. Detailed discussion about these model free techniques is presented in Chapter 2.

## 1.6 Structure of the Dissertation

This dissertation has begun with a statement of the objectives of the study, a review of the background, and the methodology employed. The review of the background has been presented to justify the objectives of the study. The methodology has been described to explain how the objectives can be achieved.

The remainder of this dissertation is organized into six chapters as following:

- Chapter 2 provides the basis of Hybrid Neuro-Fuzzy Systems. The structure generation & learning and parameter learning methods are discussed.
- Chapter 3 presents the development of the inferential model for estimating the average air temperature.
- Chapter 4 deals with the issues of data quality and how they impact on the performance of the model. Sensitivity and robustness analysis are conducted for the training data.

- Chapter 5 presents a detailed review about the current control practices for the residential forced warm-air heating systems. This chapter analyzes the scope of implementation of the proposed control strategy for improving the energy efficiency and thermal comfort.
- Chapter 6 discusses the control scheme making use of the inferential model to control the built environment. Performance of different control methods including, conventional, zoned and inferential control, is compared.
- Chapter 7 draws conclusions and makes suggestions for future work.



# Chapter 2

## Principle of Hybrid Neuro-Fuzzy Systems

### 2.1 Introduction

Over the last few decades, significant advances have been made in the areas of fuzzy logic and neural networks [22]. Fuzzy logic and neural networks are complimentary tools in the design of modern control systems. The goal of neural fuzzy systems is to provide fuzzy systems with automatic learning ability. The use of neural networks in learning membership functions (MF)s and fuzzy rules of fuzzy systems is its major benefit. Neural network learning techniques can reduce the time and cost while improving the performance of fuzzy systems. On the other hand, the training of neural fuzzy systems is generally more efficient than the traditional neural networks [19, 23].

### 2.2 Fuzzy Logic Systems

A fuzzy logic system (FLS) is a convenient way of mapping input space to output space. The mapping is achieved by transforming the input information from numerical domain (crisp domain) to linguistic domain (fuzzy domain). The transformed information is then processed using the fuzzy rules incorporated in the rule-base to yield a fuzzy output which, in turn, has to be transformed again into the numerical domain [22, 24, 25].

The main component of any fuzzy logic system is its rule-base. The rule consists of two

parts: an antecedent (premise) and a consequent. They are English-like sentences in the form of IF-THEN. The IF-part (antecedent) includes the condition for firing a rule, and the THEN-part (consequent) includes the action. A typical fuzzy rule has an expression of the form:

$$IF\ x_1\ is\ A_1\ AND\ x_2\ is\ A_2\ THEN\ y\ is\ B \quad (2.1)$$

where  $A_1$ ,  $A_2$ , and  $B$  are linguistic values defined by fuzzy sets on certain universe of discourse. A fuzzy model consisting of this form of fuzzy rules is generally called Mamdani fuzzy model, if  $B$  is a fuzzy singleton [26].

Another form of IF-THEN rules, proposed by Takagi and Sugeno [27, 28], has only fuzzy set defined in the premise part; the consequent part is described by a non-fuzzy equation of the input variables,

$$IF\ x_1\ is\ A_1\ AND\ x_2\ is\ A_2\ THEN\ y\ is\ f(x_1, x_2) \quad (2.2)$$

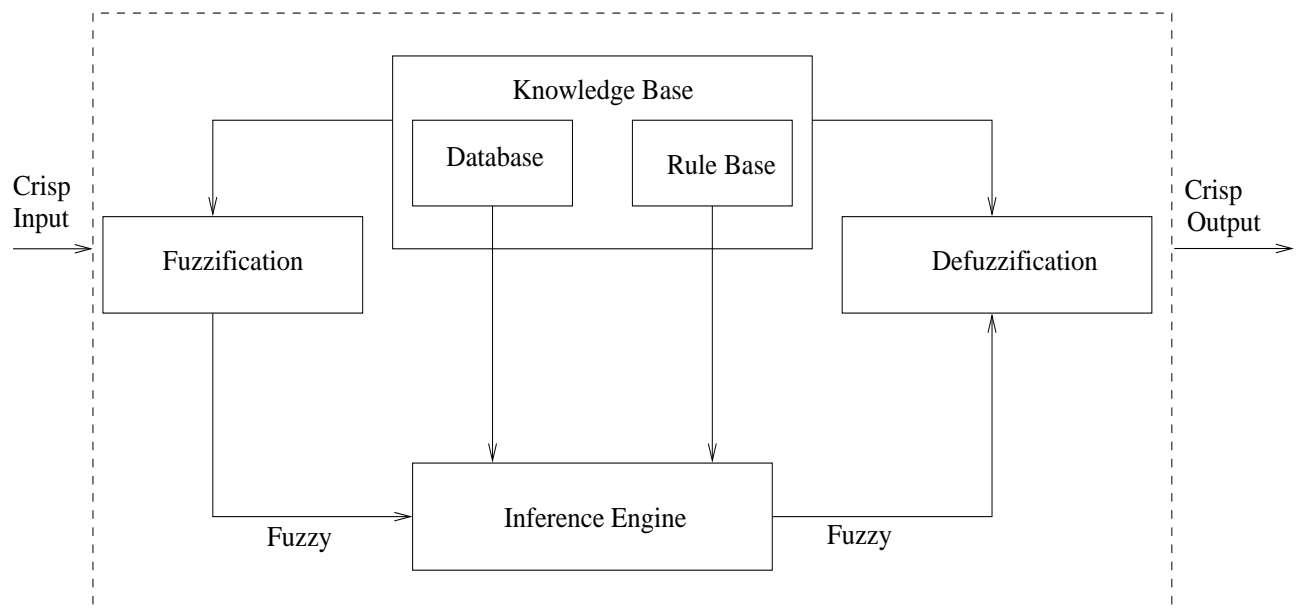
A fuzzy model consisting of these type of rules is called a TSK fuzzy model.

Traditionally, the action takes place if and only if the degree of belief for the antecedent is equal to 1 (the antecedent part is satisfied). Otherwise, if the degree of belief for the antecedent is 0 (in a bi-valued logic sense, the degree of truth is either 1 or 0), no action would be performed. In contrast to conventional rule-based systems, in FLSs, all rules are fired regardless of the value of their antecedents, and the impacts on the overall performance of FLS depends on their degree of belief. Fuzzy IF-THEN rules form the core of a fuzzy inference system, which is discussed in the next section.

## 2.3 Fuzzy Inference System

Fuzzy inference system (FIS) is a powerful framework based on fuzzy set theory, fuzzy IF-THEN rules, and fuzzy reasoning. It is also known as fuzzy model or fuzzy rule-based system. The basic structure of a fuzzy inference system is shown in Fig. 2.1, which consists of the following subsystems:

- a fuzzifier which maps from an observed input space to labels of fuzzy sets in a specified universe of discourse,
- a rule base containing a set of fuzzy IF-THEN rules,
- a database which defines the MFs used in fuzzy rules,
- an inference engine which performs the reasoning procedure upon the rules and the given conditions to derive reasoning conclusions,
- a defuzzifier which maps the fuzzy inference results back to crisp outputs.



**Figure 2.1:** Fuzzy Inference System.

In general, the steps of fuzzy inferencing are summarized as follows:

- *Fuzzification*: compare the input variables with the MF on the premise part to compute the membership values of each linguistic variable.
- *Application of T-norm*: this operator performs algebraic product or min on the membership values of the premise part to get firing strength or weight for each rule.

- *Generation of the consequent parts:* it can be fuzzy or crisp.
- *Defuzzification:* aggregate the qualified consequents to derive a crisp output.

## 2.4 Neural Networks

Neural Networks are inspired by modelling networks of biological neurons in the human brain. However, these biological models do not necessarily resemble their original biological counterparts. Thus, in literature, they are often called Artificial Neural Networks (ANN) to distinguish them from the biological ones. In a rough analogy, neural networks have a large number of interconnected processing elements (nodes) which usually operate in parallel. The collective behavior of a neural network, acting like a human brain, has the ability to learn, recall, and generalize from sampled data.

A typical biological neuron consists of three basic parts: soma (cell body), dendrites and the axon. The actual operation of a neuron is very complicated, but the basic features can be summarized as follows. The input signals from other neurons are transmitted chemically through either excitatory or inhibitory synapses to the dendrites. Excitatory synapses raise the probability of the neuron firing, while inhibitory synapses suppress it. The dendrites pass the signals into soma which sums the signals both spatially and temporally. If the summation exceeds certain threshold, a pulse or action potential of fixed strength and duration is sent out through the axon to other neurons. After firing, the neuron must recover before it can fire again. The fastest switching times are known to be in the order of  $10^{-3}$  seconds, which is quite slow compared to the computer switching speed of  $10^{-10}$  seconds. Yet, humans are able to make complex decisions surprisingly quick. This observation has led to speculate that the information processing ability must result from some highly parallel processes operating on representation that are large amount of neurons. One motivation for ANNs is to capture this parallel computational feature based on distributed representation.

A neural network is composed of a set of highly interconnected nodes. Each node collects the inputs from its input connections, performs a set of predefined mathematical operations,

and produces a single output. Inside each node, there is an integration action which serves to combine information or activation from external environment or other nodes into a net input. Another action of a node is to produce an output as a function of its net input through an activation function. Each node has associated weights that determine the exact or the incoming inputs on the activation of the node. The weights may be positive, or negative. Finally, through a learning procedure, the information can be stored in these weights in a distributed manner.

To characterize a neural network, the structure that organizes the nodes and the connection geometry should be specified. There are a variety of network structures, each of which results in different computational properties. The main distinction is between feed-forward and recurrent networks. In a feed-forward network, the weight links are unidirectional, and there is no node connecting to another node in the same layer or the preceding layer. When outputs are directed back as inputs to the nodes in the same layer or the preceding layer, the network is a feedback network. Feedback networks that have closed loops are called recurrent networks. Another important feature in specifying a neural network is the learning scheme. Complete learning involves two kinds of learning tasks: (1) structural learning, which focuses on the change in the network structure, for example, the number of nodes and the corresponding connection structure, (2) parametric learning, which is concerned about updating the weights. Parameter learning schemes can be classified into three categories: supervised learning, unsupervised learning, and reinforcement learning.

In summary, neural networks demonstrate the following properties:

- *Nonlinear mapping ability:* Neural networks can learn to approximate any continuous nonlinear input-output relation over a bounded domain from the sampled data.
- *Generalization ability:* Neural networks have the ability to generalize to situations that are different from the training data and fit the input-output mapping satisfactorily [29, 30].
- *Adaptability:* A neural network is a non-programmable dynamic system with capa-

bilities such as trainability and adaptivity that can be trained to store, process and retrieve information. With an on-line learning algorithm, neural networks can adjust the weights, even the network structures to optimize their performance.

- *Fault tolerance:* Since structured in a distributed manner, each node acts independently of the others, and relies only on local information. Large number of weight connections also give neural networks redundancy. These features provide the neural networks with inherent fault tolerance capability.

With these properties, neural networks provide a robust approach to approximating analytical functions. Neural networks can be considered as simplified mathematical models of brain like systems and they operate as parallel distributed computing networks. For certain types of problems, such as learning and generalizing to interpret sensory data, neural networks are one of the most effective learning approaches currently known. So, neural networks are well-suited to problems where nonlinear mapping must be acquired from data, and where a near-optimal solution is required in a short time.

## 2.5 Hybrid Neuro-Fuzzy Systems

A FIS can utilize human expertise by storing its essential components in the rule base and MFs. The fuzzified input data are matched with the premise parts in the IF-THEN rules and the conclusion of each rule is computed by a fuzzy inference engine. Finally, the appropriate outputs are generated through the defuzzification procedure. However, due to the vagueness and subjectivity of linguistic statements, fuzzy systems based on qualitative alone are unlikely to adequately model real systems. To circumvent such inadequacies, as in conventional approaches, available data should be used to validate the model's behavior.

On the other hand, a neural network does not rely on human experience. Instead, it employs learning procedures and a given training data set to evolve a set of parameters (weights) such that the desired functional behavior is achieved. Moreover, neural networks possess several properties such as disturbance-tolerance, and generalization abilities, which

result from their connectionist structure. However, it is difficult to extract structural knowledge from either the weights or the configuration of a neural network because the knowledge of the system is distributed into the network as synaptic weights. Also, since the hidden layers of the neural networks are opaque to the users, determination of the structure and the size of networks is still a problem to be resolved. Fortunately, it is possible to build prior knowledge about the training data into the weights or configuration of a neural network such that the efforts on the subsequent learning can be much reduced.

From the above description, it can be seen that fuzzy systems and neural networks are complementary technologies. While neural networks extract knowledge from systems to be approximated or controlled, fuzzy systems often use linguistic information from experts. A promising approach to reap their benefits and overcome their respective short-comings is to integrate them together in an appropriate way. The integrated systems should have low-level learning ability and computational power as well as high-level human like thinking and reasoning.

The synergism of fuzzy logic and neural network has produced a functional system capable of learning, high-level of thinking and reasoning. It is an improved tool for determining the behavior of imprecisely-defined complex dynamical systems. The neuro-fuzzy systems can combine the benefits of these two powerful paradigms into a single capsule. Considerable work has been performed to integrate the excellent learning capability of neural networks with FIS for deriving the initial rules of a fuzzy system and tuning the MFs [19, 20, 31–34]. The features exhibited in these systems, such as fast and accurate learning, good generalization capabilities, excellent explanation facilities in the form of semantically meaningful fuzzy rules, and the ability to accommodate both data and existing expert knowledge, make them suitable for a wide range of engineering and scientific applications [35].

Generally speaking, the merging of these two technologies can be characterized into the following categories:

- *Neural fuzzy systems:* adopting the neural learning ability to enhance the system fuzzy performance.

- *Fuzzy neural systems*: fuzzified version of conventional network models.
- *Hybrid neuro-fuzzy systems*: incorporation of fuzzy systems and neural networks into hybrid architectures.

In the first approach, neural fuzzy systems are constructed to provide fuzzy systems with automatic tuning mechanisms without altering their functionality. Since neural fuzzy systems are inherently fuzzy logic systems, most applications are found in control applications. In the second approach, fuzzy neural networks retain the basic properties and structures of neural networks but fuzzify some of their elements, where domain knowledge is formalized in terms of fuzzy sets and is applied to enhance the learning performance or augment their interpretation capability. Since the neural net architecture is preserved, they are mostly found in pattern recognition applications. In the third approach, neuro-fuzzy systems, fuzzy systems and neural networks have equal importance and serve different purposes. They are integrated to form a hybrid system. Utilizing their individual strength, they complement each other to achieve more efficient training. The structure of such hybrid systems are flexible and application oriented.

### 2.5.1 Adaptive Neuro-Fuzzy Inference Systems

Hybridization technology of neuro-fuzzy systems has resulted in the development of intelligent systems through several approaches such as GARIC, ANFIS, FUN, NEFCON, FALCON, SOFIN, FINEST, EFuNN, dmEFuNN, the evolutionary design of neuro-fuzzy systems and many others [15, 36–42]. Among these mentioned methods that fully optimize the complementary strength of ANN and FIS is the ANFIS. In comparison with the others, ANFIS has high speed of training, the most effective learning algorithm, and is simple in terms of the structure [43]. It is also regarded as one of the best in function approximation among other neuro-fuzzy models [44]. ANFIS is faster in convergence when compared to the other neuro-fuzzy models [45]; furthermore, ANFIS provides better results when applied without any pre-training [46].



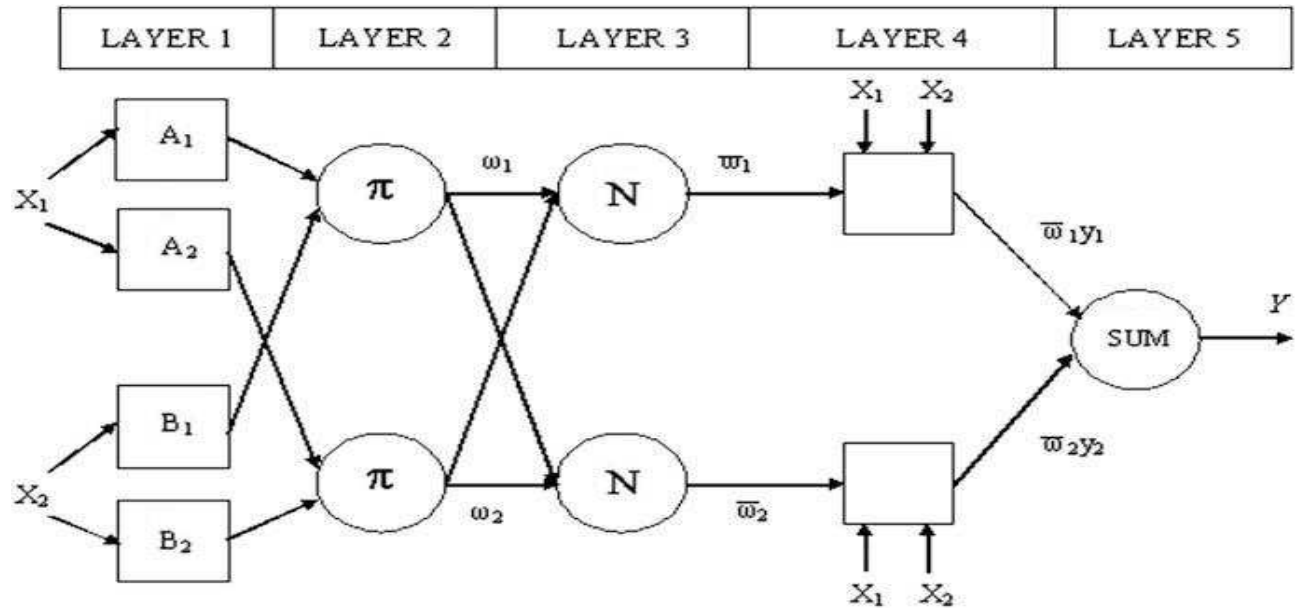
## 2.5.2 ANFIS Architecture

ANFIS is a multi-layer adaptive network-based fuzzy inference system. ANFIS consists of a total of five layers to implement different node functions to learn and tune the parameters in a FIS structure using a hybrid learning mode. To describe the architecture of ANFIS, let us consider a simple fuzzy system with two inputs ( $x_1, x_2$ ) and one output,  $y$ . Suppose that the rule base contains two fuzzy IF-THEN rules of the TSK type:

$$R_1 : \text{IF } x_1 \text{ is } A_1 \text{ AND } x_2 \text{ is } B_1 \text{ THEN } y_1 = p_1x_1 + q_1x_2 + r_1 \quad (2.3)$$

$$R_2 : \text{IF } x_1 \text{ is } A_2 \text{ AND } x_2 \text{ is } B_2 \text{ THEN } y_2 = p_2x_1 + q_2x_2 + r_2 \quad (2.4)$$

Fig. 2.2 shows a two-input, one-output ANFIS architecture. The functionality of nodes in ANFIS can be summarized as follows:



**Figure 2.2:** ANFIS Architecture.

- *Layer 1:* Nodes are adaptive; MFs of input variables are used as node functions, and

parameters in this layer are referred to as antecedent or premise parameters. The output of each node is defined by

$$O_{1,i} = \mu_{A_i}(x_1), \quad i=1,2 \quad (2.5)$$

$$O_{1,i} = \mu_{B_{i-2}}(x_2), \quad i=3,4 \quad (2.6)$$

where  $x_1, x_2$  are inputs to the fuzzy system, and  $A_i$  or  $B_i$  is a fuzzy set associated with the node. In other words, outputs of this layer are the MF values of the premise parts. In this illustrative case,  $A_i$  or  $B_i$  is described by the generalized bell-shaped function,

$$\mu_{A_i}(x) = \frac{1}{1 + \left(\frac{x-c_i}{a_i}\right)^{b_i}} \quad (2.7)$$

- *Layer 2:* Nodes are fixed with outputs representing the firing strengths of the rules. Each node performs the fuzzy AND operation using algebraic product,

$$O_{2,i} = \omega_i = \mu_{A_i}(x_1) \times \mu_{B_i}(x_2), \quad i=1,2 \quad (2.8)$$

- *Layer 3:* Nodes are fixed with outputs representing normalized firing strengths. The  $i$ th node calculates the normalized firing strength,

$$O_{3,i} = \bar{\omega}_i = \frac{\omega_i}{\omega_1 + \omega_2}, \quad i=1,2 \quad (2.9)$$

- *Layer 4:* The single node is fixed with output equal to the sum of all the input signals.

$$O_{4,i} = \bar{\omega}_i y_i = \bar{\omega}_i (p_i x_1 + q_i x_2 + r_i) \quad i=1,2 \quad (2.10)$$

where  $(p_i, q_i, r_i)$  is referred to as the consequent parameter set.

- *Layer 5:* Nodes are adaptive with node function given by Layer 1 for a first-order model, and with parameters referred to as defuzzifier of consequent parameters. The node function can be represented as given by the following equation:

$$Output = O_{5,i} = Y = \sum_i \bar{\omega}_i y_i \quad (2.11)$$

## 2.6 Neuro-Fuzzy System Identification

Essentially, system identification is a process of developing or improving a mathematical representation of a physical system using a combination of prior knowledge and empirical data. Conventionally, system identification heavily relies on prior knowledge such that mechanistic model can be established to satisfactorily describe the system. Empirical data is then used to adapt and validate these identified models. In general, there are three basic classes of such model synthesis [47]:

- *White box identification:* The entire mechanistic model is constructed from prior knowledge and physical rules. During the model construction, the observed system data are not used; they are only used for adaptation and validation. However, complete system knowledge is rarely available; consequently, in most cases, some aspects of system behavior cannot be explained by the identified model.
- *Grey box identification:* A partial model is established from prior knowledge and physical rules. Usually, such knowledge is used to satisfy the functional form of the model, and the empirical data is used to infer specific system parameters.
- *Black box identification:* No physical knowledge is available for construction of the model. This kind of identification can be viewed as function approximation or surface fitting. In this case, there exist infinitely many candidate functions that can be chosen to fit the observed data.

Although there exist different approaches for identification of an unknown system, it will be advantageous to employ as much available knowledge as possible. As mentioned above, complete physical knowledge is usually unavailable or difficult to acquire. However, prior knowledge can come in different forms. For example, a technician familiar with a specific system usually is able to describe the system operation correctly, without knowing much about the dynamics inside the system. This kind of knowledge can be very helpful for the modelling process.

Fuzzy systems have the capability of capturing such human experience or knowledge. The information is stored in a set of interpretable IF-THEN rules. If such prior knowledge is correct and complete, this fuzzy rule base approach is analogous to the white box identification. Of course, it is not realistic. Due to vagueness and subjectivity in human thinking and language, it is unlikely that a human expert can accurately describe the system behaviors in all possible situations. It is true especially when the system is complicated. This is the reason why learning-based methods such as neural networks should be employed in combination with fuzzy systems.

While fuzzy systems utilize as much prior knowledge as possible, neural networks can learn from the system observations and extract the hidden information. With reference to the different approaches described above, neuro-fuzzy modelling can be classified as grey box identification. Although neuro-fuzzy systems appear as promising modelling tools, neuro-fuzzy modelling has its limitations. First, it has to be assumed that the human knowledge and/or empirical data are of good quality. Secondly, when the dimension of the system increases, the size of a neuro-fuzzy model (the number of fuzzy rules) grows exponentially. This phenomenon, known as the “curse of dimensionality”, makes the implementation of high-dimensional neuro-fuzzy systems impractical.

As pointed out in [48], system identification involves two steps: structure identification and parameter identification. Structure identification can further be divided into two different types of tasks, (1) input variable selection and (2) determination of functional expression.

*Structure Identification:* Generally speaking, structure identification of an unknown system is to solve two problems: to select a set of input variables and to establish an input-output relation. In phase I of the structure identification, we need to find most relevant input variables for the system. In some identification works, appropriate input variables to the system are assumed to be known. However, in many real-world applications, this is not the case. While linear functions have a unique representation, there is an infinite number of candidates for the nonlinear systems in the absence of prior knowledge. In phase II of structure identification, which is concerned about the input-output relation, there are two parts of an IF-THEN rule: the premise part and the consequent part. So the rules have two structures in principle. It is clear that the number of rules corresponds to the number of subspaces. This partition of input space is the premise structure in a fuzzy model. So phase II of structure identification determines how the input space should be partitioned.

*Parameter Identification:* In conventional system identification, parameters are the coefficients in a system functional model. In a fuzzy model, the parameters also include those in the MFs. In the literature, many existing fuzzy identification schemes perform those two identification tasks separately. However, structure identification and parameter identification cannot be solved separately.

## 2.6.1 Structure Identification

### Variable Selection

Regarding all the aforementioned tasks to be done in identification, the most important task is to identify an optimal set of variables. This is because variable selection is the preliminary step before operating regression or any other identification tasks. One has to include enough variables to explain system behaviors. However, too many variables will artificially increase the complexity of the identified model. This over-complex model will usually over fit the data and, therefore, will have less generalization capability. A variable selection technique typically consists of the following elements:

- a feature evaluation criterion for comparing subsets of variables,

- a search procedure for searching the set of possible variable combination,
- and a stop criterion as a significance threshold for determining the evaluation criterion or the final feature space dimension.

Depending on the task and the model to be used, several evaluation criteria based on either statistics or heuristics, have been proposed for measuring the importance of a variable subset. Some selection methods only consider the data for computing the relevant variables; others take into consideration the model to be used for the medullization task.

### Feature Space Partition

Construction of a fuzzy rule base can be viewed as partitioning the multidimensional input space and extracting IF-THEN rules from the sampled data. Three commonly used partitioning methods are described here by highlighting their basic ideas and characteristics.

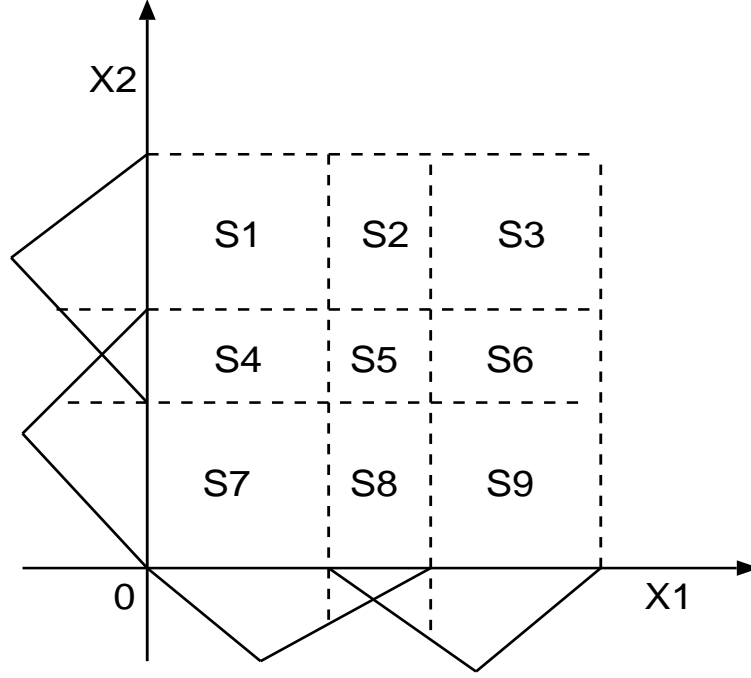
*Fuzzy Grid Partition:* Grid partitioning divides the data space into rectangular subspaces using axis-paralleled partitions based on a pre-defined number of MFs and their types in each dimension, as shown in Fig. 2.3. Based on grid partitioning, fuzzy rules are generated from input-output data.

It determine the degrees of given data pairs in different regions. For example, suppose we are given the following set of desired input-output data pairs:

$$(x_{1,1}, x_{2,1}; y_1), (x_{1,2}, x_{2,2}; y_2), \dots \text{ having values as } (0.6, -0.2; 0.2), (0.4, 0; 0.4), \dots \quad (2.12)$$

where the first two numbers  $(x_1, x_2)$  are inputs and the third one  $(y)$  is the output for each data pair. Fuzzy grid partitioning consists of the following steps:

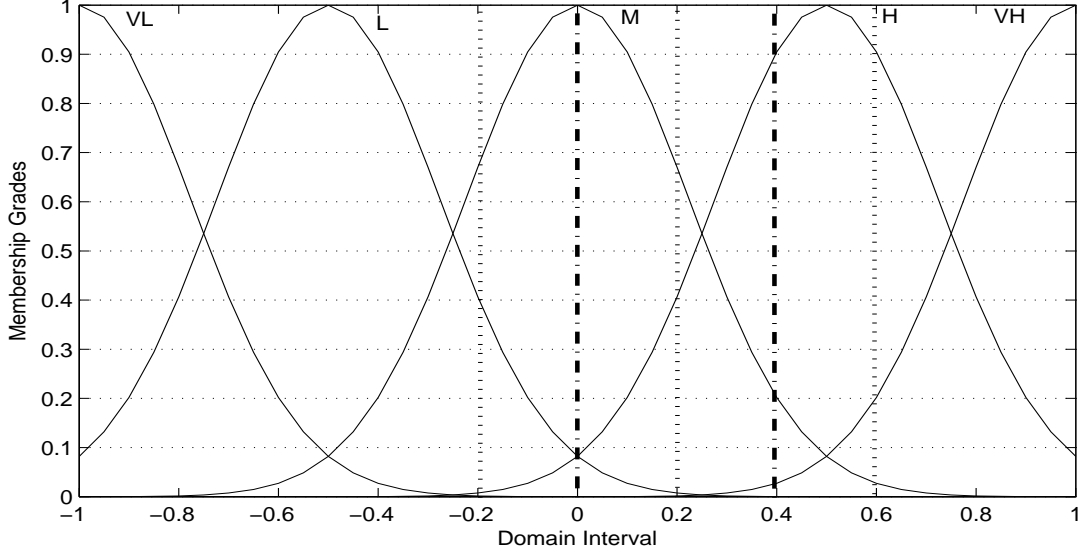
- *Step 1: Divide the input and output space into fuzzy regions.* After the number of MFs associated with each input and output are fixed, the initial values of parameters are set in such a way that the centers of the MFs are equally spaced along the range of each input and output variable. Moreover, these MFs satisfy the condition of  $\epsilon$ -completeness with  $\epsilon=0.5$ , which means that given a value  $x$  of one of the inputs in



**Figure 2.3:** Grid Partitioning.

the operating range, we can always find a linguistic label  $A$  such that  $\mu_A(x) \geq \epsilon$ . In this manner, the FIS can provide smooth transitions and sufficient overlap from one linguistic label to another. One has to choose the intervals for the linguistic values of each input and output linguistic variables in such a way that they do overlap and also cover the entire space of the corresponding input-output linguistic variables. For instance assume that the domain intervals of  $x_1$ ,  $x_2$  and  $y$  are  $[-1, 1]$ . One would divide each domain interval into  $N$  regions and also assign each region a fuzzy MF. Fig. 2.4 shows an example where the domain intervals of  $x_1$ ,  $x_2$  and  $y$  are divided into five regions. In this example, Gaussian MFs are used. Of course, other divisions of the domain regions and other shapes of MFs are possible.

- *Step 2: Generate fuzzy rules from given input-output data pairs.* Determine the degrees of the given  $x_{1,i}$ ,  $x_{2,i}$ , and  $y_i$  in different regions. For example,  $x_{1,1}$  in Fig.2.4 has a degree of 0.9 in H, a degree of 0.2 in VH. Similarly  $x_{2,2}$  has degree 1 in M and smaller degrees



**Figure 2.4:** Initial MFs for  $x_1$ ,  $x_2$  and  $y$ .

than  $x_{2,2}$  in all other regions. Assign  $x_{1,i}, x_{2,i}$  and  $y_i$  to a region with maximum degree: e.g.  $x_{1,1}$  in Fig. 2.4 is assigned to H and  $x_{2,2}$  is assigned to M. Finally, obtain one rule from one pair of desired input-output data, for example,

$$\begin{aligned}
 & (x_{1,1}, x_{2,1}; y_1) \\
 & \implies [x_{1,1} (0.9 \text{ in H}), x_{2,1} (0.7 \text{ in M}), y_1 (0.7 \text{ in M})] \\
 & R_1 \rightarrow \text{IF } x_1 \text{ is H and } x_2 \text{ is M, THEN } y \text{ is M}
 \end{aligned}$$

$$\begin{aligned}
 & (x_{1,2}, x_{2,2}; y_2) \\
 & \implies [x_{1,2} (0.9 \text{ in H}), x_{2,2} (0.7 \text{ in M}), y_2 (0.9 \text{ in H})] \\
 & R_2 \rightarrow \text{IF } x_1 \text{ is H and } x_2 \text{ is M, THEN } y \text{ is H}
 \end{aligned}$$

- *Step 3: Assign a degree of certainty to each rule.* To resolve a possible conflict problem, i.e. rules having the same IF-part but a different THEN-part, and to reduce the number of rules, we assign a degree to each rule generated from data pairs and accept only



the rule from a conflict group that has a maximum degree. In other words, this step is performed to delete the redundant rules, and therefore obtain a concise fuzzy rule base. The effect of these rules is that each of the rules is activated to a certain degree represented by the weight value associated with that rule. The following product strategy is used to assign a degree to each rule. The degree of the rule denoted by

$$R_i \rightarrow \text{IF } x_1 \text{ is A and } x_2 \text{ is B , THEN } y \text{ is } \omega_i \quad (2.13)$$

is defined as

$$\omega_i = \mu_a(x_1)\mu_b(x_2)\mu_c(y) \quad (2.14)$$

For example  $R_1$  has a degree of

$$\omega_1 = \mu_H(x_1)\mu_M(x_2)\mu_M(y) = 0.9 \times 0.7 \times 0.7 = 0.441 \quad (2.15)$$

and  $R_2$  has a degree of

$$\omega_2 = \mu_H(x_1)\mu_M(x_2)\mu_H(y) = 0.9 \times 1 \times 0.9 = 0.81 \quad (2.16)$$

Note, that if two or more generated fuzzy rules have the same preconditions and consequents, then the rule that has maximum degree is used. In this way, assigning the degree to each rule, the fuzzy rule base can be adapted or updated by the relative weighting strategy.

After fuzzy rules have been established, the overall output can be computed by any defuzzification methods. In general, the performance depends heavily on the definition of grids; the finer the grid is, the better the performance that will be. However, it is likely that the fuzzy regions used in this approach will not cover all training data and some regions remain undefined. Adaptive fuzzy grid can be used to optimize the system mapping. First, a uniformly partitioned grid is used as the initial state. As the process goes on, the parameters in the antecedent MF will be adjusted. Consequently, the

fuzzy grid evolves. The gradient descent method optimizes the size and location of the fuzzy regions and the overlapping degree among them. No matter what grid partition method is used, there exists one major drawback in this scheme: the performance suffers from an exponential explosion as the number of inputs or the MFs on each input increase. For example, a fuzzy model with five inputs and five fuzzy sets on each input would possibly result in  $5^5=3125$  IF-THEN rules. It is usually referred to as the “*curse of dimensionality*”.

*Fuzzy k-d Tree Partitioning:* The tree partition method is more suitable for multi-input-single-output (MISO) systems. To illustrate the ideas of tree partition, the procedure to construct a Mamdani type of fuzzy system is described. Suppose there exists a set of input-output data from a system with  $n$  inputs,  $x_i|i = 1, \dots, n$ , and one output. The input space is divided into a set of subspaces or bins, to create a fuzzy rule for each subspace. Then IF-THEN rule for bin  $b$  is in the form of the Mamdani type,

$$R_b \rightarrow \text{IF } x_1 \text{ is } \mu_{1,b}(x_1) \text{ and } x_2 \text{ is } \mu_{2,b}(x_2), \dots \text{and } x_n \text{ is } \mu_{n,b}(x_n) \text{ THEN } y \text{ is } C_b \quad (2.17)$$

where  $\mu_{1,b}(x_i)|i = 1, \dots, n$  are input fuzzy MFs and  $C_b$  is the output for rule  $b$ .

We begin by viewing the entire compact set defined by the range of the inputs as one bin. We divide the initial bin into two smaller bins. Then, we divide these smaller bins recursively until the difference between the two resulting bins can be ignored or the number of training data points within a bin reaches a minimum. As we divide the bins, a fuzzy rule for each bin is created. Thus, a fuzzy rule base is dynamically created to represent the input-output data as the partition proceeds.

In bin  $b$ , the minimum and maximum values of each variable  $x_i$  define an interval,  $[\min_{i,b}, \max_{i,b}]$ . The goal is to divide or cut the interval  $[\min_{i,b}, \max_{i,b}]$  at the place that yields the maximum difference  $|C_{b,1} - C_{b,2}|$  between the two fuzzy outputs  $C_{b,1}$  and  $C_{b,2}$ . There are several strategies for deciding which dimension to cut, where to cut at each step and when to stop [49, 50]. Experiment has shown that 35%, 50%, and 65% of the length

of the interval  $[min_{i,b}, max_{i,b}]$  are good positions for the trial cuts. Thus, the partition procedure becomes the following: we cut each interval at these trial positions, calculate the difference between the two fuzzy rule outputs for each trial cut, and choose the cut that yields the maximum difference. Since there are three trials per dimension, a problem with  $n$  dimensions requires  $3n$  different cuts to generate two fuzzy rules [51]. Note that because the input space is cut in one dimension each time, only two rules will be generated after each cut. Hence, the number of rules will not grow exponentially. After the MFs have been defined, the output of a fuzzy rule corresponding to bin  $b$  is computed as

$$C_b = \frac{\sum_{i=1}^M \mu_b(x_1, \dots, x_n) \times y_i}{\sum_{i=1}^M \mu_b(x_1, \dots, x_n)} \quad (2.18)$$

where  $M$  is the number of training data points, and  $\mu_b(x_1, \dots, x_n)$  is the firing strength of rule  $b$ , calculated by the definition of T-norm. This flexible tree partition algorithm relieves the problem of an exponential increase of the number of rules. However, more MFs will be needed for each input variable, and these usually do not have clear linguistic meanings. Consequently, the resulting fuzzy model is less descriptive.

*Clustering/ Scatter Partitioning:* Clustering/scatter partitioning is the most attractive approach. The clustering method or scatter partition, the fuzzy clustering algorithm in particular, provides a flexible partitioning technique. Clustering can be used to detect possible groups that exist, groups that have similarity in behaviors, and the groups that can be used to establish some hypothesis about the structure presented in the data. Instead of covering the whole space, scatter partitioning tries to find subspaces that characterize the fuzzy region of the input space. Scatter partitioning gives the most efficient partition with a smaller amount of computing time compared with other methods. The drawback of scatter partitioning is how the quality of the fuzzy system depends on the completeness of the data set in representing the whole operation region of the system. Summarizing the fact, the scatter partition/clustering groups the input-output data pairs into clusters and one fuzzy rule represents one cluster. The number of rules in the neuro-fuzzy system is equal to the number of clusters. Systems with different composition and number of clusters can

be formed by varying parameters in the clustering algorithm.

*Data Clustering:* Data clustering algorithms are used to categorize and organize data. These categorized data can then be used for applications such as data compression, model building, etc. The clustering in the fuzzy system is useful for reducing the dimension of fuzzy system rules while still representing the overall system. Clustering partitions a data into several clusters where each data point in a cluster has more similarity than the one among the clusters. In neuro-fuzzy systems, clustering is used to determine the initial locations and the number of IF-THEN rules. There are several clustering techniques that are used for this purpose and the most common ones are: (1) K-means, (2) Fuzzy C-means, (3) Mountain clustering method, and (4) Subtractive Clustering.

- *K-Means Clustering:* The k-means clustering is also known as the hard c-means clustering since a point belongs to only a particular cluster and not others. The opposite of this method is the fuzzy clustering in which the data point can belong to all clusters with different degrees of membership.

This clustering algorithm partitions a collection of  $n$  data points  $x_1, x_2, \dots, x_n$ , into a cluster  $c$ . The cost function that minimizes the distance between the data points that belong to a cluster with cluster center  $v_i$  can be defined by

$$J = \sum_{i=1}^c \sum_{j=1}^n \|x_j - v_i\|^2 \quad (2.19)$$

The clusters are defined by its cluster center and a  $c \times n$  binary membership matrix  $P$ , where the element  $p_{i,j}$  is 1 if the  $j$ th data point  $x_j$  belongs to the  $i$ th cluster and 0 otherwise. The process of determining the cluster center  $v_i$  and membership matrix  $P$  is iterative. The cluster centers are initialized randomly. The membership matrix  $P$  is then calculated as follows:

$$p_{i,j} = \begin{cases} 1 & \text{if } \|x_j - v_i\|^2 \leq \|x_j - v_k\|^2, \text{ for each } k \neq i \\ 0 & \text{otherwise} \end{cases} \quad (2.20)$$

The cost function is computed and the iteration can be stopped if the value is below some tolerance or if the improvement over previous value is below some threshold. The cluster centers are updated using the new membership matrix  $P$  as:

$$v_i = \frac{\sum_{j=1}^n p_{i,j} x_j}{\sum_{j=1}^n p_{i,j}} \quad (2.21)$$

The process is repeated until a satisfactory result is found or a maximum number of iterations has been reached.

- *Fuzzy C-Means (FCM) Algorithm:* Conventional clustering algorithms locate a hard partition of a given data set where each entry of the data belongs to one partition or the other. On the other hand, the fuzzy clustering finds a soft partition of a given data set. Each entry of data can belong to a multiple clusters. The degree of an entry in data to a cluster is given by a degree of membership. A widely used type of the fuzzy clustering algorithm is the fuzzy c-means or ISODATA [52].

Data set  $X$  with  $n$  data points  $x_1, x_2, \dots, x_n$  can be clustered into  $c$  fuzzy sets using the fuzzy c-mean clustering method. The criterion in most instances is to optimize an objective function that acts as a performance index of clustering. The end result of the fuzzy clustering can be expressed as a partition matrix  $P$ :

$$P = p_{i,j} \text{ with } i = 1, \dots, c \text{ and } j = 1, \dots, n \quad (2.22)$$

where  $p_{i,j}$  is a numerical value between 0 and 1 and expresses the degree to which the data point  $x_j$  belongs to the  $i$ th cluster. The objective function of the FCM algorithm takes the form of

$$J(p_{i,j}, v_k) = \sum_{i=1}^c \sum_{j=1}^n \|x_j - v_i\|^2, \quad m > 1 \quad (2.23)$$

where  $m$  is the fuzziness factor, which influences the degree of fuzziness of the cluster partition. If  $m$  is a large number, a point with less membership in the cluster will have

less influence on the calculation of the new cluster center.  $v_i$  is a cluster center for the  $i$ th cluster. To solve this minimization problem, the objective function is differentiated w.r.t.  $v_i$  (for fixed  $p_{i,j}, i = 1, 2, \dots, c, j = 1, 2, \dots, n$ ) and with respect to  $p_{i,j}$ , (for fixed  $v_i, i = 1, 2, \dots, c$ ).

$$v_i = \frac{\sum_{j=1}^n (p_{i,j})^m x_j}{\sum_{j=1}^n (p_{i,j})^m}, \quad i = 1, 2, \dots, c \quad (2.24)$$

$$p_{i,j} = \sum_{k=1}^c \left[ \left( \frac{\|x_j - v_i\|^2}{\|x_j - v_k\|^2} \right)^{\frac{1}{m-1}} \right]^{-1}, \quad i = 1, 2, \dots, c, j = 1, 2, \dots, n \quad (2.25)$$

After determining the number of clusters  $c$  ( $2 \leq c \leq n$ ) and fuzziness factor, the initial partition matrix  $P$  is chosen randomly. Cluster centers and the partition matrix can be calculated iteratively from the above equations. If the difference of the previously calculated center and/or partition matrix and the current value is less than the predetermined threshold, the process can be stopped.

- *The Mountain Clustering Algorithm:* The mountain clustering method is a grid-based method for identifying the approximate locations of the cluster centers [53]. Unlike, fuzzy c-means, this method does not require a predetermination of the number of clusters. Grid points on the data space provide the potential cluster centers. A finer grid does not only increase the number of potential cluster centers but it also increases the required computation effort. The grid is generally evenly spaced, but it is not required. Uneven spaced grids that reflect the prior knowledge of the data space can be formed.

Grid point selection for a cluster center is based on the mountain function. The height of the mountain function at a grid point  $g$  is equal to

$$H(g) = \sum_{i=1}^n \exp \left( -\frac{\|g - x_i\|}{2\gamma^2} \right) \quad (2.26)$$

where  $x_i$  is the  $i$ th data point and  $\gamma$  is an application specific constant. The closer the data point  $x_i$  to the grid point, the more it contributes to the height of the mountain function. The value of the mountain function reflects the density of data points in the vicinity of each grid point. The higher the mountain function value at a grid point, the larger it's potential for being a cluster center. The grid node with the highest score of the mountain function is selected and becomes the first cluster center  $v_i$ . The next cluster center could not be selected yet since the first cluster center is usually surrounded by a number of grid points which also have high density values. The effect of the first center must be eliminated by sequentially destructing the mountain function. In order to do so a revised mountain function is formed:

$$H_{new}(g) = H(g) - H(v_i) \exp\left(-\frac{\|g - v_1\|^2}{2\beta^2}\right) \quad (2.27)$$

After the subtraction, the new mountain function value at  $v_1$  is zero and its effect on surrounding points is eliminated. The second cluster center then can be selected from the grid point with the highest value of the new mountain function. This process is repeated until the new mountain function value is less than a stopping constant value.

- *Subtractive Clustering*: the mountain clustering method is simple and very effective in finding cluster centers that can be the base of fuzzy system MFs. However, the number of calculations required grows exponentially with the dimensions of the data set. For data set of 3 variables and 10 grid points for each variable, 1000 points must be evaluated. Adding another variable to the data set multiplies the grid points by 10 or 10,000 grids.

A variation of the mountain method called subtractive clustering solved this problem [54]. Instead of using grid points, data points are used as candidates of the cluster centers. By doing so the computation at effort needed for calculation is proportional to the number of data points and is independent of the dimension of the problem/variables. This rough calculation of the cluster centers is particularly suit-

able if the clustering method is used to find the initial structure of a fuzzy system that will be optimized later by the neural network learning algorithm.

For a data set of  $n$  data points, a density measure at data point  $x_i$  is defined as

$$D_i = \sum_{j=1}^n \exp \left( -\frac{\|x_i - x_j\|^2}{\left(\frac{r_a}{2}\right)^2} \right) \quad (2.28)$$

where  $r_a$  is a positive constant. A data point will have a high density value if it has many neighboring data points.

As in the mountain method, the data point with the highest density measure is selected as the first cluster center  $v_1$ . The next step is to eliminate the influence of the first cluster center to the surrounding data points which also have high density values. The density measure of each data point is revised as

$$D_{i(new)} = D_i - D_{v_1} \exp \left( -\frac{\|x_i - x_{v_1}\|^2}{\left(\frac{r_b}{2}\right)^2} \right) \quad (2.29)$$

where  $r_b$  is a positive constant. The density measure of data points in the neighborhood of the cluster center  $v_1$  is reduced and the one at the first cluster center is zero. The effect of the first cluster center on the surrounding points is eliminated. The constant  $r_b$  defines a neighborhood that has significant reduction in density measures after the revision. The constant  $r_b$  is usually larger than  $r_a$  to prevent closely spaced cluster centers. Generally  $r_b$  is chosen to be equal to  $1.5 \times r_a$ .

The point with the highest density measure is selected again as the next cluster center. This process is iterated until the highest density measure is lower than a predetermined stopping constant value or sufficient number of cluster centers has been determined. The result can be used for developing the Takagi-Sugeno fuzzy model. Cluster centers  $v_i$  are fuzzy system rules. The degree of fulfillment of the fuzzy  $i$ th rule is defined by



$$\mu_i = \exp \left( -\frac{\|x - v_i\|^2}{\left(\frac{r_a}{2}\right)^2} \right) \quad (2.30)$$

After completing this procedure, a more accurate system can be constructed using optimization scheme like gradient descent algorithm.

## 2.6.2 Parameter Identification

Most existing neuro-fuzzy systems are rule based systems which consist of the Mamdani model or the TSK model. This type of neuro-fuzzy model has the following form,

$$R_i \Rightarrow IF \ x_1 \text{ is } A_{1,i} \text{ AND } \cdots \text{ AND } x_p \text{ is } A_{p,i} \text{ THEN } y = f_i(x_1, \dots, x_p), i = 1, \dots, M \quad (2.31)$$

and  $f_i(x_1, \dots, x_p)$  is given by

*Mamdani Type:*

$f_i = c_i$ , where  $c_i$  is a constant consequent.

*TSK Type:*

$$f_i = \sum_{j=1}^p a_{i,j} x_j \quad (2.32)$$

The total output of the model is computed by the COA defuzzification method,

$$y = \frac{\sum_{i=1}^M \omega_i f_i}{\sum_{i=1}^M \omega_i} = \sum_{i=1}^M \bar{\omega}_i f_i \quad (2.33)$$

where  $\omega_i$  denotes the firing strength of the  $i$ th rule and  $\bar{\omega}_i$  is the normalized firing strength of the  $i$ th rule defined by

$$\bar{\omega}_i = \frac{\omega_i}{\sum_{i=1}^M \omega_i} \quad (2.34)$$

In the least squares method, the output of a model  $y$  is given by the parameterized expression

$$y = \theta_1 f_1(x) + \theta_2 f_2(x) + \dots + \theta_n f_n(x) \quad (2.35)$$

where  $x = [x_1, \dots, x_n]$  is input vector for the model,  $f_1, \dots, f_n$  are known functions of  $x$ , and  $\theta_1, \dots, \theta_n$  are unknown parameters to be optimized. To identify these unknown parameters  $\theta_i$ , usually a training data set of data pairs  $(x_i, y_i), i = 1, \dots, m$  is taken; substituting each data pair in Eq. (2.35) a set of linear equations is obtained, which can be written as

$$A\theta = y \quad (2.36)$$

in matrix form. Where  $A$  is a  $m \times n$  matrix.

$$A = \begin{bmatrix} f_1(u_1) & \cdots & f_n(u_1) \\ \vdots & \vdots & \vdots \\ f_1(u_m) & \cdots & f_n(u_m) \end{bmatrix} \quad (2.37)$$

where  $\theta$  is an  $n \times 1$  unknown parameter vector

$$\theta = \begin{bmatrix} \theta_1 \\ \vdots \\ \theta_n \end{bmatrix} \quad (2.38)$$

and  $y$  is an  $m \times 1$  output vector

$$y = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} \quad (2.39)$$

Since generally  $m > n$ , instead of exact solution of (2.35) an error vector  $e$  is introduced to account for the modelling error as

$$A\theta + e = y \quad (2.40)$$

and searched for  $\theta = \hat{\theta}$  which minimizes the sum of squared errors.

$$E(\theta) = \sum_{i=1}^m (y_i - a_i^T \theta)^2 = e^T e \quad (2.41)$$

where  $E(\theta)$  is called the overall error objective function. The squared error in (2.41) is minimized when  $\theta = \hat{\theta}$ , called least squares estimator (LSE) that satisfies the normal equation

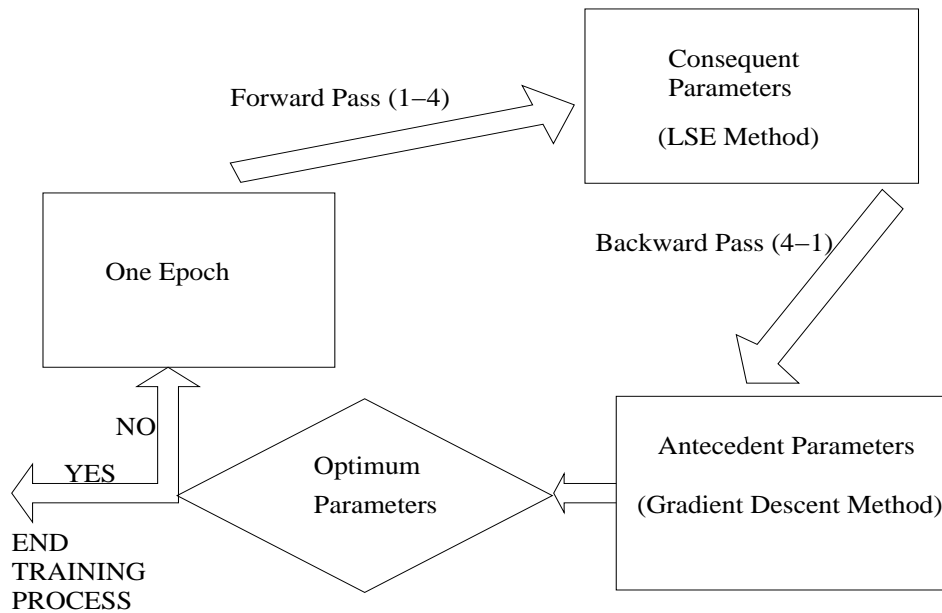
$$A^T A \hat{\theta} = A^T y \quad (2.42)$$

If  $A^T A$  is a non singular,  $\hat{\theta}$  is unique and is given by

$$\hat{\theta} = (A^T A)^{-1} A^T y \quad (2.43)$$

There are two kinds of parameters in a neuro-fuzzy model: the parameters in the consequent parts and the parameters in the antecedent parts. According to Eq. (2.40), the consequent parameters are simply  $\theta$ , and the antecedent parameters are generally those used to define the MFs. Though one can apply the gradient methods such as back propagation (BP) to identify both the antecedent and the consequent parameters, the gradient descent methods are generally slow and likely to be trapped in local minima. One can identify the consequent parameters by the well known least squares (LS) methods. After the consequent parameters have been identified, the BP algorithm can be used to update the antecedent parameters. This observation leads to a hybrid learning algorithm which combines the neural learning rules and the least squares estimators for faster parameter identification. In the hybrid learning scheme, each epoch is composed of a forward phase and a backward phase, as shown in Fig. 2.5. In the forward phase, one present an input vector and compute the corresponding row vectors in matrices  $A$  and  $Y$ . This process is repeated until all the input vectors have been presented. In the backward phase, first the least-squares formula in Eq. (2.43) is used to compute  $\hat{\theta}$ . After  $\hat{\theta}$  is identified, one calculates the error signal for each training data entry. Then the error signals can be back propagated from the output end toward the input end, so as to update the premise parameters. This is a batch learning scheme since the parameters are updated after all the training data have been presented. For the given premise parameters at any training epoch, the  $\theta$  found by LS estimation is guaranteed to be globally optimal. Thus, this hybrid learning algorithm not only can reduce the dimension of each of the search space in the BP algorithm, but also can substantially

reduce the time needed for convergence.



**Figure 2.5:** Hybrid Learning Algorithm: One Epoch.

## 2.7 The Feedback Hybrid Neuro-Fuzzy System

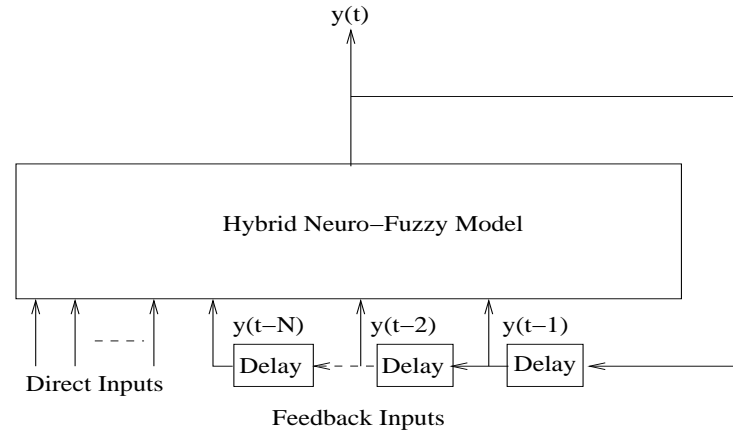
Due to the nature of its feed forward network the hybrid neuro-fuzzy model doesn't allow for effectively capturing the time varying properties of the dynamic processes. Unlike recurrent networks it does not have any dynamic features. In the past few years, many studies related to recurrent neural networks (RNN) have been presented to solve this problem by using attractor dynamics and information storage ability [55–57]. However, RNNs training is found to be more difficult as compared to the feed forward networks because RNNs have complex network structures and therefore heavier computational efforts are also required making it unsuitable for the targeted applications [58]. Furthermore, with feed forward network structures the performance of a fuzzy neural network has been demonstrated to be better than a neural network [59]. Based on these observations, researchers have attempted the construction of a recurrent fuzzy neural network composed of internal feedbacks and

time-delay synapses for control and identifications systems [60–65].

For the model discussed in Section 2.5.1, the dynamic property is achieved by feeding the output back through a feedback loop as shown in Fig. 2.6. Introducing this feedback connection enables the system to remember previous states and use both the previous and current states to calculate the new output values. From a practical point of view, time delay embedded with the feedback is time variant which can lead to a different dynamical behavior and such behavior is vital in improving the ability to provide better and more accurate prediction capability. With the output feedback system, the output at time  $t$  is given by

$$y(t) = f[y(t-1), \dots, y(t-N), X(t-1), \dots, X(t-M)] \quad (2.44)$$

By doing so, one will be able to train the canonical hybrid neuro-fuzzy system as a recurrent of the order of  $N$  by providing the time-delay element at the output.



**Figure 2.6:** The Feedback System for Neuro-Fuzzy Model.

## 2.8 Concluding Remarks

Conventional neuro-fuzzy systems cannot effectively cope with dynamic processes, such as the heating systems of the buildings, due to the feed forward network structure. To overcome

this problem, the existing hybrid system is incorporated with a feedback loop so that it can model the dynamical behavior of the process. In the following Chapter, we develop a RenFIS based inferential model.

# Chapter 3

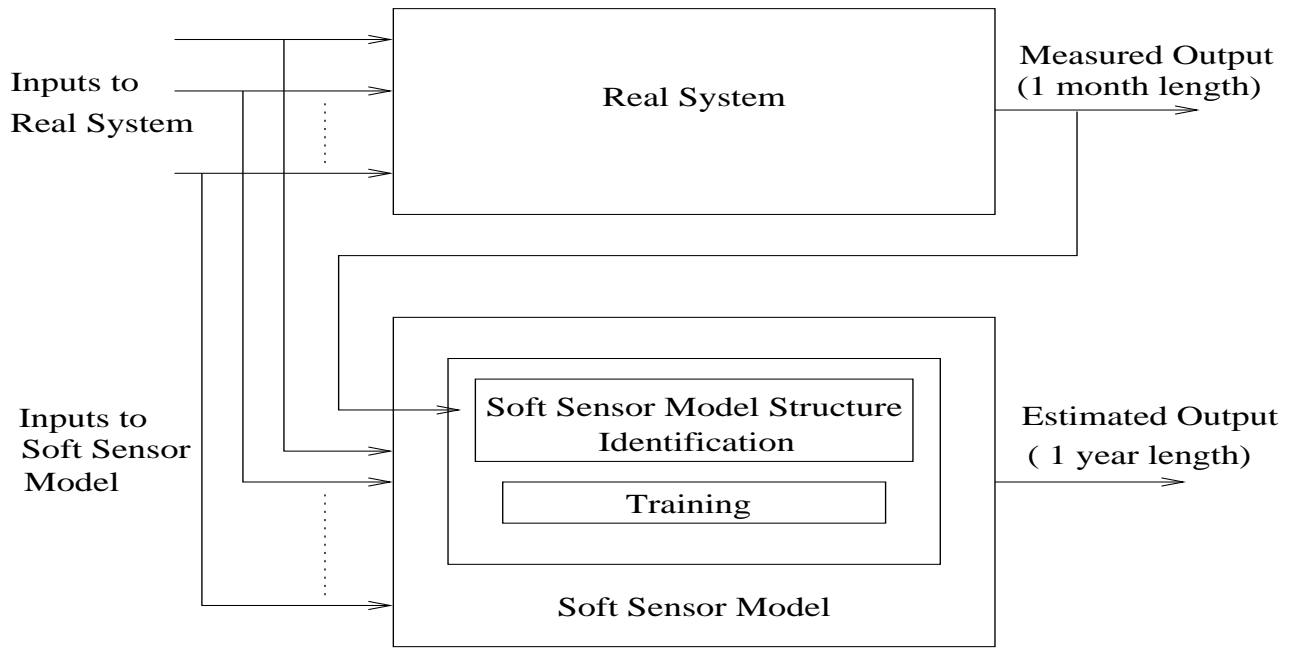
## Development of the RenFIS based Inferential Model

### 3.1 Introduction

An average air temperature based inferential control strategy was proposed in Chapter 1. The strategy controls the furnace operation according to the average air temperature representative of the temperature in all the zones of a multiple-zone building. A method for estimating the average air temperature in the building is needed in order to make use of the proposed scheme in practice.

Soft sensing is an attractive technique for modelling the dynamic behavior of the space heating systems, providing practical methods for estimating the value of the critical control variables that are otherwise difficult, if not impossible, to measure using conventional physical sensors. A more representative measurement of the average air temperature in the buildings, which is essential for optimal control of the relevant equipment, can be obtained using an inferential sensor [9,66]. The inferential model trained with the available short-term measurements can be used for long-term prediction/estimation of the average air temperature, Fig. 3.1 shows that with one month long measurements the inferential model can estimate the variable values for a period up to and beyond one year.

The soft sensing technology was originally developed to improve the control of chemical and biological processes [67–71]. Over the last three decades enormous efforts have been



**Figure 3.1:** Real System vs. Soft Sensor Model.

made to develop soft sensors and soft sensor based control schemes [9, 11, 72]. The model relating the un-measurable variables to the measurable variables, the inferential model, is the heart of a soft sensor.

The statistics show that the number of applications using nonlinear mathematical models is growing while the trend of using linear mathematical models is diminishing. However, it is difficult to achieve accurate nonlinear mathematical models for complex nonlinear systems. Researchers have started augmenting physical models with artificial intelligence techniques for system identification and designs [73, 74]. Also, the physical-model based inferential sensor intends to estimate the time-dependent value of the average air temperature, which is difficult to commission. The physical-model based inferential sensor fails to function properly if the quality of commissioning data is not good or if the optimization process used for commissioning terminated at local extremes. This thesis aims to make contributions to solve the above said problem.

If the system structure is not completely known, the temperature estimation should be based on data or heuristic information. The inherent characteristics of fuzzy logic are suitable



for development of the inferential models for complex systems. The nonlinear mapping characteristic of a fuzzy model, with fast and robust implementation, the capacity to embed prior knowledge and the ability for generalization can be beneficial for the inferential model for temperature estimation. With these advantages, a fuzzy model is a natural tool to deal with nonlinear uncertain conditions in buildings [75, 76].

The fuzzy logic based approach has been used to deal with the thermal and visual comfort problems in the passive buildings without demanding a precise mathematical model. Performance comparison of knowledge-based control systems with classical control systems showed that fuzzy control systems are able to achieve the goals of energy minimization through optimal resource management in the design of passive buildings [77–81]. Fuzzy sets have also been intensively applied in the modelling and control of building thermal systems. Fuzzy systems enhanced with rule indices, allow the fuzzy model rule based structure and parameters to be estimated from the training data without establishing a complete rule list which ultimately make it more flexible [82]. Fuzzy systems add more flexibility, user friendliness and ability to analyze more precisely & deeply the whole set of possible solutions to the optimum design and control of thermal systems [83, 84].

Even though a fuzzy system is easy to comprehend because it uses linguistic terms and structure of IF-THEN rules, it does not have a learning algorithm. Trial and error or expert knowledge is used in tuning the fuzzy system parameters and it can take a long time to finally find an acceptable system [85]. At the same time, neural networks have many learning algorithms but it is extremely difficult to use prior knowledge about the system. It is also almost impossible to explain the behavior of the neural system in a particular situation.

Independently and in parallel to fuzzy systems, neural networks have many applications to buildings including but not limited to modelling, design, control, prediction, operative temperature estimation, energy and optimization studies [86–93]. Yalcintas used ANN for energy auditing and energy saving predictions due to building retrofits and found ANNs have faster learning capabilities, adaptability to seasonal climate variations and changes in the building energy use as compared to other statistical and simulation tools [94]. ANNs

also have applications in solar buildings in predicting the energy demand resulting in energy savings while maintaining thermal comfort [13].

The combination of fuzzy logic and neural networks has produced a functional system capable of learning, simulating high-level thinking and reasoning. It is an improved tool for determining the behavior of imprecisely-defined complex dynamical systems [95]. The neuro-fuzzy systems can combine the benefits of these two powerful paradigms into a single capsule. The purpose of a neuro-fuzzy system is to apply neural learning techniques to identify and tune the parameters and/or structure of neuro-fuzzy systems. The features exhibited in these systems, such as fast and accurate learning, good generalization capabilities, excellent explanation facilities in the form of semantically meaningful fuzzy rules, and the ability to accommodate both data and existing expert knowledge, make them suitable for a wide range of engineering and scientific applications [35]. A particular neuro-fuzzy system named ANFIS was proposed by Jang [20]. Since Jang proposed ANFIS, its applications are numerous in various fields, including engineering, management, health, biology and even social sciences [96–113]

Due to the nature of its feed forward network the hybrid neuro-fuzzy model doesn't allow for effectively capturing the time varying properties of dynamic processes [114, 115]. This research has attempted the development of a RenFIS structure by feeding the output back through a feedback loop. Introducing this feedback connection enables the system to remember previous states and uses both the previous and current state to calculate the new output values.

## **3.2 Neuro-Fuzzy Modelling for Average Air Temperature Estimation**

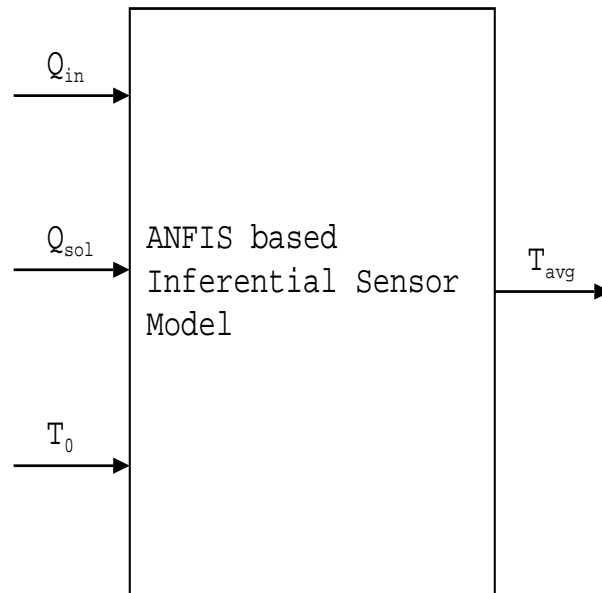
Neuro-fuzzy modelling is used for developing the average air temperature estimator. For hybrid neuro-fuzzy modelling the major steps are: (1) determine input and output variables, (2) experimental data monitoring (3) fuzzy partitioning of the I/O space, (4) shapes and initial parameters for the membership functions, (5) short-term training data selection, (6)

parameters and algorithm for the training process, and (7) validation of the model for long term testing data.

The ANFIS and RenFIS are applied for the estimation of average air temperature in the buildings for two types of heating systems: (1) Hot-Water heating system and (2) Forced Warm-Air heating system.

### 3.2.1 Input and Output Variable Selection

When generating a FIS system structure, first important step is to select the relevant input and output variables. The output variable is hard to measure variable, and the inputs are easily measurable variables. As shown in Fig. 3.2 , there are three inputs and one output for the model:



**Figure 3.2:** Inferential Model Showing Input and Output Variables.

- *Input variables:* energy consumed by the heating plant ( $Q_{in}$ ), solar radiation ( $Q_{sol}$ ), and exterior temperature ( $T_0$ ).
- *Output variable:* the average air temperature in the building ( $T_{avg}$ ).

The three variables,  $Q_{in}$ ,  $Q_{sol}$ ,  $T_0$ , are used as input variables and  $T_{avg}$  is the output variable for ANFIS. For RenFIS in addition to  $Q_{in}$ ,  $Q_{sol}$ ,  $T_0$ , the estimated temperature,  $\hat{T}_{avg}$  is also used as an extra input by feeding back the output.

### 3.2.2 Experimental Data

I/O data pairs are required for short-term training and long-term testing of the developed models. The experimental data are obtained for two types of heating systems, monitored under different research projects. Hot-water heating system data was collected in a laboratory heating system and the length of the experiment was 8 months. Forced warm-air heating system data was collected in a residential building and the length of the experiment was 16 months.

#### *Hot-Water Heating System*

Experimental data obtained from a laboratory heating system are used for training and testing of the developed models. The laboratory heating system was monitored in an EU CRAFT project [116]. The laboratory located in Milan, Italy is a three story building with one zone at each floor. Multiple sensors were used to monitor the air temperature in each zone and their algebraic average was treated as the representative measurement of the room temperature in that zone. Because each zone has the same floor area, the building air temperature is represented by the algebraic average of the air in all three zones. The data are collected for the months of January, February, March, April, November and December of the year 2000 and January and February of the year 2001.

A gas meter was used to monitor the energy consumption  $e$ , where

$$e = [e(1) \ e(2) \ \cdots \ e(N_e)] \quad (3.1)$$

and  $e(i)$  [ $i = 1, \dots, N_e$ ] is the total gas consumption recorded at the  $i$ th time step.  $N_e$  is the total number of samples.

Therefore, the input  $Q_{in}$  is given by

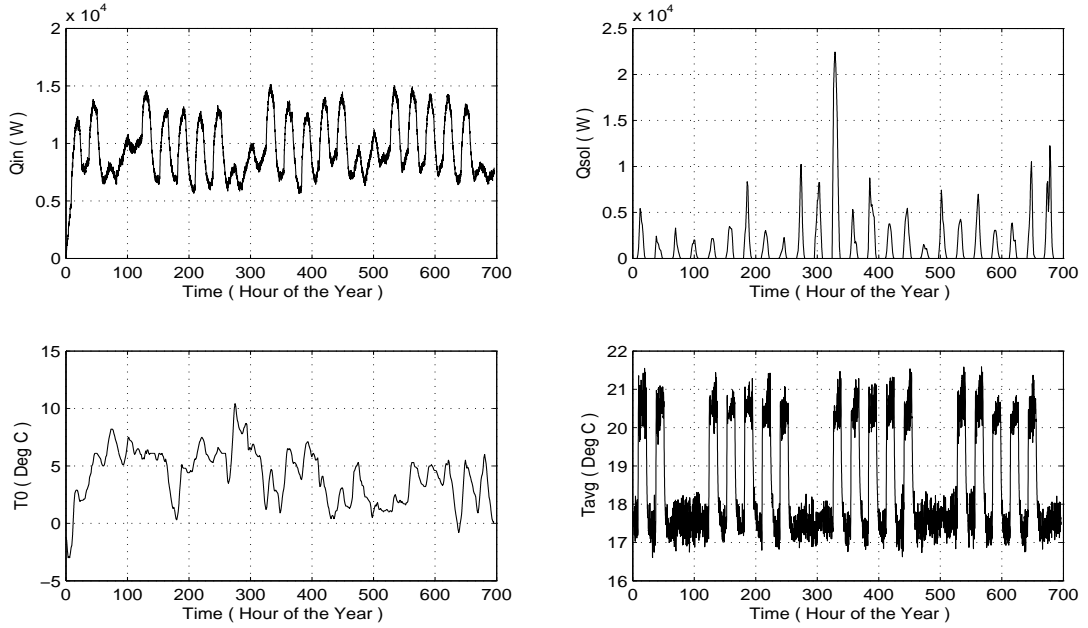
$$Q_{in}(j) = (e(j+1) - e(j)) / \Delta\tau \quad (3.2)$$

where  $\Delta\tau$  is the sampling interval. First order filter is used to convert  $Q_{in}(j)$  obtained from equation (3.2) into a continuous signal.

$$Q_{in}(j) = \alpha Q_{in}(j-1) + (1-\alpha)Q_{in}(j) \quad (3.3)$$

where  $\alpha$  is determined by the the heating equipment parameters.

The other two inputs,  $T_0$  and  $Q_{sol}$ , were monitored regularly by a metrological station next to the laboratory. Fig. 3.3 shows the experimental data for the four variables for the month of February 2000.



**Figure 3.3:** Experimental Data for Hot-Water Heating System (February 2000).

#### *Forced Warm-Air Heating System*

Experimental data are obtained for a residential building located in Markham, Ontario, Canada. The building is a single detached house with three levels, including basement, ground floor and second floor. Each level is divided into zones. Multiple sensors were used

to monitor the air temperature in each zone and their algebraic average was treated as the representative measurement of the room temperature in the zone. It is assumed that each zone has the same floor area; the building air temperature,  $T_{avg}$ , is represented by the algebraic average of the air temperature in all the zones in all the three levels. The sampling interval is 5 minutes.

The time for ON and OFF states of the furnace is recorded. This data is a discrete signal. First order filter is used to convert this discrete signal into a continuous signal. This continuous signal is multiplied by the furnace capacity to compute the actual energy consumed by the furnace system,  $Q_{in}$ .

The exterior temperature is measured with the two sensors.  $T_0$  is represented by the algebraic average of the two exterior temperature measurements. The third input,  $Q_{sol}$ , was monitored by metrological station, University of Toronto, Mississauga, Ontario, Canada. This weather station is operated by the department of geography. CNR1 net radiometer is used for the measurement of net radiation at the earth's surface [117]. The experiment is conducted from December 2007 to April 2009. Fig. 3.4 shows the data collected for the four variables for the month of Jan 2008.

### 3.2.3 Space Partitioning

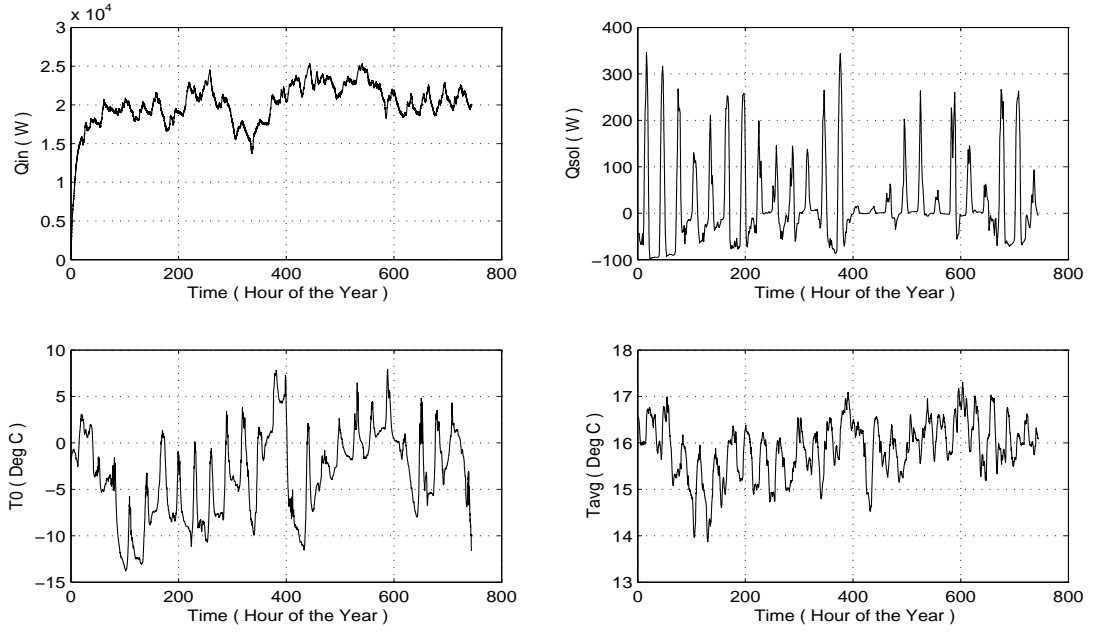
Two methods have been used for space partitioning: (1) Grid Partitioning for ANFIS-GRID structure as shown in Fig. 3.5, and (2) Subtractive Clustering for ANFIS-SUB structure as shown in Fig. 3.6.

### 3.2.4 Number and Type of MFs

For the given set of input-output data pairs, the number and type of MFs are selected for each individual input and output variable. Minimum testing error (MTE) is considered as a criterion for the selection of these variables.

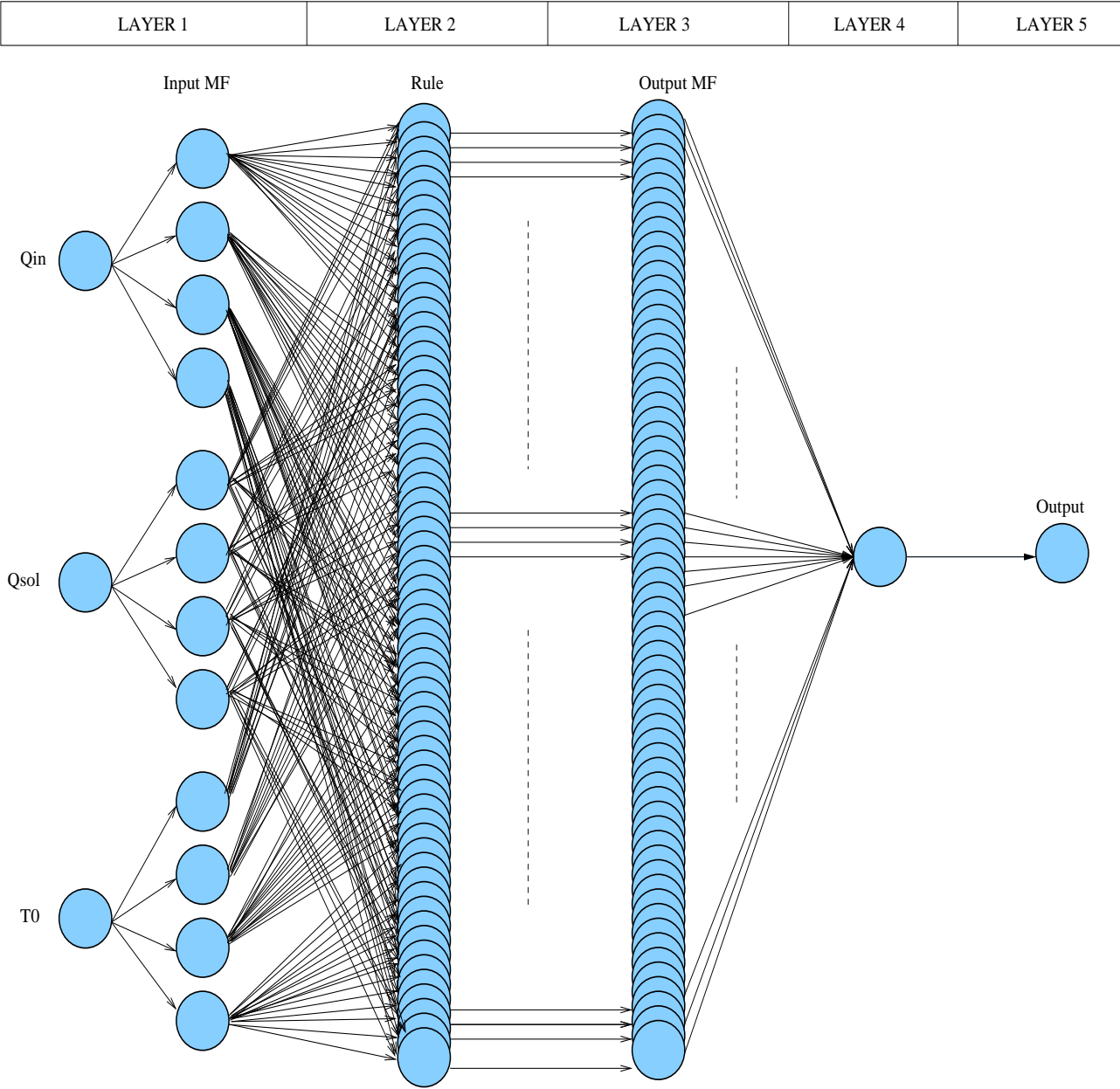
#### *Hot-Water Heating System*

For specific training and testing data sets, Fig. 3.7 and Fig. 3.8 present the impact



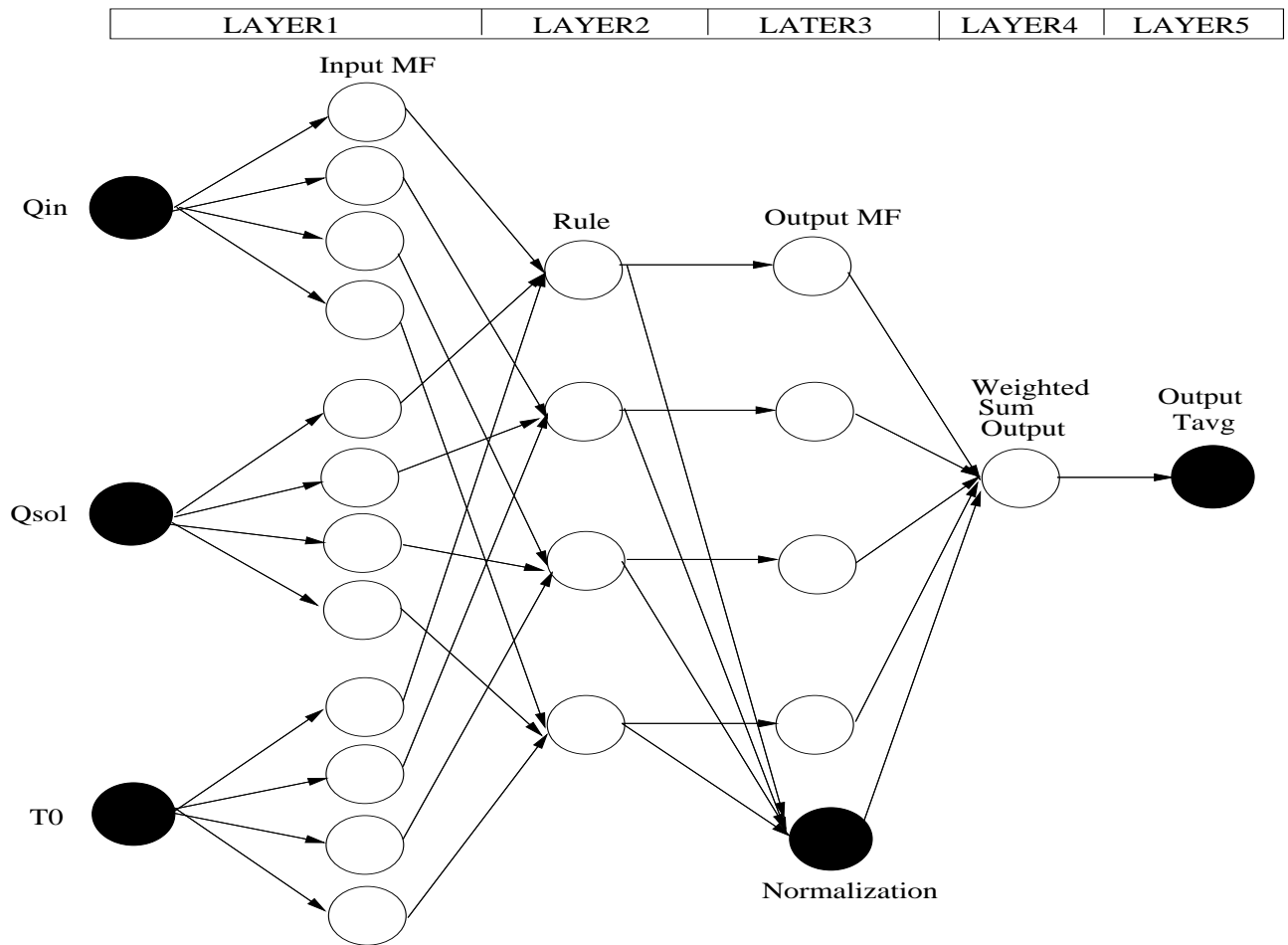
**Figure 3.4:** Experimental Data for Forced Warm-Air Heating System (January 2008).

of number and type of MFs on the training/testing errors. Fig. 3.7 shows that the different combinations of number of MFs affect the training and testing errors significantly but increasing the number of MFs does not necessarily improve the performance of FIS. Four training and testing data sets are selected with different combinations of number of MFs for each input variable. The number of MF combination is: [4 5 4] for trn1 and test1, [4 4 5] for trn2 and test2, [5 4 4] for trn3 and test3 and [4 4 4] for trn4 and test4. For trn4 and test4 data set each input variable has four MFs, [4 4 4], and error curves are with minimum training and testing errors. Fig. 3.8 concludes that the selection of the type of MF depends upon the shape of the training and testing data variations. The different MF types are: [tri tri gauss] for trn1 and test1, [gauss2 gbell gauss] for trn2 and test2, [gauss gauss gauss] for trn3 and test3. For data set three, trn3 and test3, the training and testing error curves give optimal values as MTE occurs at 54 epochs and it is the minimum out of the three testing error curves. The best combination of number of MFs is [4 4 4], i.e. each individual input variable have four MFs. MF type selected is Gaussian MF for all the variables [118].

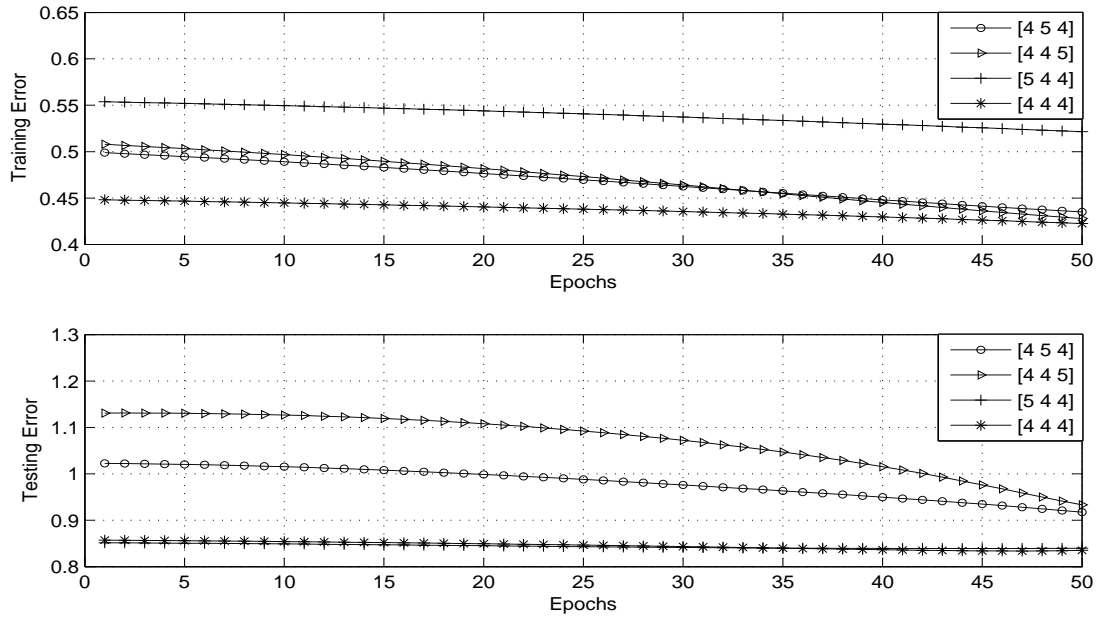


**Figure 3.5:** ANFIS-GRID Model Structure.

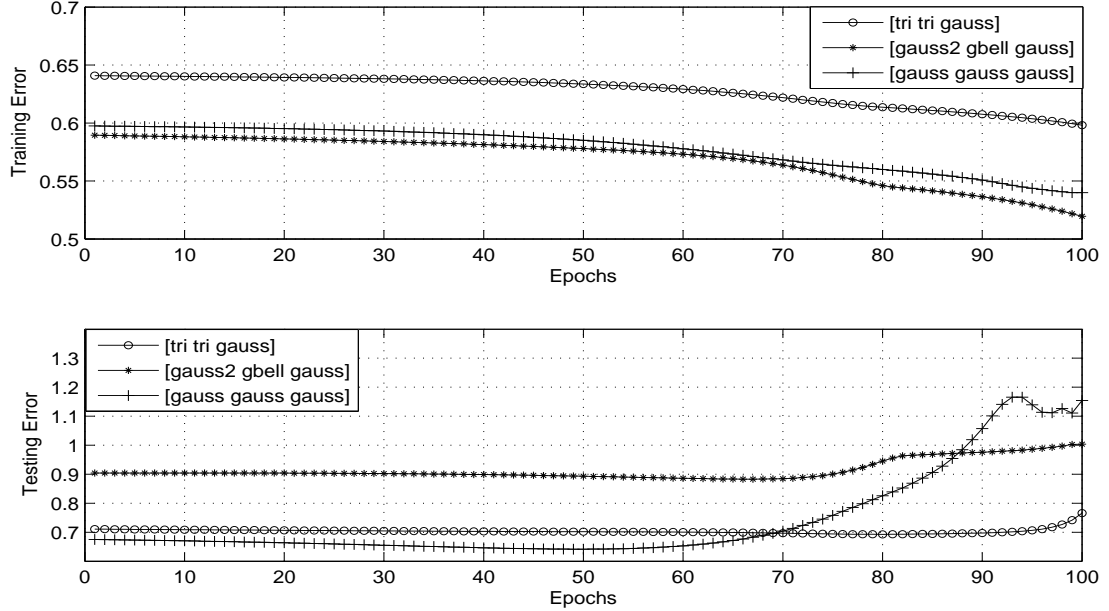




**Figure 3.6:** ANFIS-SUB Model Structure.



**Figure 3.7:** Training and Testing Errors Obtained by the Neuro-fuzzy Model using Different Combinations of Number of MFs for Hot-Water Heating System.



**Figure 3.8:** Training and Testing Errors Obtained by the Neuro-fuzzy Model using Different Types of MFs for Hot-Water Heating System.

### *Forced Warm-Air Heating System*

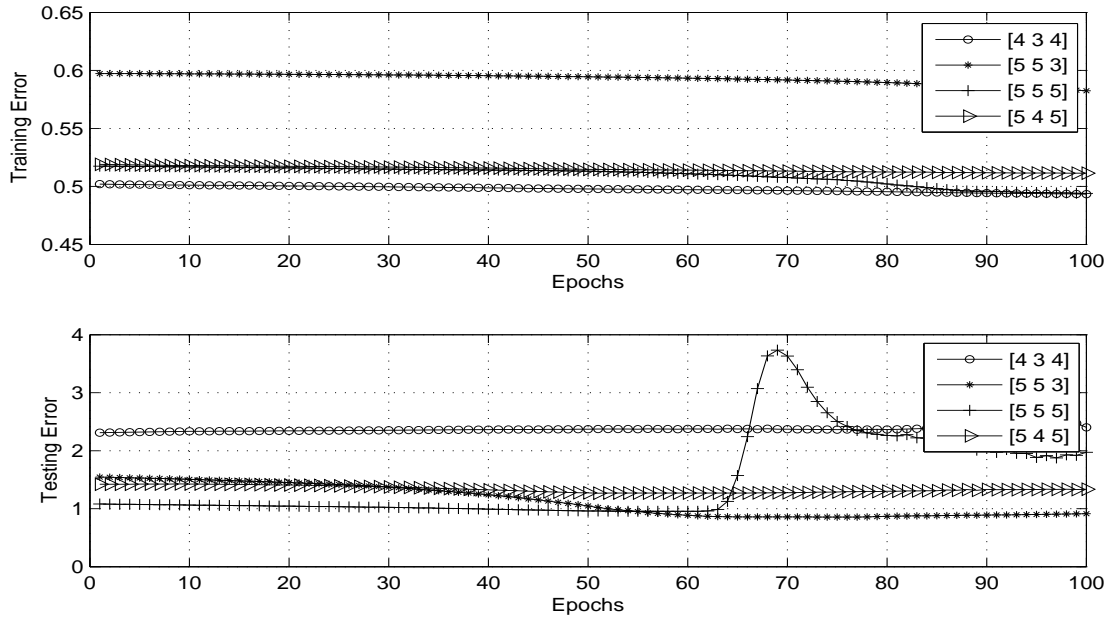
Fig. 3.9 and Fig. 3.10 present the impact of number and type of MFs on the training/testing errors. Fig. 3.9 shows that the different combinations of the number of MFs affect the training and testing errors significantly. The number of MF combination is: [4 3 4] for trn1 and test1, [5 5 3] for trn2 and test2, [5 5 5] for trn3 and test3 and [5 4 5] for trn4 and test4. For trn3 and test3 data set each input variable has five MFs, [5 5 5], and error curve is with minimum testing error. Fig. 3.10 concludes that the selection of the type of MF depends upon the shape of the training and testing data variations. The MF type combination is: [trimf trimf gaussmf] for trn1 and test1, [gauss2mf gbellmf gaussmf] for trn2 and test2, [gbellmf gaussmf gbellmf] for trn3 and test3, and [gaussmf gaussmf gaussmf] for trn4 and test4. For data set four, trn4 and test4, the testing error curve give optimal value as MTE occurs at 60 epochs and it is the minimum out of the four testing error curves. In this case, the input-output training data points in the training data set are large enough as compared to the testing data, and the best combination of number of MFs is [5 5 5], i.e. each individual input variable have five MFs. MF type selected is Gaussian MF.

### **3.2.5 Learning of the Model Parameters**

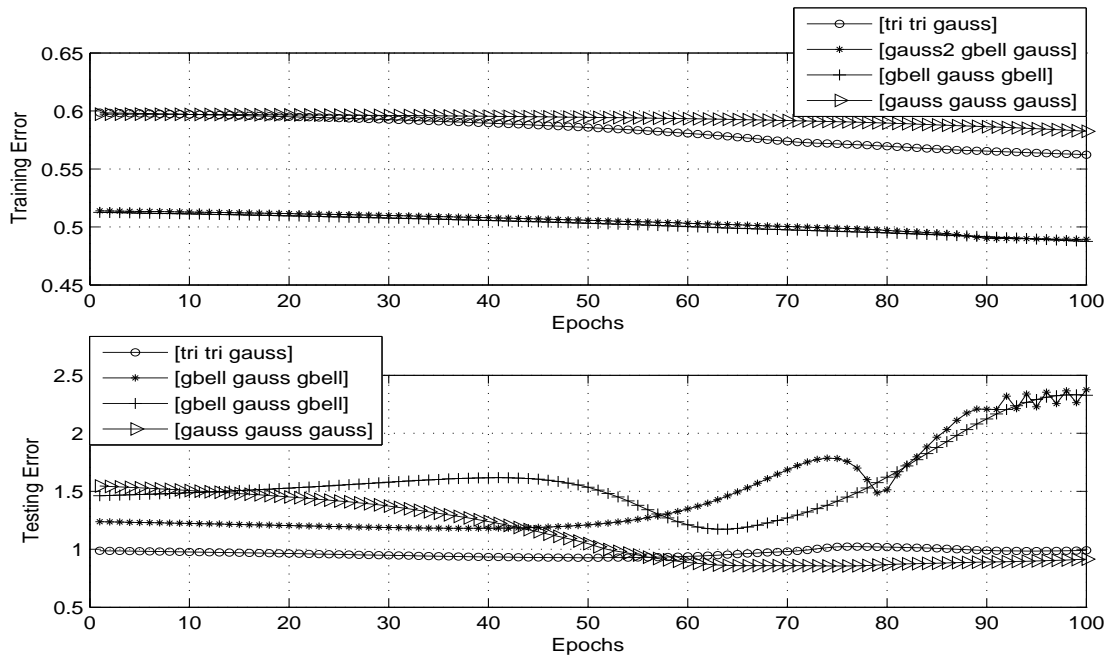
The model is trained using a hybrid learning algorithm through which the suitable values for the parameters are determined in this phase.

#### *Hot-Water Heating System*

Figs. 3.11, 3.12, 3.13 show the training and testing errors using different parameters for the training process. Fig. 3.11 shows that initial step size does not affect the values of MTE, while it does affect the training epochs when the MTE appears. The larger the initial step size, the earlier the MTE comes. For data set trn1 and test1, with initial step size 0.01, the MTE occurs at 74 epochs and for data set trn4 and test4, with initial step size 0.07, MTE occurs at 22 epochs. Fig. 3.12 presents the effect of step size increase rate, SSINC, on the training and testing errors. A similar conclusion can be drawn as the impact of initial step size. The larger the increase rate, the faster the FIS achieves MTE. Fig. 3.13 shows that

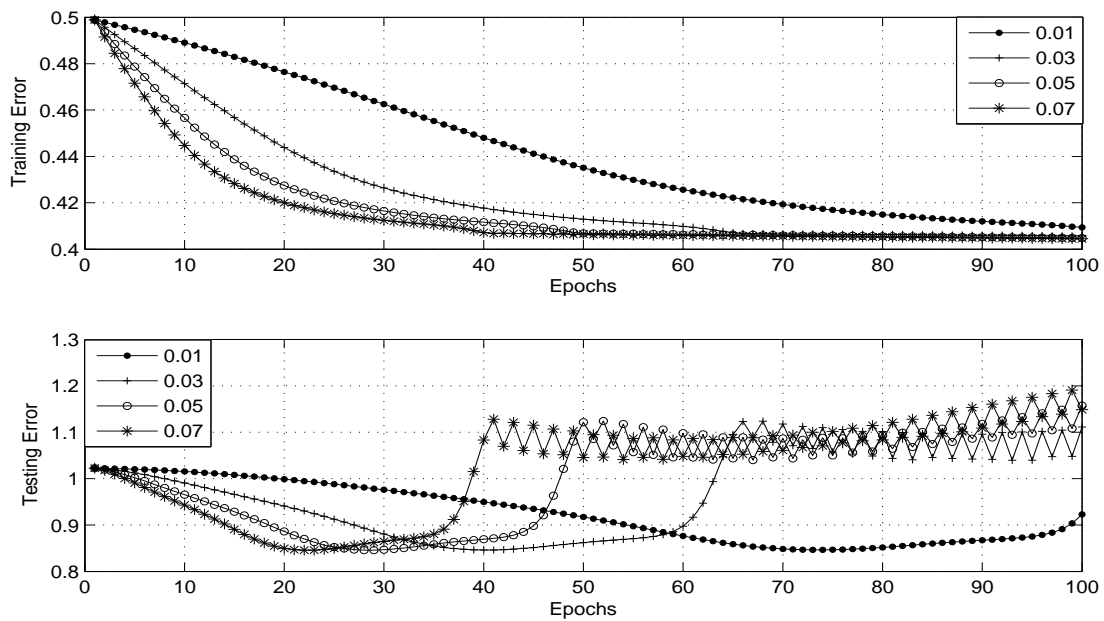


**Figure 3.9:** Training and Testing Errors Obtained by the Neuro-fuzzy Model using Different Number of MFs for Forced Warm-Air Heating System.



**Figure 3.10:** Training and Testing Errors Obtained by the Neuro-fuzzy Model using Different Types of MFs for Forced Warm-Air Heating System.

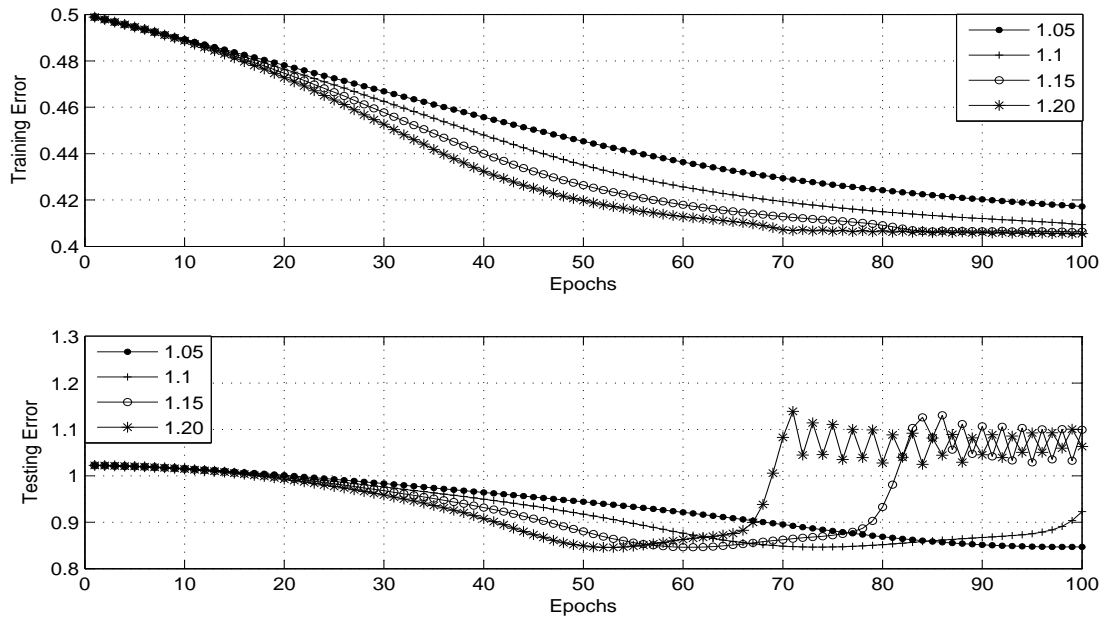
no significant difference exists for FIS obtained using different step size decrease rates. For four different SSDEC values as 0.80 for first data set trn1 and test1, 0.85 for second data set trn2 and test2, 0.90 for third data set trn3 and test3 and 0.95 for fourth data set trn4 and test4, the testing and training error curves are approximately same and all the four curves in each subplot are superimposing each other. Based on this analysis, the training process parameters selected are: SS=0.01, SSINC=1.1, SSDEC=0.9, number of training data pairs = 1801 [118].



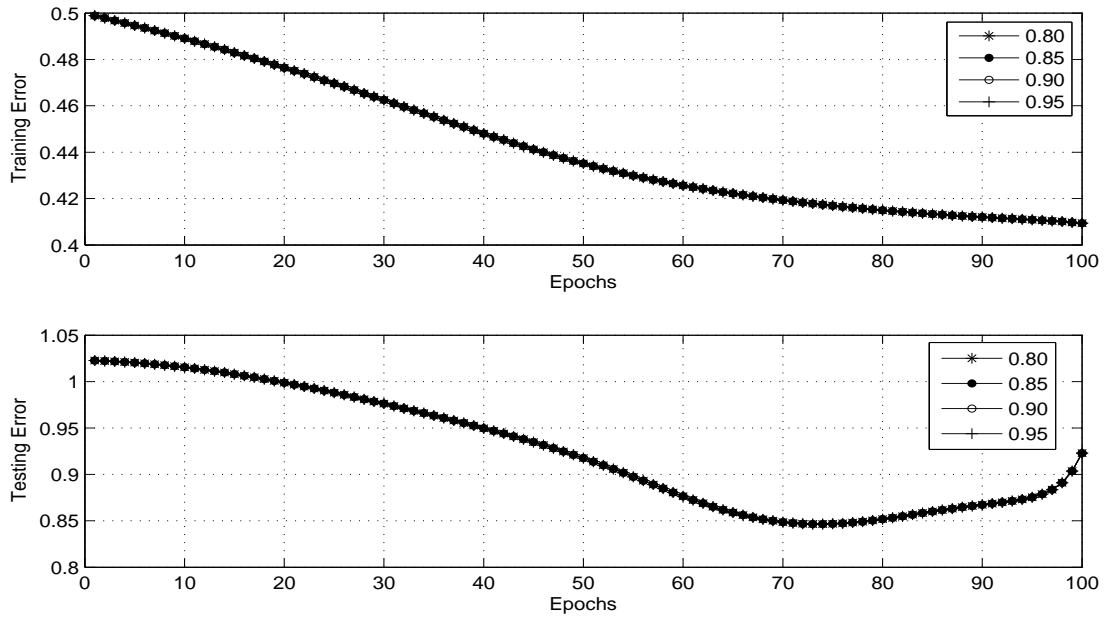
**Figure 3.11:** Training and Testing Errors Obtained by the Neuro-Fuzzy Model for Different Step Sizes for Hot-Water Heating System.

#### *Forced Warm-Air Heating System*

Figs. 3.14, 3.15, 3.16 show the training and testing errors using different parameters for training process. Fig. 3.14 shows that the initial step size does not affect the values of MTE, while it does affect the training epochs when the MTE appears. The larger the initial step size, the earlier the MTE comes. For data set trn1 and test1, with initial step size 0.01, the MTE occurs at 60 epochs and for data set trn4 and test4, with initial step size 0.07, MTE occurs at 16 epochs. Fig. 3.15 presents the effect of step size increase rate, SSINC, on the

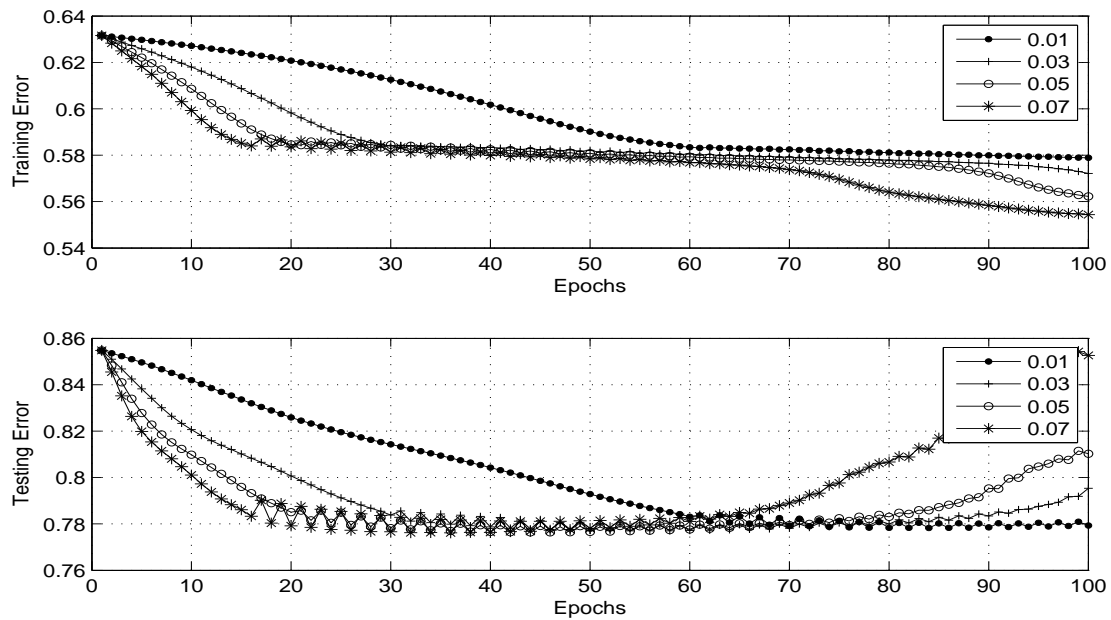


**Figure 3.12:** Training and Testing Errors Obtained by the Neuro-Fuzzy Model for Different Step Size Increase Rates for Hot-Water Heating System.



**Figure 3.13:** Training and Testing Errors Obtained by the Neuro-fuzzy Model for Different Step Size Decrease Rates for Hot-Water Heating System.

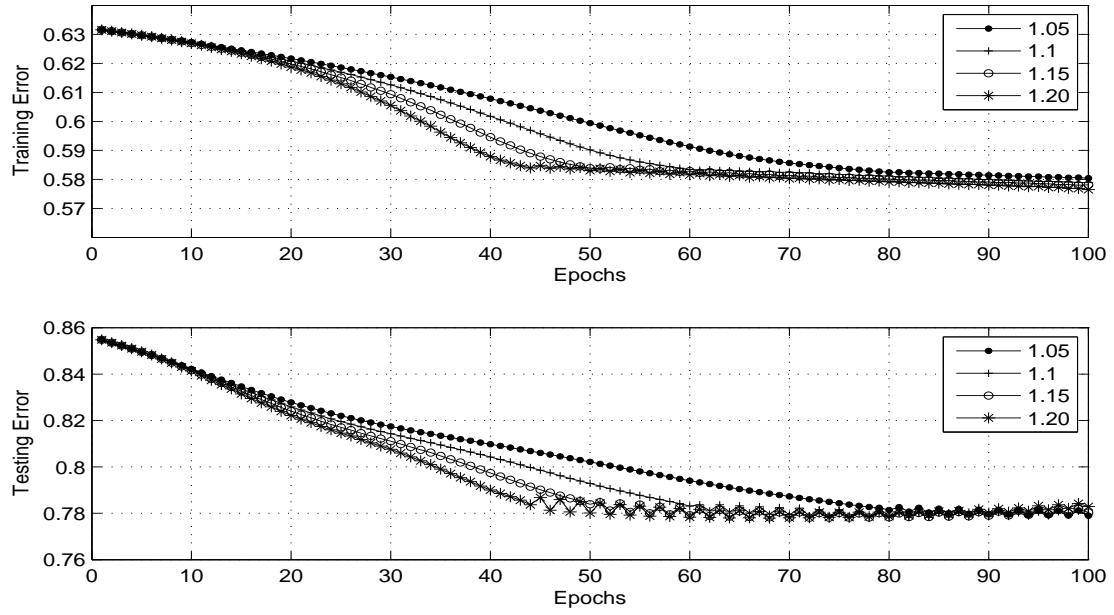
training and testing errors. A similar conclusion can be drawn on the impact of initial step size. The larger the increase rate, the faster the FIS achieves MTE. Fig. 3.16 shows that no significant difference exists for FIS obtained using different step size decrease rates. For four different SSDEC values as 0.80 for first data set trn1 and test1, 0.85 for second data set trn2 and test2, 0.90 for third data set trn3 and test3 and 0.95 for fourth data set trn4 and test4, the testing and training error curves are approximately same and all the four curves in each subplot are superimposing each other. Based on this analysis and choosing MTE as the deciding criterion, the training process parameters selected are: SS=0.01, SSINC=1.1, SSDEC=0.85, number of training data pairs = 3000, number of epochs = 59.



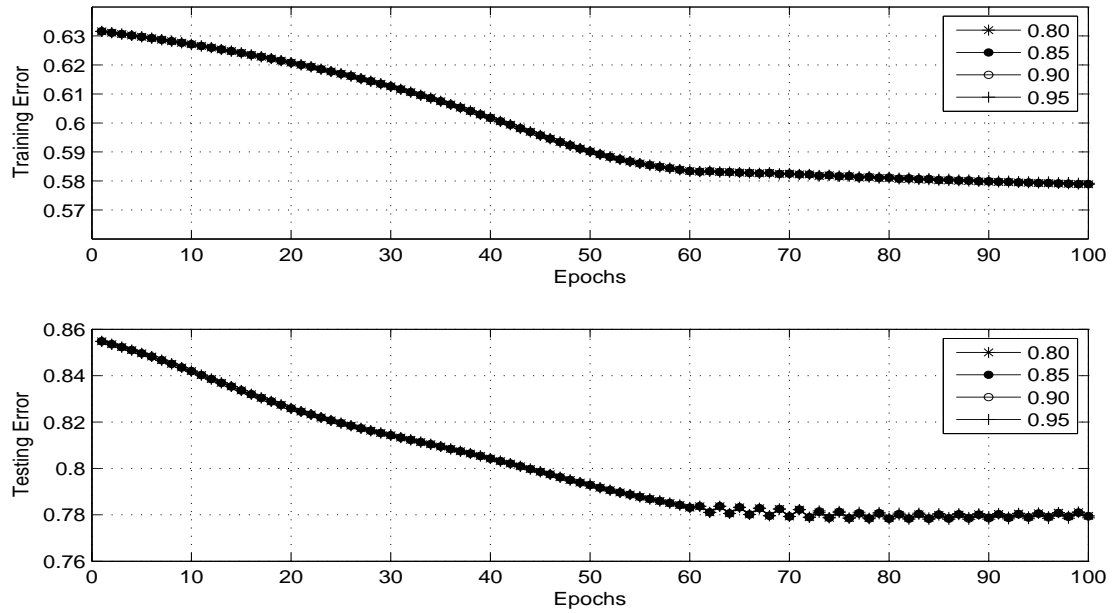
**Figure 3.14:** Training and Testing Errors Obtained by the Neuro-Fuzzy Model for Different Step Sizes for Forced Warm-Air Heating System.

Table 3.1 present the architectures and training parameters for both type of models for the two types of space heating systems.

The training and testing data sets are chosen from the experimental data collected from the laboratory and the residential heating systems. Statistical analysis is conducted to check if the range of the training and the testing data sets is closely related, otherwise if



**Figure 3.15:** Training and Testing Errors Obtained by the Neuro-Fuzzy Model for Different Step Size Increase Rates for Forced Warm-Air Heating System.



**Figure 3.16:** Training and Testing Errors Obtained by the Neuro-fuzzy Model for Different Step Size Decrease Rates for Forced Warm-Air Heating System.



**Table 3.1:** Model Architectures and Training Parameters.

Parameters	Hot-Water Heating System		Forced Warm-Air Heating System	
	ANFIS-GRID	ANFIS-SUB	ANFIS-GRID	ANFIS-SUB
Number of layers	5	5	5	5
Number of inputs	3	3	3	3
Size of training data	$1801 \times 4$	$1801 \times 4$	$3000 \times 4$	$3000 \times 4$
Type of MF	Gaussian	Gaussian	Gaussian	Gaussian
Number of MFs	4	4	5	5
Number of fuzzy rules	64	4	125	5
Training Algorithm	Hybrid Learning Algorithm	Hybrid Learning Algorithm	Hybrid Learning Algorithm	Hybrid Learning Algorithm
Number of epochs	74	60	59	60
Size of testing data	$7132 \times 4$	$7132 \times 4$	$9872 \times 4$	$9872 \times 4$

training data set range is different than the testing data set range, the prediction accuracy will decrease. Four statistical indices, minimum value, maximum value, mean and standard deviation are used. Table 3.2 shows the values of these indices for training and testing data sets. The values show that the training data set range is closely related to the testing data set range.

**Table 3.2:** Statistical Analysis between Training and Testing Data Sets.

	Hot-Water Heating System							
	$Q_{in}(W)$		$Q_{sol}(W)$		$T_0(^{\circ}C)$		$T_{avg}(^{\circ}C)$	
	Training	Testing	Training	Testing	Training	Testing	Training	Testing
Min.	262	0	0	0	-0.8	-3.0	16.75	16.61
Max.	14253	15024	12252	11979	6	10.4	21.37	21.59
Mean	8652	9509	1535	1129	3.32	4.42	18.33	18.61
SD	2687	2452	2552	2052	1.69	2.33	1.35	1.46
	Forced Warm-Air Heating System							
	$Q_{in}(W)$		$Q_{sol}(W)$		$T_0(^{\circ}C)$		$T_{avg}(^{\circ}C)$	
	Training	Testing	Training	Testing	Training	Testing	Training	Testing
Min.	0	0	-97.5	-97.71	-13.8	-13.79	14.82	13.87
Max.	23127	25356.84	346.5	346.8	3.101	7.9309	16.99	17.31
Mean	8762	9670	14.76	19.57	-5.91	-3.35	15.6	15.87
SD	2971	2752	106.96	87.18	4.41	4.61	0.66	0.61

### 3.3 Model Validation

#### 3.3.1 Performance Indicators

Various statistical indices are proposed in the literature to check the predictive performance of the models [10,119]. In this research the model performance is measured using the following statistical indicators to provide a numerical description of the goodness of the estimates. One of the most common indicators used with hybrid models is the root mean square error (RMSE). This is calculated according to equation (3.4):

$$RMSE = \sqrt{\frac{1}{T} \sum_{i=1}^T (K_i - L_i)^2} \quad (3.4)$$

where  $T$  is the number of data points,  $L_i$  the measured data point and  $K_i$  the estimated data point.

The Coefficient of determination ( $R^2$ ), tells us how much of the experimental variability is accounted for by the estimate model.  $R^2$  is calculated according to equation (3.5).

$$R^2 = \frac{\sum_{i=1}^T [K_i - \bar{L}]^2}{\sum_{i=1}^T [L_i - \bar{L}]^2} \quad (3.5)$$

where  $\bar{L}$  is the average of measured data.

#### 3.3.2 Results

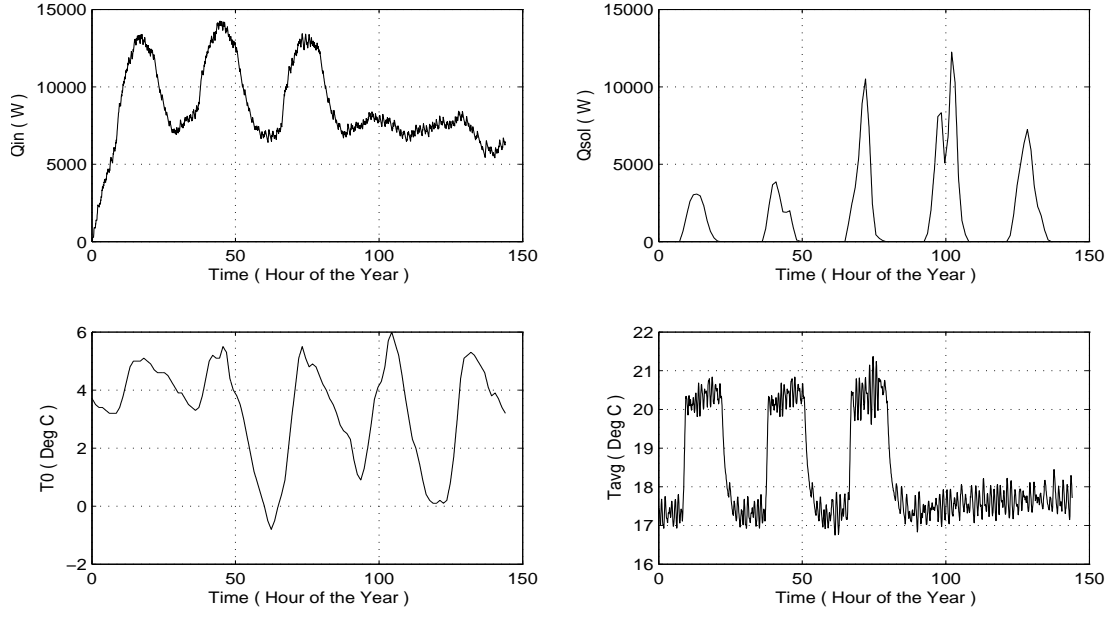
##### Hot-Water Heating System

###### *ANFIS-SUB*

The developed ANFIS-SUB model has been trained with six days of the experimental data pairs as shown in Fig. 3.17. The following parameters are selected for the learning process: SS=0.01, SSINC=1.1, SSDEC=0.9, number of epochs = 60.

Fig. 3.18 presents a comparison between the measured and estimated values of the average air temperature during training phase.

The overall error function of the network is expressed as follows:



**Figure 3.17:** Training Data (day 21 to day 27: February 2000).

$$E = \frac{1}{T} \sum_{i=1}^T E_i^2 = \frac{1}{T} \sum_{i=1}^T \left( \hat{T}_{avg} - T_{avg} \right)^2 \quad (3.6)$$

where  $\hat{T}_{avg}$  is the ANFIS estimated output and  $T_{avg}$  is the experimentally measured output.

If  $\omega$  is a parameter of the network, then simplifying Eq. (3.6) further,

$$\frac{\partial E}{\partial \omega} = \sum_{i=1}^T \frac{\partial E_i}{\partial \omega} \quad (3.7)$$

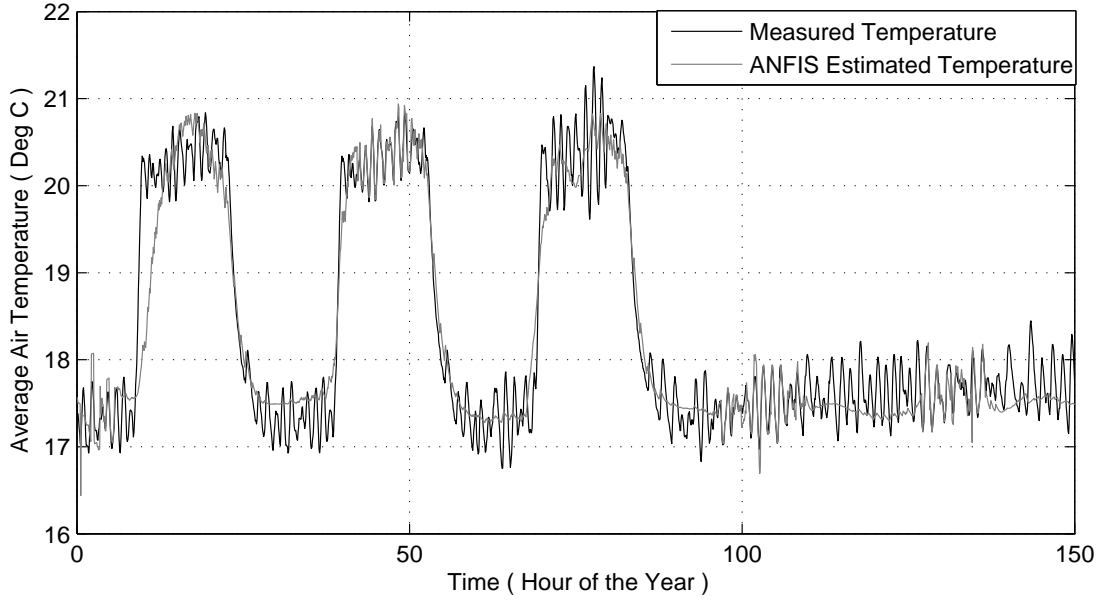
Accordingly, the updated formula for the parameter  $\omega$  is

$$\Delta \omega = -\eta \frac{\partial E}{\partial \omega} \quad (3.8)$$

In Eq. (3.8),  $\eta$ , is a learning rate and is given by,

$$\eta = \frac{SS}{\sqrt{\sum_{\omega} \left( \frac{\partial E}{\partial \omega} \right)}} \quad (3.9)$$

where SS, step size, is the length of each gradient transition in the parameter space.



**Figure 3.18:** Training Data and ANFIS-SUB Output after Training the Model.

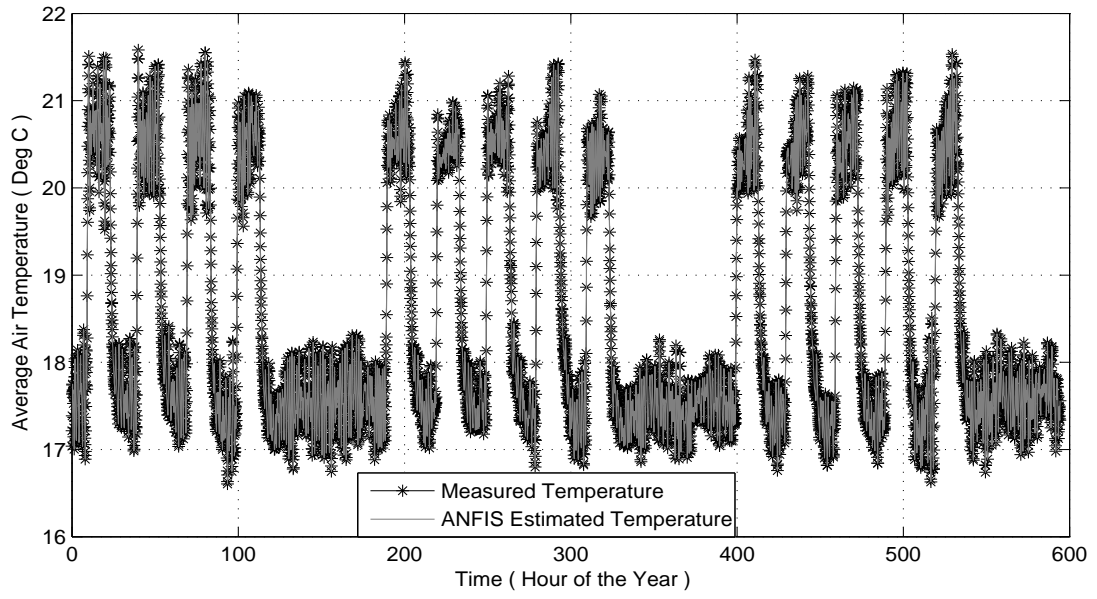
The strong agreement between measured and estimated temperature values in Fig. 3.18 indicates that it is possible to tune the model so that it can represent the real situation during the learning phase. Fig. 3.19 presents the comparison for the testing data. Good agreement can be observed. This indicates that the model used for estimation is correctly structured and can effectively estimate the building air temperature [120].

#### *ANFIS-GRID*

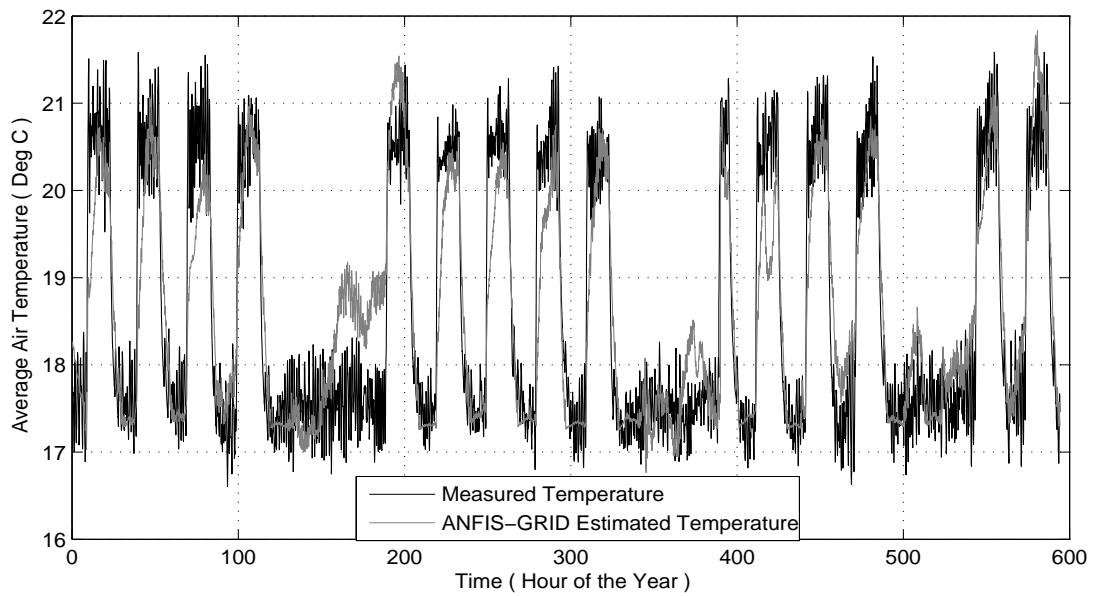
The developed ANFIS-GRID model has been trained with the same set of training data, as shown in Fig. 3.17. The training parameters are similar to the ANFIS-SUB as:  $SS=0.01$ ,  $SSINC=1.1$ ,  $SSDEC=0.9$ , number of epochs = 74.

Fig. 3.20 presents the comparison for the testing data. Good agreement can be observed for this structure as well [118].

Both the models are tested for the year around testing data collected from a laboratory heating system. Table 3.3 gives the RMSE and  $R^2$  values for eight different testing data sets. Testing data sets I to VI are for the year 2000 and testing data sets VII and VIII are for the year 2001. For February 2000 the testing data set is day 1 to day 20 long and rest of



**Figure 3.19:** Comparison of Experimental Output and ANFIS Output.



**Figure 3.20:** Average Air Temperature Estimation using ANFIS-GRID.

the testing data sets are a month long time.

**Table 3.3:** Hot-Water Heating System Results.

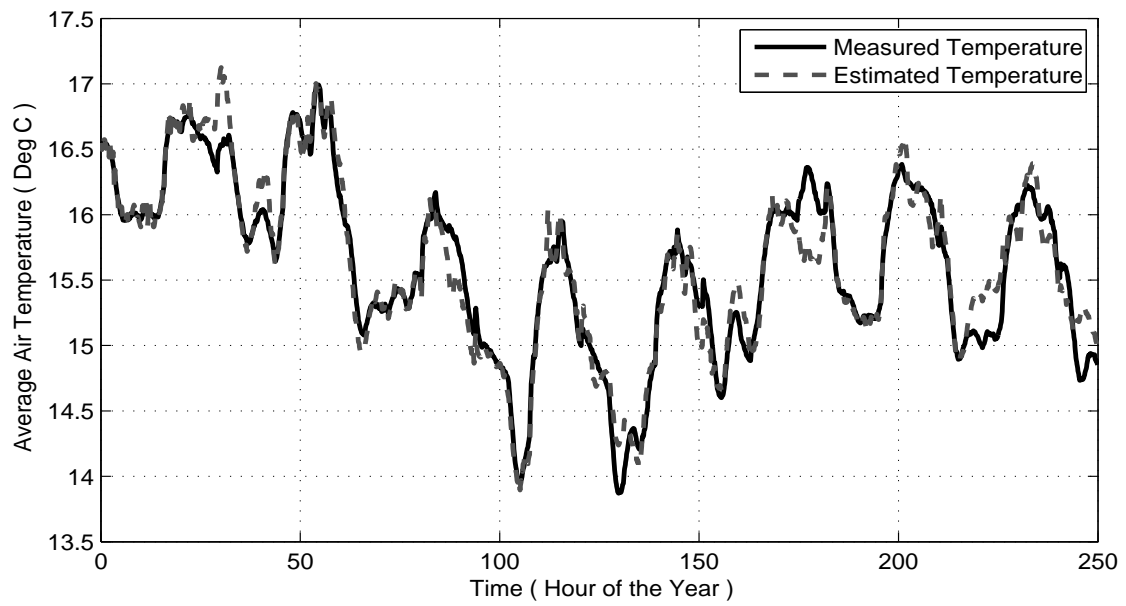
Test Data Set	RMSE ( $^{\circ}\text{C}$ )	
	ANFIS-GRID	ANFIS-SUB
Testing Data Set I (January)	0.3542	0.573
Testing Data Set II (February)	0.2188 (Fig. 3.19)	0.566 (Fig. 3.20)
Testing Data Set III (March)	0.2092	0.552
Testing Data Set IV (April)	0.3671	0.634
Testing Data Set V (November)	0.4982	0.632
Testing Data Set VI (December)	0.2175	0.534
Testing Data Set VII (January)	0.5782	0.743
Testing Data Set VIII (February)	0.3459	0.5672

### Forced Warm-Air Heating System

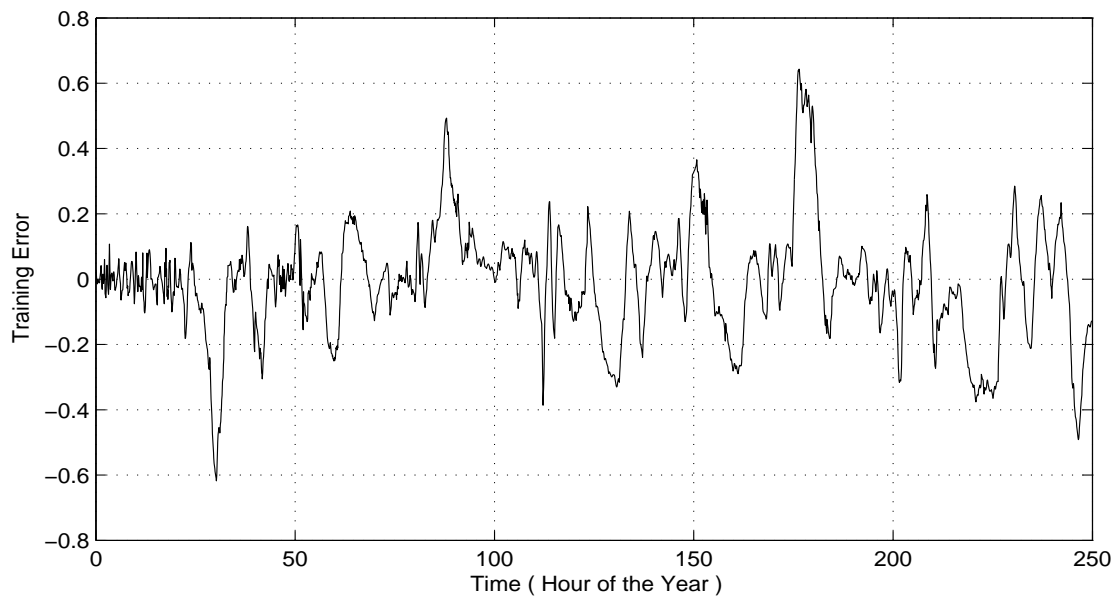
After applying both types of structures to the Hot-Water heating system application, ANFIS-GRID is applied to the Forced Warm-Air heating systems, as this application has only three input variables and structure will not face the problem of “curse of dimensionality”.

The ANFIS-GRID model is trained with following parameters for 59 epochs:  $SS=0.01$ ,  $SSINC=1.1$ ,  $SSDEC=0.85$ , number of training data pairs = 3000. The comparison of the measured and estimated average air temperature values for the training data set is shown in Fig. 3.21. The graph in Fig. 3.22 shows the error for each training data pair.

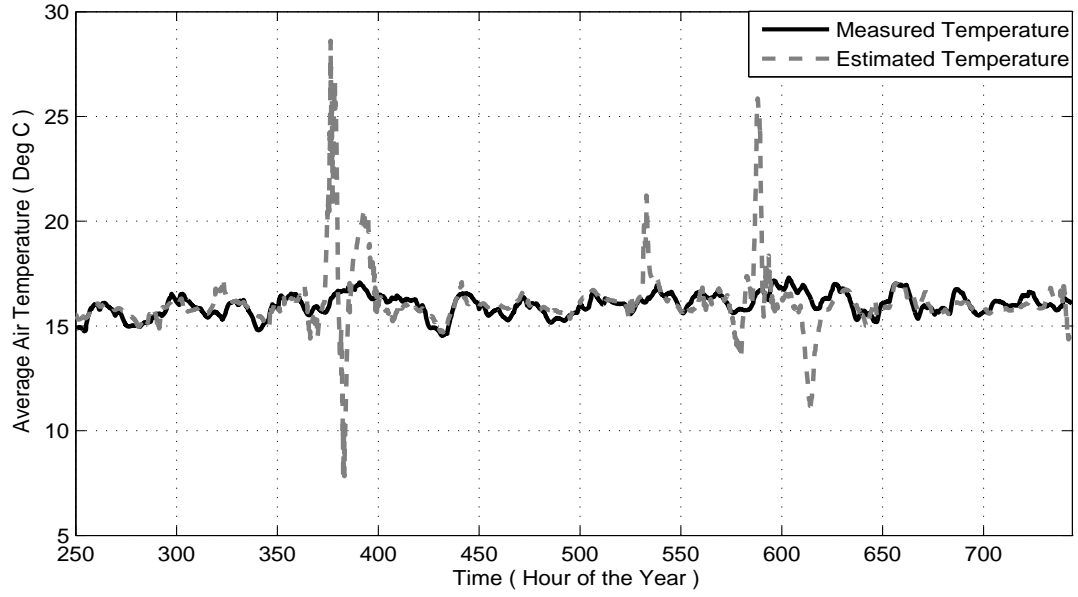
The testing data, as shown in Fig. 3.4, is used for checking the performance of the trained model. It is clear from Fig. 3.23, for few data pairs the error is quite large of the order of  $11^{\circ}\text{C}$  and  $-13^{\circ}\text{C}$ . This is because the dynamics of the building thermal system is not captured by the model.



**Figure 3.21:** Training Results.



**Figure 3.22:** Training Error.



**Figure 3.23:** Testing Results without Feedback.

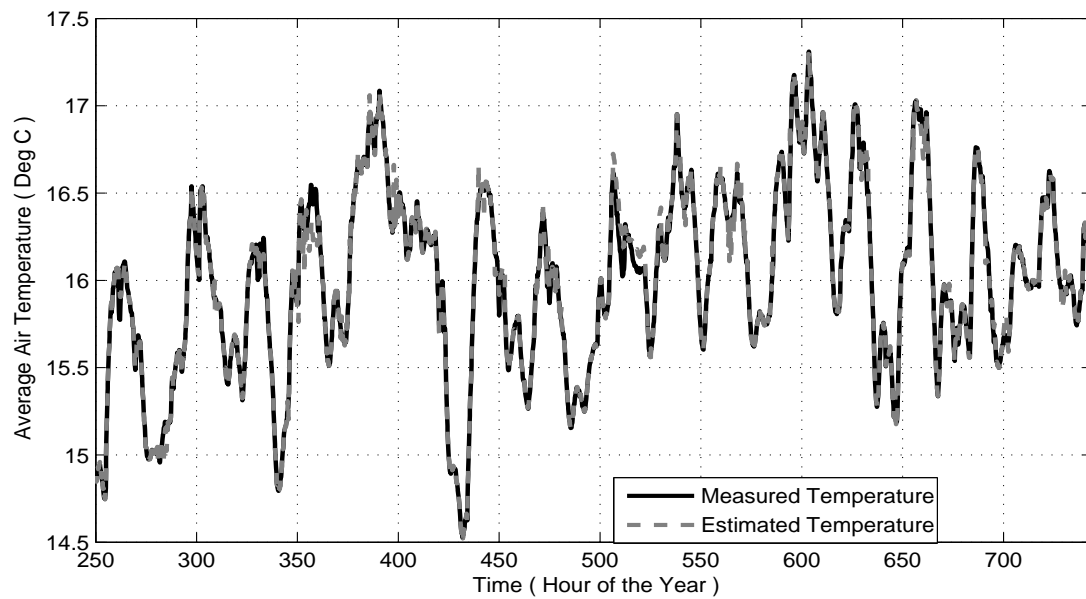
To resolve this problem the feedback loops are implemented using time-delay element at the output and the model performance is tested for the same testing data sets. The previous value of the predicted average air temperature is fed back, the resulted recurrent neuro-fuzzy system structure is trained. Fig. 3.24 shows a comparison of the measured and estimated average air temperatures for the same testing data set of Fig. 3.4. The graph in Fig. 3.25 shows the testing error for each testing data pair. As compared to the static model, the maximum error range is reduced to  $0.35^{\circ}C$  to  $-0.75^{\circ}C$ .

Nine different sections of the experimental data obtained from the same experimental set on a residential house have been tested. Table 3.4 gives the values of RMSE, model performance indicator, for these testing data sets. Testing data sets I to VI are for year 2008 and testing data sets VII to IX are for year 2009.

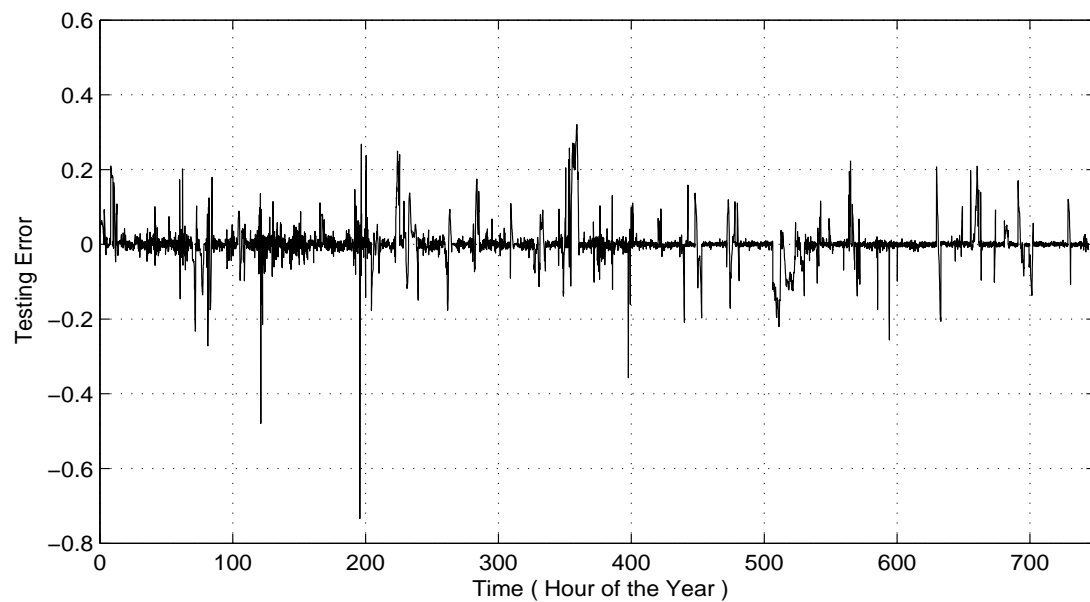
### 3.4 Concluding Remarks

An average air temperature estimator model based on feedback neuro-fuzzy system for capturing the dynamic properties of the space heating system is structured and trained. The





**Figure 3.24:** Testing Results with Feedback System.



**Figure 3.25:** Testing Error with Feedback System.

**Table 3.4:** Forced Warm-Air Heating System Results.

Test Data Set	RMSE ( $^{\circ}\text{C}$ )	
	Hybrid Neuro-fuzzy Model without Feedback	Hybrid Neuro-fuzzy Model with Feedback
Testing Data Set I (January)	1.69	0.045
Testing Data Set II (February)	1.45	0.041
Testing Data Set III (March)	1.53	0.043
Testing Data Set IV (April)	1.72	0.057
Testing Data Set V (October)	1.42	0.040
Testing Data Set VI (November)	1.61	0.043
Testing Data Set VII (December)	1.80	0.062
Testing Data Set VIII (January)	1.59	0.045
Testing Data Set IX (February)	1.39	0.040

developed model can accurately estimate the average air temperature based on the information available to the heating equipment once it is trained using short-term monitoring data. The results show that the dynamical behavior of each system can be better represented using feedback systems.

The inferential model performance, measured by RMSE, is significantly improved with the implementation of the feedback loop. The positive and negative peak for error is also reduced from  $11^{\circ}C$  to  $0.35^{\circ}C$  and from  $-13^{\circ}C$  to  $-0.75^{\circ}C$  respectively. The estimation results show that the model is correctly structured and trained and is suitable for use as an inferential model for incorporation of soft sensing control scheme to the conventional control systems in the residential buildings.

# Chapter 4

## Impact of Data Quality on Model Performance

### 4.1 Introduction

Data quality is generally recognized as a multidimensional concept [121]. While no single definition of data quality has been accepted by researchers working in this area, there is agreement that data accuracy, currency, completeness, and consistency are important areas of concern [122–125]. This research is primarily considering the data accuracy, defined as conformity between a recorded value and the actual data value.

Several studies have investigated the effect of data errors on the outputs of the computer based models. Bansal et al. studied the effect of errors in the test data on the predictions made by neural network and linear regression models [126]. The training data set applied in the research was free of errors. The research concluded that the error size had a statistically significant effect on the predictive accuracy of both linear regression and neural network models.

O’Leary investigated the effect of data errors in the context of a rule-based artificial intelligence system [127]. He presented a general methodology for analyzing the impact of data accuracy on the performance of an artificial intelligence based system designed to generate rules from data stored in a database. The methodology can be applied to artificial intelligence systems that analyze the data and generate a set of rules of the form “if X then

Y". It is often assumed that a subset of the generated rules is added to the system's rule base on the basis of the measure of the "goodness" of each rule. O'Leary showed that the data errors can affect the subset of rules that are added to the rule base and that inappropriate rules may be retained while useful rules are discarded if data accuracy is ignored. Klein has conducted experimental analysis to see the effect of training and testing data errors on the performance of a neural network model [128, 129].

Wei et al. analyzed the effect of data quality on the predictive accuracy of ANFIS model [130]. The ANFIS model is developed for predicting the injection profiles in the Daqing Oilfields, China. They concluded that the cleaning of the data has improved the accuracy of ANFIS model from 78% to 86.1%.

In this research, the hybrid neuro-fuzzy based inferential model is trained and tested using the experimental data collected from a laboratory heating system. The data collected has some uneven patterns. In this Chapter we will discuss the experiments conducted to examine the impact of data quality on the predictive performance of the inferential model.

## 4.2 Experimental Methodology

Data errors may affect the accuracy of the ANFIS based models in two ways. First, the data used to build and train the model may contain errors. Second, even if training data are free of errors, once the developed model is used for estimation tasks a user may use input data containing errors to the model.

The research in this area has assumed that data used to train the models and data input to make estimation of the processes are free of errors. In this research we relax this assumption by asking two questions: (1) What is the effect of errors in the test data on the estimation accuracy of ANFIS based models? (2) What is the effect of errors in the training data on the predictive accuracy of ANFIS based models?

While many sources of error in a data set are possible, we assume that the underlying cause of errors affect data items randomly rather than systematically. One source of inaccuracy that may effect a data set in this way is the measurement errors caused by reading the

equipment. This type of error may affect any data item in the data set and may understate or overstate the actual data value. This study does not address the effect of systematic data errors on the estimations made by the ANFIS based models.

Two experiments are conducted to examine the research targets. Both the experiments used the same application and the same data set.

- *Experiment 1 examines the first question:* What is the effect of errors in the test data on the estimation ability of ANFIS based models?
- *Experiment 2 examines the second question:* How do errors in the training data effect the accuracy of the ANFIS based models?

### 4.2.1 Experimental Factors

There are two factors in each experiment: (1) fraction-error and (2) amount-error. Fraction-error is the percent of the data items in the appropriate part of the data set (the test data in experiment 1 and the training data in experiment 2) that are perturbed. Amount-error is the percentage by which the data items identified in the fraction-error factor are perturbed.

#### *Fraction-error*

Since fraction-error is defined as a percent of the data items in a data set, the number of data items that are changed for a given level of fraction-error is determined by multiplying the fraction-error by the total number of data items in the data set.

- *Experiment 1:* The test data used in experiment 1 has four data items (one value for each of the four input and output variables for one entry of the total data pairs). This experiment examines all of the possible number of data items that could be perturbed. These four levels for fraction-error factor are: 25% (one data item perturbed), 50% (two data items perturbed), 75% (three data items perturbed), and 100% (4 data items perturbed)
- *Experiment 2:* For the training data, four levels of the fraction-error factor are tested: 5% (90 and 150 data items are perturbed for hot-water heating system and forced

warm-air heating system, respectively), 10% (180 and 300 data items are perturbed for hot-water heating system and forced warm-air heating system, respectively), 15% (270 and 450 data items are perturbed for hot-water heating system and forced warm-air heating system, respectively), and 20% (360 and 600 data items are perturbed for hot-water heating system and forced warm-air heating system, respectively).

#### *Amount-error*

For both experiments, the amount-error factor has two levels: (1) plus or minus 5% and (2) plus or minus 10%. The amount-error applied to the data set can be represented by the following set of equations:

$$\acute{d} = d \pm 0.05 \times d \quad (4.1)$$

$$\acute{d} = d \pm 0.1 \times d \quad (4.2)$$

For equations (4.1) and (4.2),  $\acute{d}$  is the value of the variable after adding or subtracting the noise error to the unmodified variable  $d$ .

### **4.2.2 Experimental Design**

The experimental design is shown in Table 4.1. Both the experiments have four levels for the fraction-error factor and two levels for the amount-error. For each combination of fraction-error and amount-error, four runs with random combinations of the input and output variable are performed.

Although the levels of fraction-error are different in the two experiments, the sampling procedure is the same. For each fraction-error level, the variables are randomly selected to be perturbed. This is repeated a total of four times per level. Table 4.2 shows the combinations of the variables for experiment 1.

Second, for each level of the amount-error factor, each variable is randomly assigned either a positive or a negative sign to indicate the appropriate amount-error to be applied.

**Table 4.1:** Experimental Design.

<b>Experiment 1 (Errors in the Test Data)</b>	Hot-Water Heating System	Forced Warm-Air Heating System
Fraction-error levels (25%, 50%, 75%, and 100%)	4	4
Amount-error levels (5%, and 10 %)	2	2
Number of random combinations of the variables considered within each fraction-error level	4	4
Total number of samples considered	7132	9872
<b>Experiment 2 (Errors in the Training Data)</b>		
Fraction-error levels (5%, 10%, 15%, and 20%)	4	4
Amount-error levels (5%, and 10 %)	2	2
Number of random combinations of the variables considered within each fraction-error level	4	4
Total number of samples considered	1801	3000

**Table 4.2:** Four Combinations of the Variables for each Fraction Error Level in Experiment 1.

Fraction-Error Level	Input and Output Variable Combination			
	1	2	3	4
25%	$(Q_{in})$	$(Q_{sol})$	$(T_0)$	$(T_{avg})$
50%	$(Q_{in}, T_0)$	$(Q_{in}, T_{avg})$	$(Q_{sol}, T_0)$	$(T_{avg}, T_0)$
75%	$(Q_{in}, T_0, Q_{sol})$	$(Q_{in}, T_{avg}, T_0)$	$(Q_{in}, T_{avg}, Q_{sol})$	$(T_0, T_{avg}, Q_{sol})$
100%	$(Q_{in}, T_0, Q_{sol}, T_{avg})$	$(Q_{in}, T_0, Q_{sol}, T_{avg})$	$(Q_{in}, T_0, Q_{sol}, T_{avg})$	$(Q_{in}, T_0, Q_{sol}, T_{avg})$



Table 4.3 shows the randomly assigned amount-error levels in experiment 1. The procedure for experiment 2 differs only in the number of variables that were randomly selected to be perturbed for the four levels of the fraction-error factor.

**Table 4.3:** Randomly assigned Percentage Increase (+) or Decrease (-) for a given Amount-Error Level in Experiment 1.

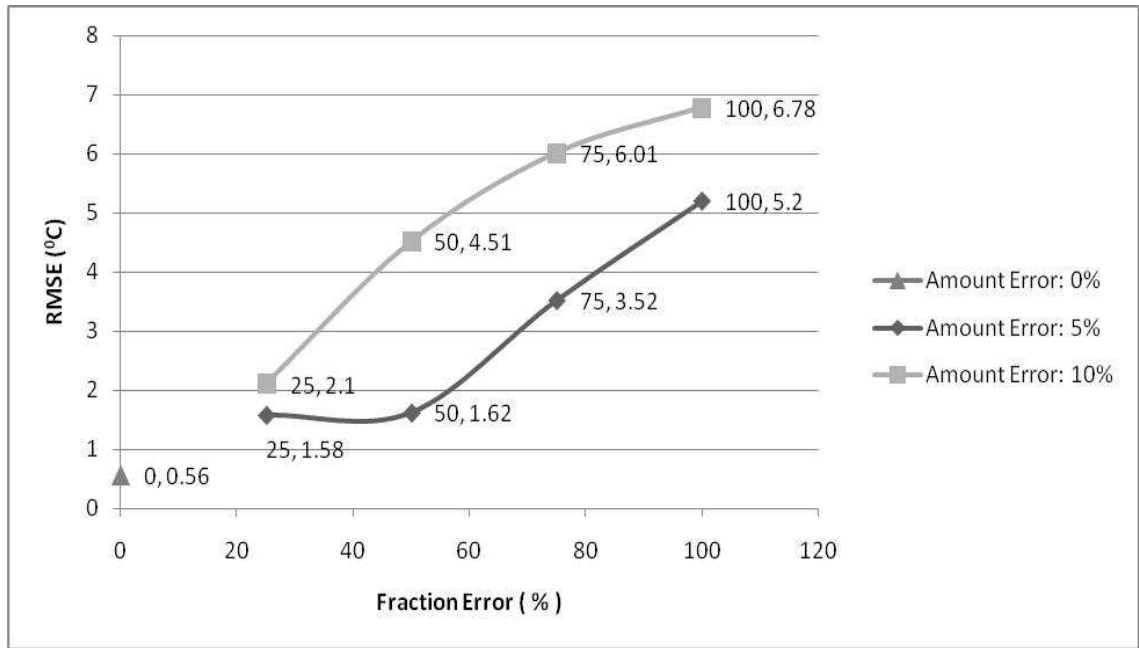
Fraction-Error Level	Input and Output Variable Combination			
	1	2	3	4
25%	$(Q_{in})$ (-)	$(Q_{sol})$ (+)	$(T_0)$ (+)	$(T_{avg})$ (-)
50%	$(Q_{in}, T_0)$ (+, -)	$(Q_{in}, T_{avg})$ (-, -)	$(Q_{sol}, T_0)$ (+, +)	$(T_{avg}, T_0)$ (-, +)
75%	$(Q_{in}, T_0, Q_{sol})$ (+, -, -)	$(Q_{in}, T_{avg}, T_0)$ (-, +, -)	$(Q_{in}, T_{avg}, Q_{sol})$ (+, +, -)	$(T_0, T_{avg}, Q_{sol})$ (+, -, +)
100%	$(Q_{in}, T_0, Q_{sol}, T_{avg})$ (-, +, -, +)	$(Q_{in}, T_0, Q_{sol}, T_{avg})$ (+, +, +, +)	$(Q_{in}, T_0, Q_{sol}, T_{avg})$ (-, -, +, +)	$(Q_{in}, T_0, Q_{sol}, T_{avg})$ (-, -, -, +)

## 4.3 Experimental Results

For both the experiments, the measured average air temperature values and ANFIS-GRID estimated average air temperature values are compared using RMSE index as a measure of estimation accuracy.

### 4.3.1 Experiment 1 Results: Errors in the Test Data

Estimation accuracy results, using the simulated inaccuracies for amount-error and fraction-error for the average air temperature estimation are given in Fig. 4.1. It shows that as fraction-error increases from 25% to 100%, RMSE increases resulting in a decrease in predictive accuracy. As amount-error increases from 5% to 10%, RMSE increases also indicating a decrease in estimation accuracy. Both fraction-error and amount-error have an effect on predictive accuracy.



**Figure 4.1:** RMSE ( $^{\circ}\text{C}$ ) Values as Error Level in the Test Data Varies.

### 4.3.2 Experiment 2 Results: Errors in the Training Data

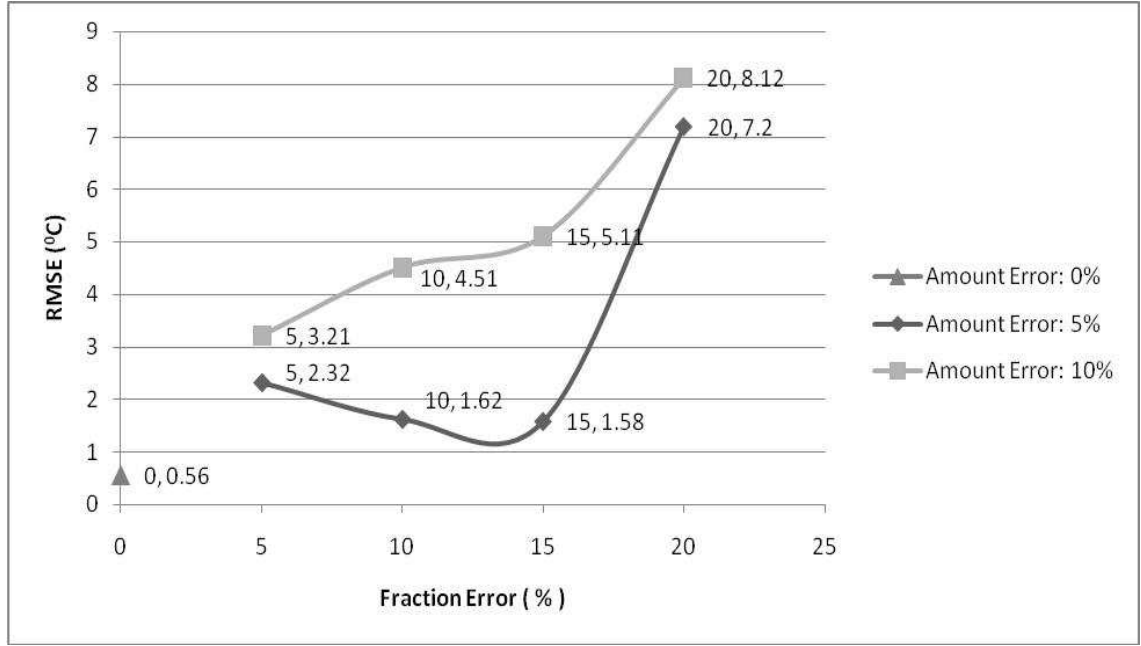
Predictive accuracy results, using the simulated inaccuracies for amount-error and fraction-error for the average air temperature estimation are given in Fig. 4.2. It shows that as fraction-error increases from 5% to 20%, RMSE increases indicating a decrease in predictive accuracy.

## 4.4 TANE Algorithm for Noisy Data Detection

Data quality analysis results show that the errors in the training data as well as in the testing data affect the predictive accuracy of the ANFIS based soft sensors. This section discusses an efficient algorithm, TANE algorithm, to identify the noisy data pairs in the data set.

### 4.4.1 Functional Dependencies

The raw data are analyzed using approximate functional dependence mining method. An approximate dependency, or an approximate functional dependency, is a functional dependency



**Figure 4.2:** RMSE ( $^{\circ}\text{C}$ ) Values as Error Level in the Training Data Varies.

that is almost valid with the exception of data tuples. A functional dependency studies the relationship of attributes in one or several tables, and claims that the value of an attribute is uniquely determined by the values of some other attributes. The discovery of functional dependencies in databases leads to useful knowledge and data quality problems.

More formally, a functional dependency over a relation is expressed as  $X \rightarrow A$ , where  $X \subseteq R$  and  $A \subseteq R$ . The dependency is valid in a given relation  $r$  if for all pairs of records  $t, u \in r$ , the following statements hold: if  $t(B) = u[B]$  for all  $B \in X$ , then  $t(A) = u[A]$ . A functional dependency  $X \rightarrow A$  is trivial if  $A \in X$ . The task in functional dependency mining is to find all minimal non-trivial dependencies that hold in  $r$  [131].

Approximate dependencies arise in many databases when there are natural dependencies between attributes, but some records contain errors and inconsistencies. For example, the relationship between zip code and the combination of city and state in a country. Another example is the social security number (SSN) and a corresponding person residing in the USA. Theoretically, these attributes have consistent relationships, as one person associated with one SSN, and one zip code associated with one combination of city, state in a country. But

if errors are somehow introduced, the relationships between these attributes will be violated, which leads to the approximate dependencies.

#### 4.4.2 TANE Algorithm

The TANE algorithm, which deals with discovering functional and approximate dependencies in large data files, is an effective algorithm in practice [132]. The TANE algorithm partitions attributes into equivalence partitions of the set of tuples. By checking if the tuples that agree on the right-hand side agree on the left-hand side, one can determine whether a dependency holds or not. By analyzing the identified approximate dependencies, one can identify potential erroneous data in the relations.

In this research, the relationship of three input parameters ( $Q_{in}$ ,  $Q_{sol}$ , and  $T_0$ ) and the average air temperature ( $T_{avg}$ ) is analyzed using TANE algorithm. For equivalence partition, all the four parameters are rounded off to zero decimal points.

After data pre-processing, four approximate dependencies are discovered, as shown in Table 4.4. Although all these dependencies reflect the relationships among the parameters, the first dependency is the most important one because it shows that the selected input parameters have consistent association relationship with the average air temperature except a few data pairs, which is a very important dependency for average air temperature estimation.

**Table 4.4:** Approximate Functional Dependencies Detected using the TANE Algorithm.

Index	Approximate Dependencies	Number of Rows with Conflicting Tuples	
		Hot-Water Heating System	Forced Warm-Air Heating System
1	$Q_{in}, Q_{sol}, T_0 \rightarrow T_{avg}$	42	39
2	$Q_{in}, T_0, T_{avg} \rightarrow Q_{sol}$	47	56
3	$Q_{in}, Q_{sol}, T_{avg} \rightarrow T_0$	43	37
4	$Q_{sol}, T_0, T_{avg} \rightarrow Q_{in}$	54	61

To identify exceptional tuples by analyzing the approximate dependencies, it is required to investigate the equivalence partitions of both left-hand and right-hand sides of an approximate dependency. It is non-trivial work that could lead to the discovery of problematic data. By analyzing the first dependency, conflicting tuples are identified. For the first approximate

dependency from Table 4.4, 42 and 39 conflicting data pairs are present which were fixed using data interpolations for better performance of ANFIS-GRID model.

#### 4.4.3 Improvement in Estimation Accuracy

The effect of data quality on the performance of ANFIS-GRID model is checked. The training and testing data sets are cleaned using TANE algorithm. The conflicting data pairs are replaced with the appropriate data pairs using data interpolation. Then the cleaned data set is applied for the training and the testing of ANFIS-GRID model. Table 4.5 clearly show the effect of data quality on predictive accuracy of inferential model. The RMSE is improved by 37.5% [133] and 20% for two types of heating systems, respectively. The predictive accuracy is improved with RMSE decreasing.

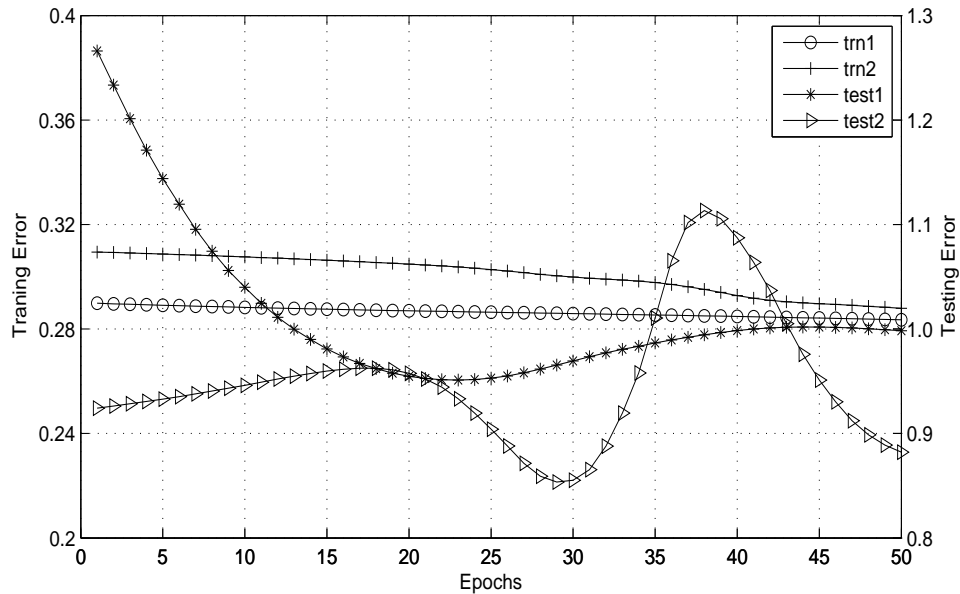
**Table 4.5:** Comparison of Results.

Model	Hot-Water Heating System		Forced Warm-Air Heating System	
	RMSE( $^{\circ}C$ )	$R^2$	RMSE( $^{\circ}C$ )	$R^2$
Inferential model using raw data	0.56	0.7831	1.69	0.43
Inferential model using cleaned data	0.35	0.8945	1.35	0.76

### 4.5 Training Data: Sensitivity and Robustness Analysis

In most of the cases, when modelling practical problems using neuro-fuzzy systems, training data set is selected arbitrarily. But, if the training data is not selected properly, the testing data does not validate the model obtained using the training data, as shown in Fig. 4.3. For the second training and testing data sets, trn2 and test2, although the testing errors (the plot for test2) are not large compared with the plot for test1, the MTE is achieved within the first epoch. This concludes that the testing data presented to ANFIS for testing are sufficiently different from the training data. Hence, the trained FIS does not capture the features of the testing data. Properly selected training data set should have error curves as

trn1 and test1 for any testing data set, testing error decreases with training proceeding until a jump point occurs, as shown in Fig. 4.3. Overfitting occurs when the training passes that point.



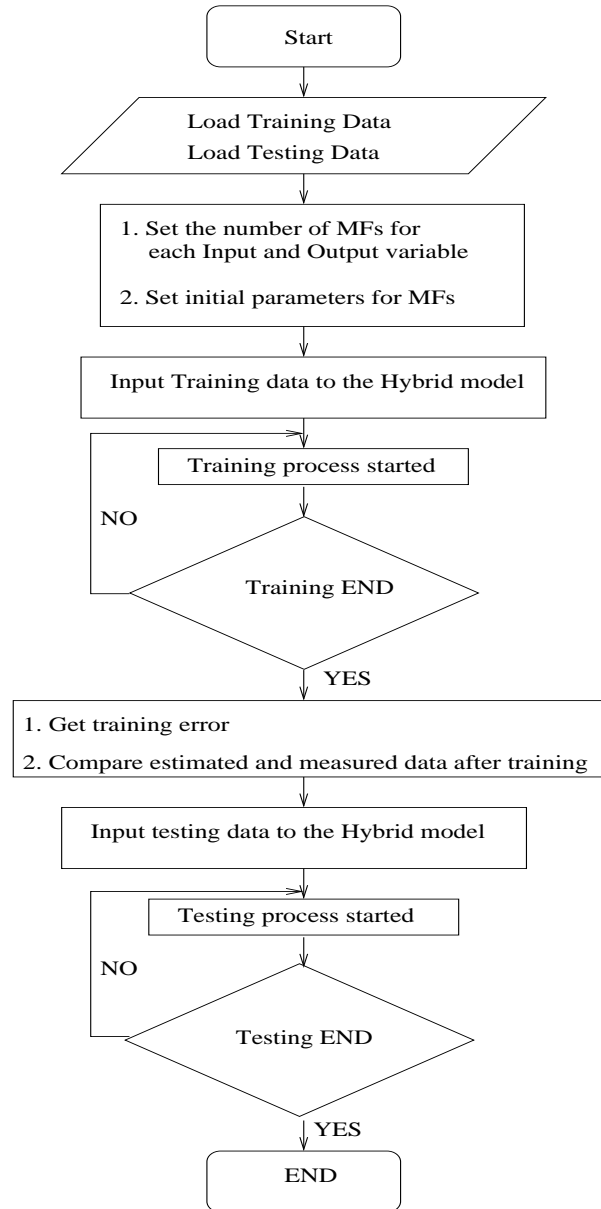
**Figure 4.3:** Training Data Selection.

### 4.5.1 Sensitivity Analysis

Before choosing a training data set, sensitivity analysis is carried out to validate the selection of the training data set. The flow chart in Fig. 4.4 shows the steps a neuro-fuzzy model follows for training and testing of the developed model.

Short-term training data set is selected and model is trained for this data. In sensitivity analysis, three different training data sets are selected. The three data sets represent different conditions while collecting the experimental data.

- *Weekend Only training data set (TD1):* The I/O training data pairs are from the weekend data from Fig. 3.3 and Fig. 3.4 data sets mentioned in Chapter 3. Since the data sets are representing one month data, so there is less number of weekends



**Figure 4.4:** Flow Chart for Training and Testing Process.

as compared to week days. The building temperatures have setback settings for the weekends. The average air temperature variations for the weekends are different than the weekdays. Weekend temperature variations are least representative for the full length one month temperature variations.

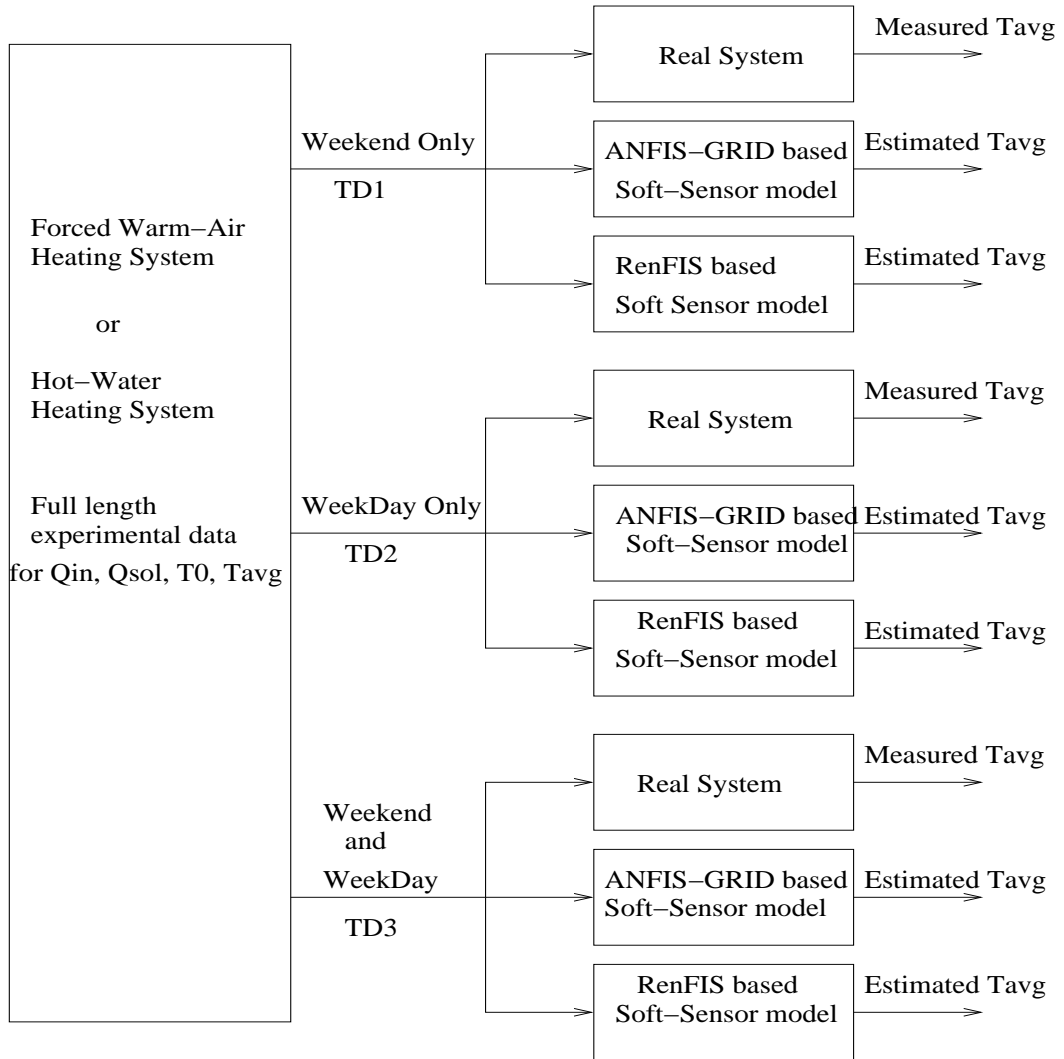
- *Week Day Only training data set (TD2)*: The I/O training data pairs are from weekday data from Fig. 3.3 and Fig. 3.4 data sets mentioned in Chapter 3. For the same month data there are more number of weekdays. The buildings temperatures have setback settings for night time. Weekday data set is more representative for the full one month data as there is more frequency of the similar kind of variations in temperature values.
- *Weekend and Week Day training data set (TD3)*: The I/O training data pairs are from the weekend as well as the week day data from Fig. 3.3 and Fig. 3.4 data sets mentioned in Chapter 3. This data set is most representative out of the three selected training data sets as it is representing the characteristics of both the weekday as well as the weekend temperature variations.

Fig. 4.5 shows the schematic diagram for choosing three different training data sets from the full length experimental data. These data sets are applied to train the developed ANFIS-GRID and RenFIS models. The models follow the steps given in the flow chart for short-term training and long-term testing, as shown in Fig.4.4. The output of the models for long-term testing data is the estimated average air temperature.

#### *Hot-Water Heating System*

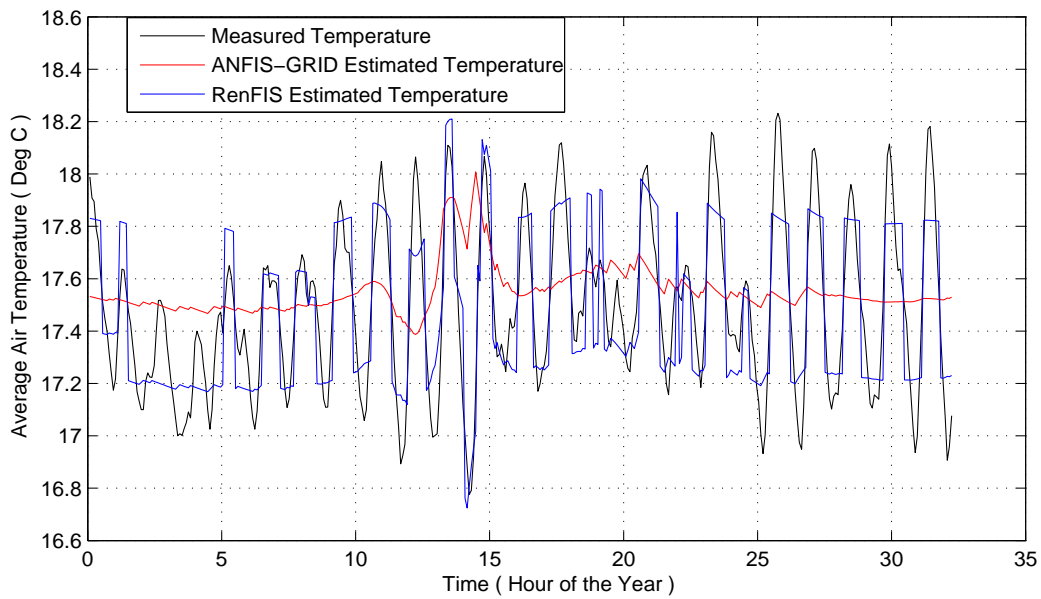
Figs. 4.6, 4.7, and Fig. 4.8 compare the ANFIS-GRID and RenFIS estimated temperatures values with that of the experimentally measured temperature values. These figures show a small part of the whole universe of the testing results. Both models are giving poor performance if those are trained with TD1 training data set as this data set is the least representative for the whole universe of experimentally collected data out of the three training data sets. ANFIS-GRID has the best performance for TD2 as there is more number of weekdays as compared to weekends in the full length of data set. RenFIS has the best





**Figure 4.5:** Schematic Diagram for Sensitivity Analysis.

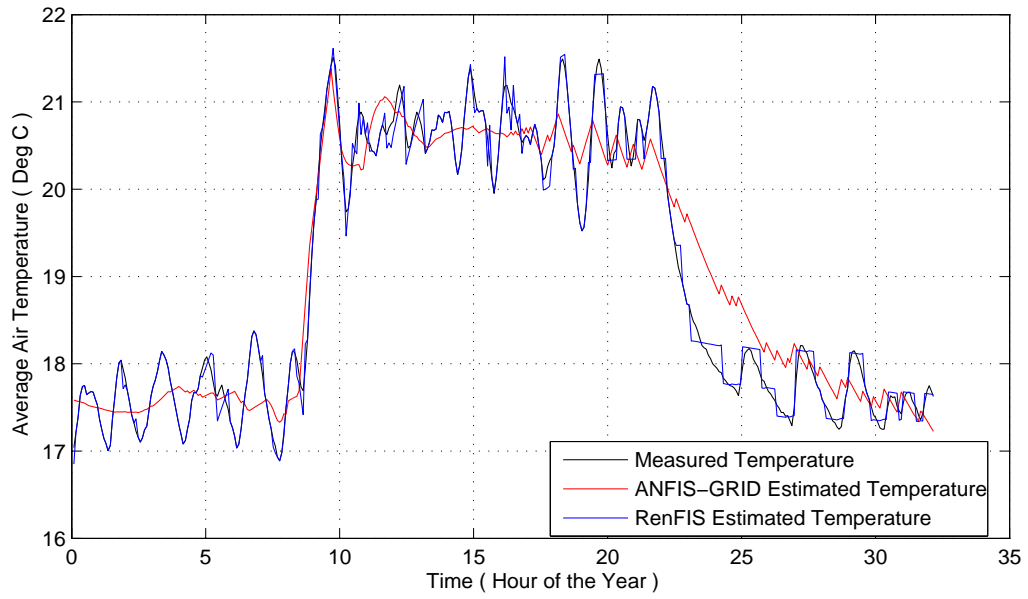
performance if trained with TD3. The analysis concludes that both models are sensitive to the selection of training data set. ANFIS-GRID gives better estimation of the process if trained with data representing most repetitive quality of the whole universe of the data set. While RenFIS, having a loop for feeding back the output, gives better performance if trained with the data set representing overall qualities of the whole experimentally collected data set.



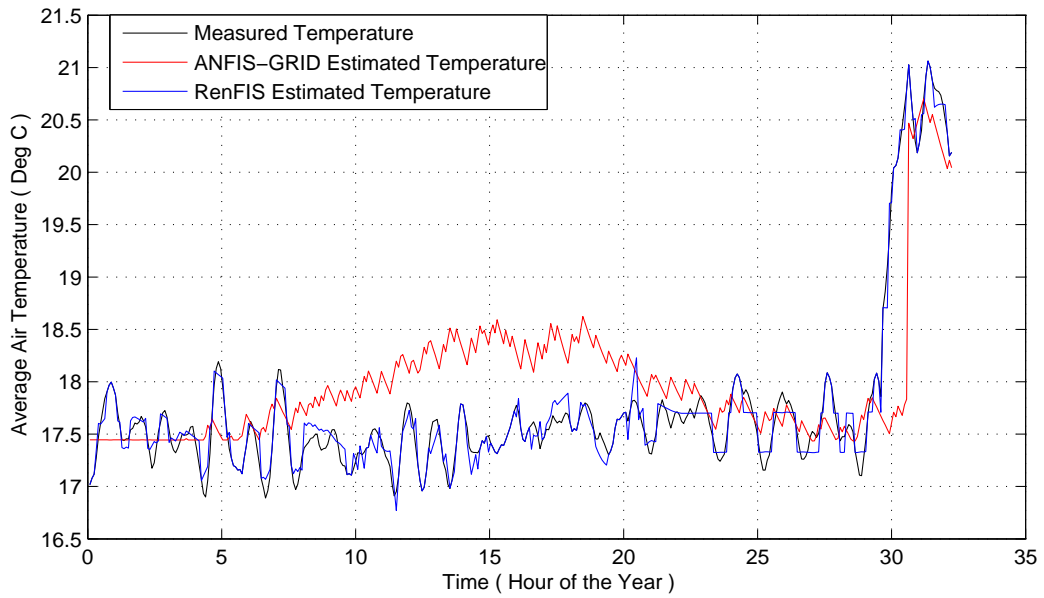
**Figure 4.6:** Comparison of Measured and Estimated Temperatures for the Models Trained with TD1 for Hot-Water Heating System.

#### *Forced Warm-Air Heating System*

Figs. 4.9, 4.10, and Fig. 4.11 compare the ANFIS-GRID and RenFIS estimated temperatures and the experimentally measured temperature values for forced warm-air heating system. These figures present a small part of the whole universe of the testing results. The results follow the same trends as that for hot-water heating system testing results. Both models are giving poor performance if those are trained with TD1. ANFIS-GRID has the best performance for TD2. The analysis concludes that both models are sensitive to the selection of training data set. ANFIS-GRID gives better estimation of the process if trained

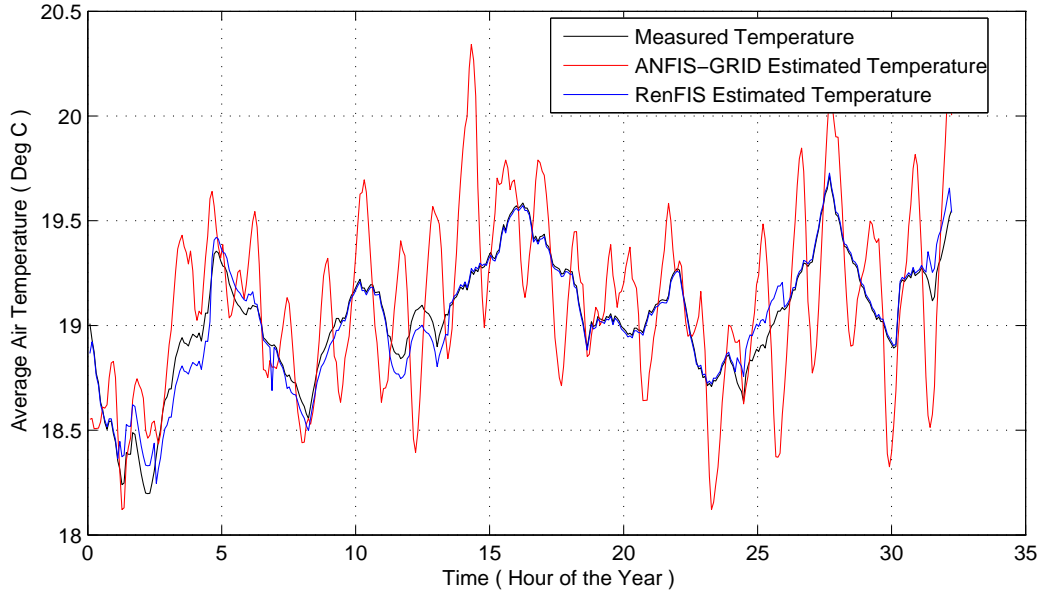


**Figure 4.7:** Comparison of Measured and Estimated Temperatures for the Models Trained with TD2 for Hot-Water Heating System.



**Figure 4.8:** Comparison of Measured and Estimated Temperatures for the Models Trained with TD3 for Hot-Water Heating System.

with data representing the repetitive quality of the whole universe of the data set. RenFIS on the other hand, having a loop for feeding back the output, gives better performance if trained with the data set representing overall qualities of the whole experimentally collected data set.

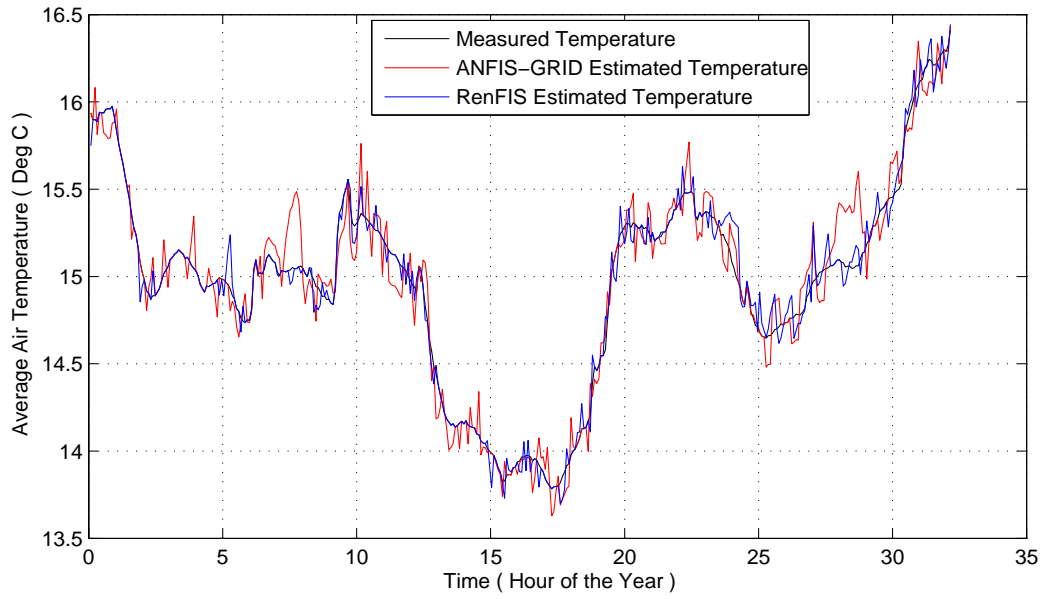


**Figure 4.9:** Comparison of Measured and Estimated Temperatures for the Models Trained with TD1 for Forced Warm-Air Heating System.

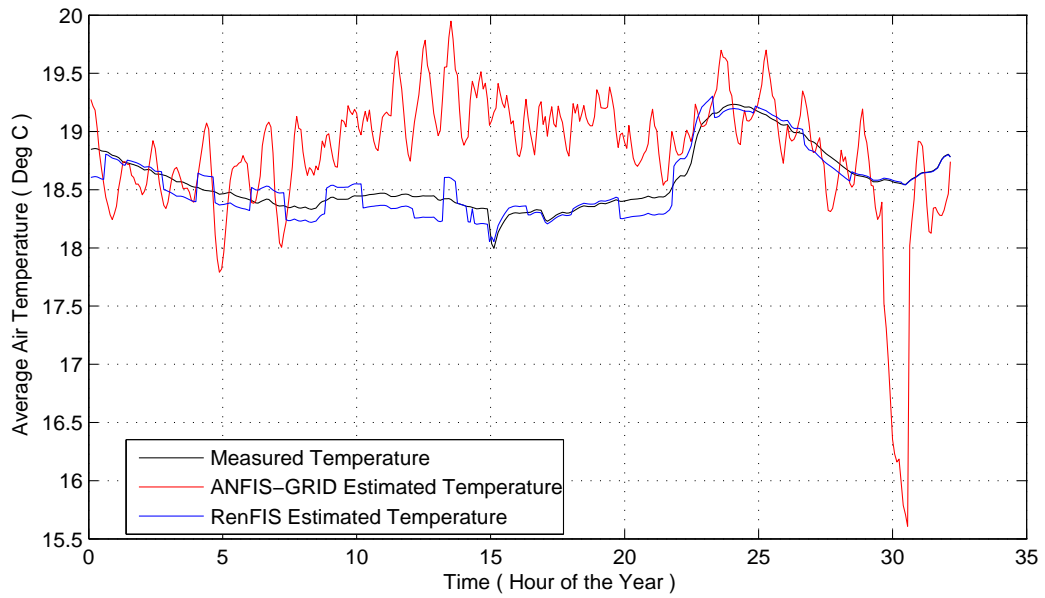
All the training data sets are of the equal length. TD1 is the least representative for the entire universe of experimentally collected data, as it represents the temperatures under setback condition only. TD2 is not including the setback temperature readings, but as there are more number of weekdays as compared to weekends, TD2 represents most repetitive characteristic of the entire data. TD3 has a combination of weekend and weekday, covers more characteristics of the process under control.

#### 4.5.2 Robustness Analysis

Final target of the research is to make use of the developed soft-sensor model in the development of robust control schemes for the building heating systems. The robustness analysis

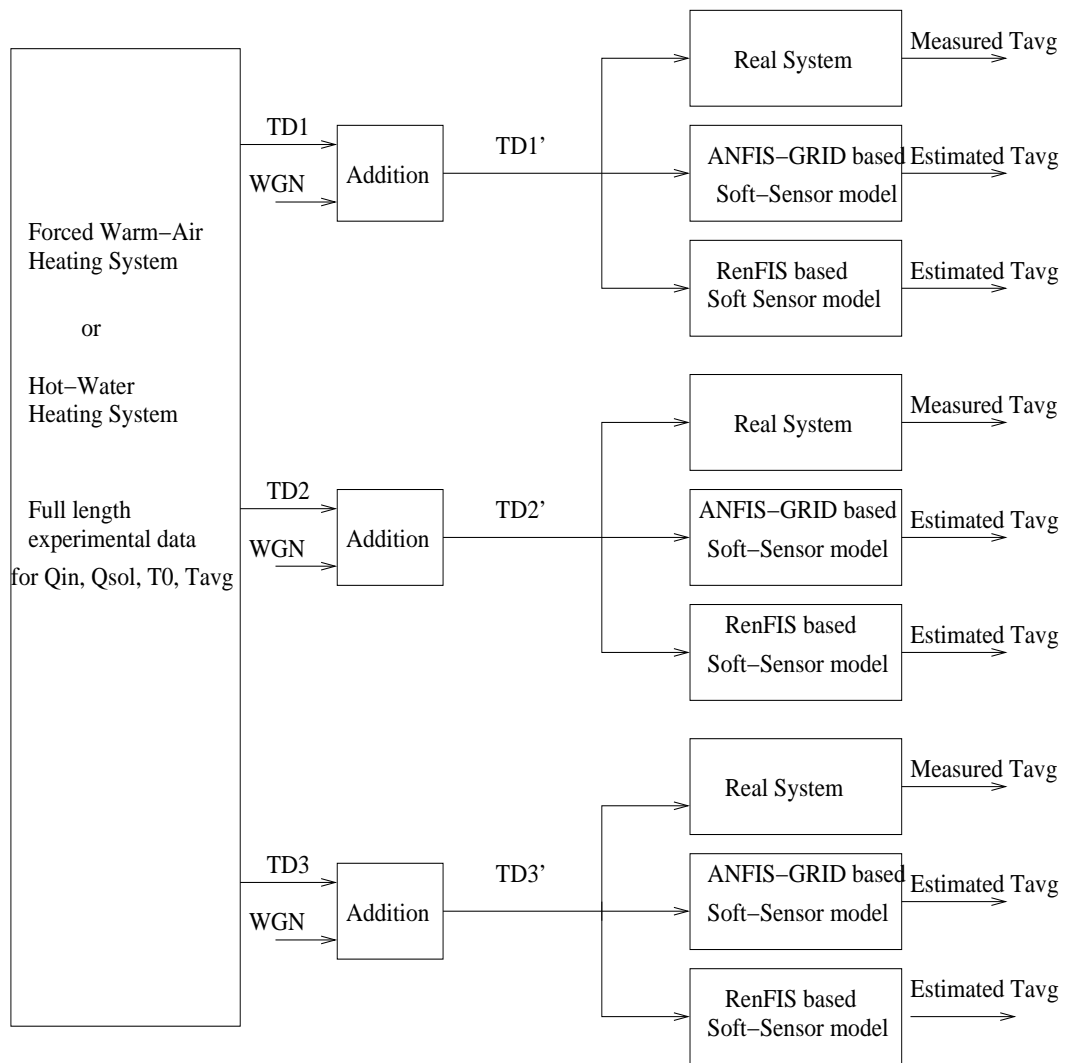


**Figure 4.10:** Comparison of Measured and Estimated Temperatures for the Models Trained with TD2 for Forced Warm-Air Heating System.



**Figure 4.11:** Comparison of Measured and Estimated Temperatures for the Models Trained with TD3 for Forced Warm-Air Heating System.

is conducted by adding noise to the training data sets. As the experimental data is used for the training and testing of the developed models, the data is not accurate due to measurement errors, instrument errors and human errors. The white Gaussian noise (WGN) adds disturbance to the full length data, which can best replace the errors in the experimentally collected data. Disturbance is introduced in the form of white noise to all the three training data sets, which were collected under different conditions. All the input variable signals include the disturbance. Fig. 4.12 shows a schematic diagram for conducting the robustness analysis.

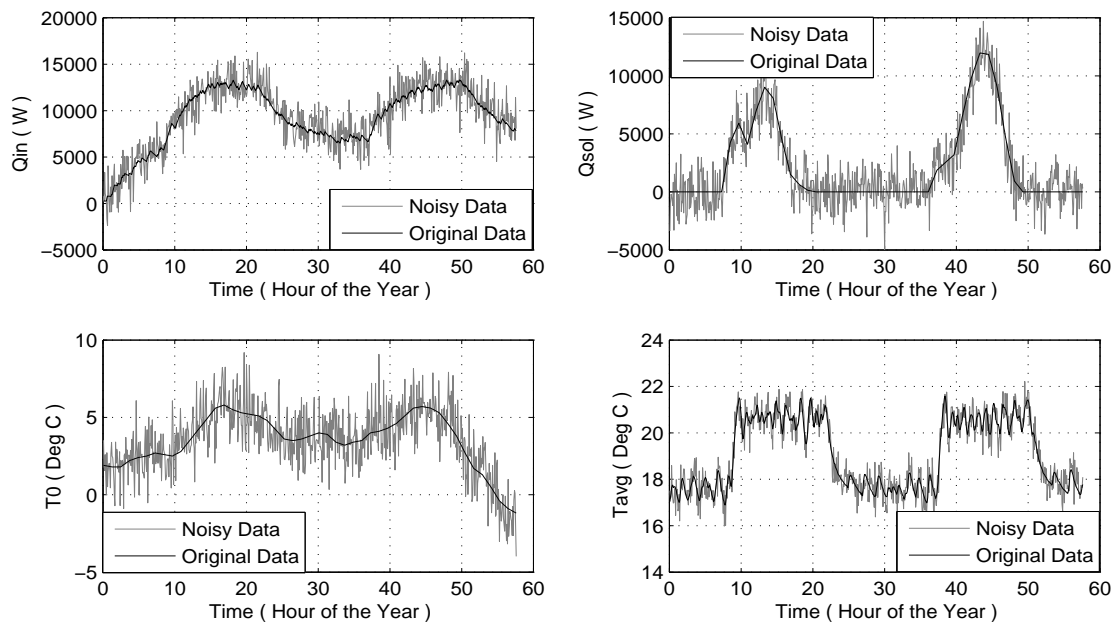


**Figure 4.12:** Schematic Diagram for Robustness Analysis.

In the addition block, noise signal is added to the original training data according to the formula given in equation (4.3).

$$TD' = TD + s \times randn[size(TD)] \quad (4.3)$$

where  $TD$  represents a general training data set.  $s$  is the strength of the noise signal. The value of  $s$  depends on the peak to peak value of the signal. Figs. 4.13 and 4.14 show the data set after adding noise to the data set shown in Figs. 3.3 and 3.4 in Chapter 3.

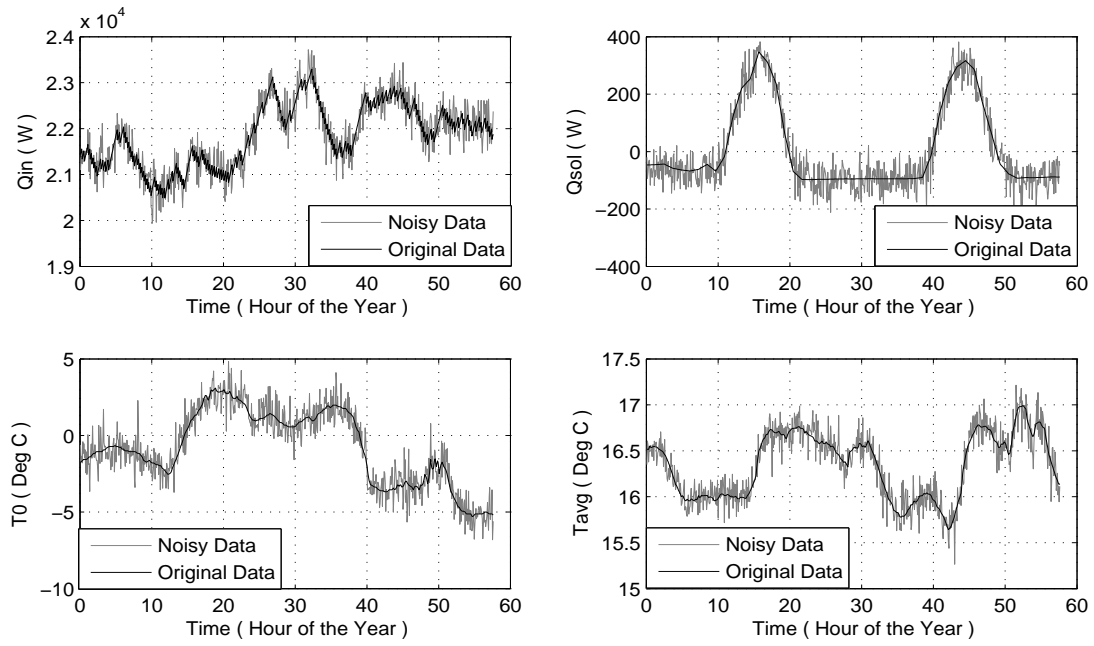


**Figure 4.13:** Noisy Experimental Data for Hot-Water Heating System.

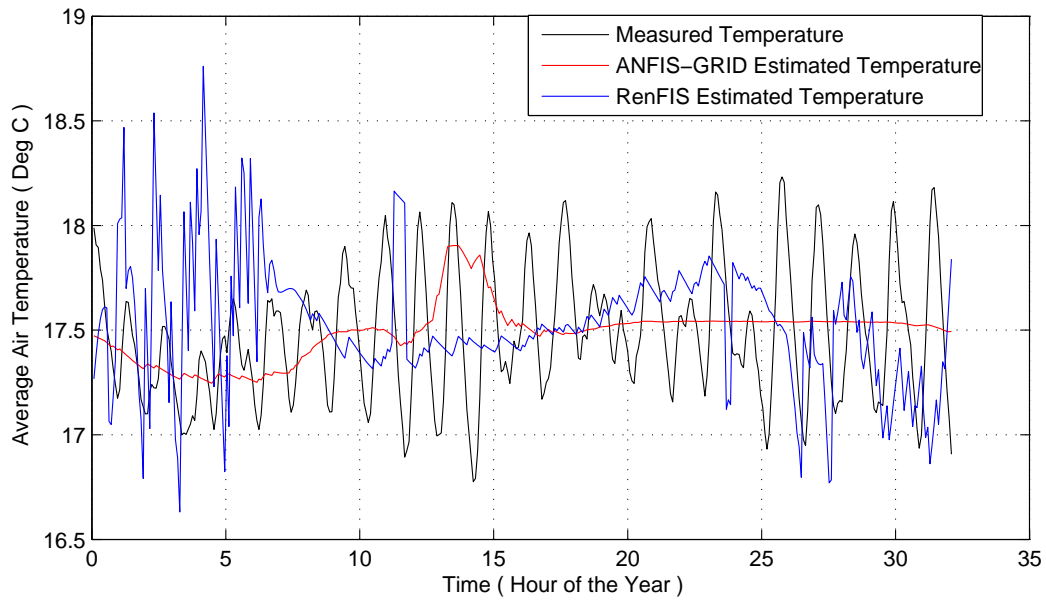
### *Hot-Water Heating System*

The testing results are shown in Figs. 4.15, 4.16, and 4.17. The results show that the noisy signal is reducing the performance of both the models. RenFIS estimation capability compared with ANFIS-GRID has better estimation results because of the presence of the feedback loop.

### *Forced Warm-Air Heating System*

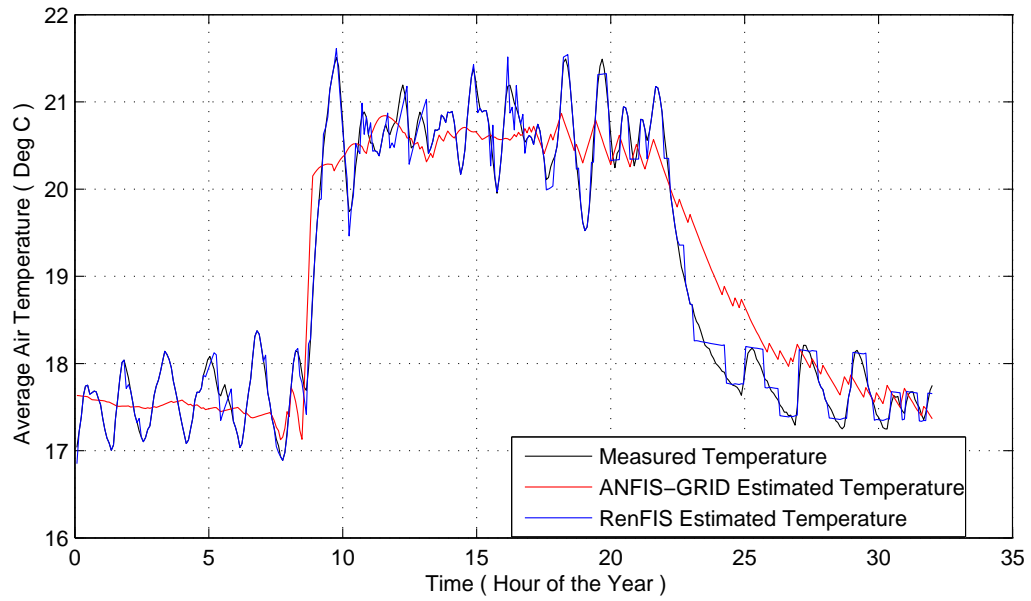


**Figure 4.14:** Noisy Experimental Data for Forced Warm-Air Heating System.

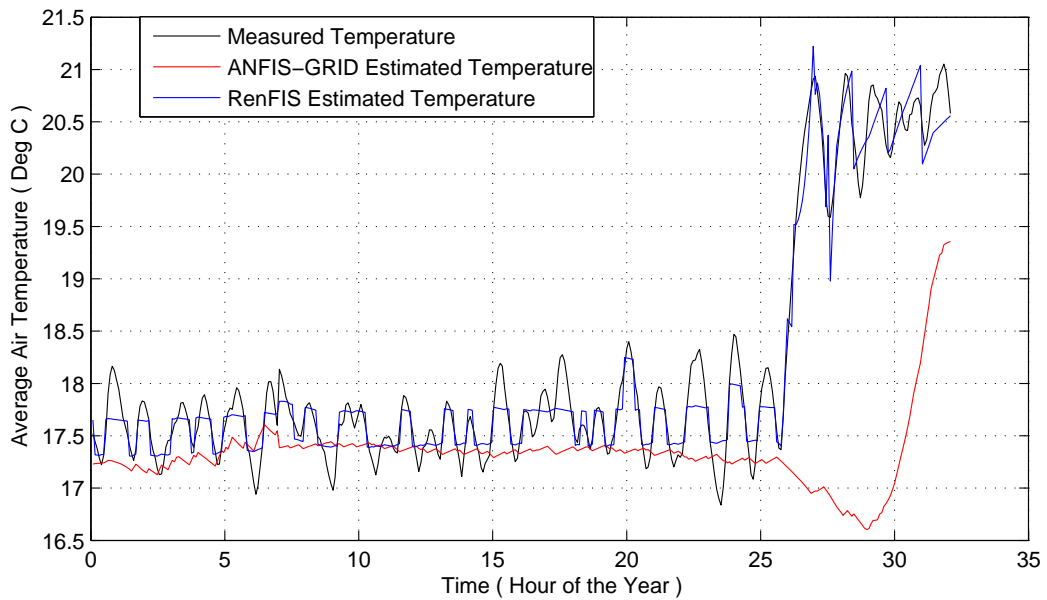


**Figure 4.15:** Comparison of Measured and Estimated Temperatures for the Models Trained with  $TD1$  for Hot-Water Heating System.



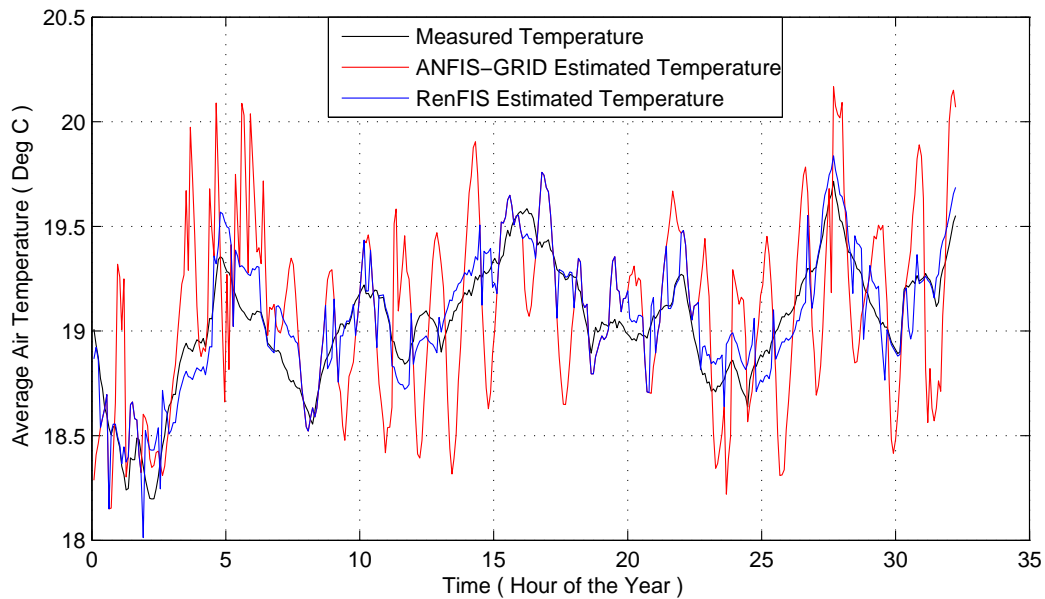


**Figure 4.16:** Comparison of Measured and Estimated Temperatures for the Models Trained with  $TD2$  for Hot-Water Heating System.



**Figure 4.17:** Comparison of Measured and Estimated Temperatures for the Models Trained with  $TD3$  for Hot-Water Heating System.

Figs. 4.18, 4.19, and Fig. 4.20 compare the ANFIS-GRID and RenFIS estimated temperatures values with that of the experimentally measured temperature values for forced warm-air heating system. The results show the same trends as that for hot-water heating system testing results. As a result of analyzing the robustness, good predictive performance was seen for each condition in case of RenFIS.

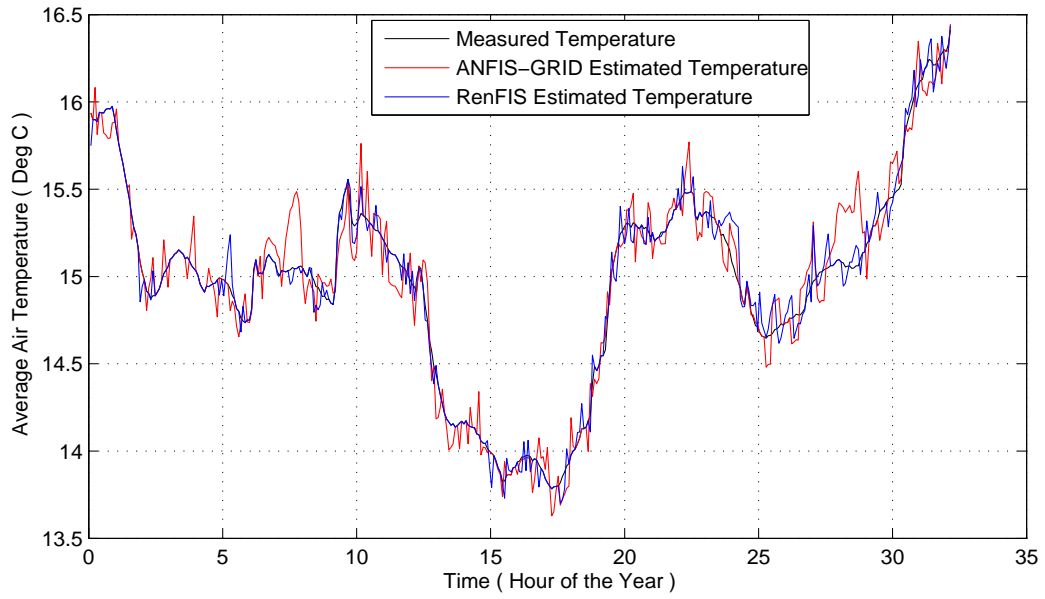


**Figure 4.18:** Comparison of Measured and Estimated Temperatures for the Models Trained with *TD1* for Forced Warm-Air Heating System.

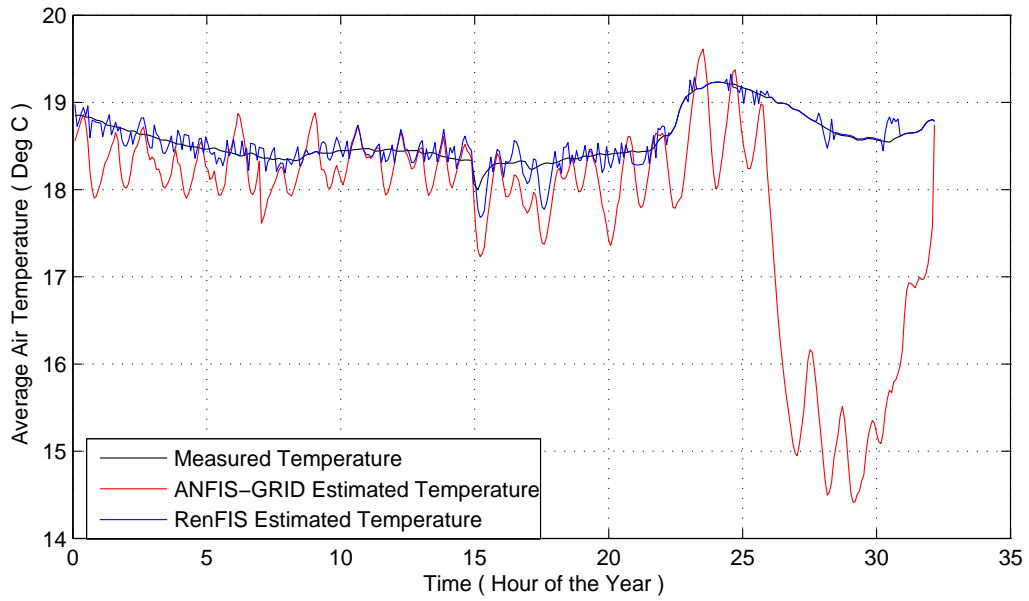
## 4.6 Concluding Remarks

The values for three performance indicators are given in Tables 4.6 and 4.7 for hot-water heating system and forced warm-air heating system, respectively.

ANFIS-GRID has the best performance when trained with training data set *TD2* as RMSE and Bias have minimum and  $R^2$  has maximum value. RenFIS has the best performance when trained with training data set *TD3* as RMSE and Bias have minimum and  $R^2$  has maximum value. Both models have poor performance for training data set *TD1*, as RMSE and Bias have maximum and  $R^2$  has minimum value. RenFIS compared with



**Figure 4.19:** Comparison of Measured and Estimated Temperatures for the Models Trained with  $TD2$  for Forced Warm-Air Heating System.



**Figure 4.20:** Comparison of Measured and Estimated Temperatures for the Models Trained with  $TD3$  for Forced Warm-Air Heating System.

**Table 4.6:** Sensitivity Analysis for Hot-Water Heating System.

Training Data Type	Training Data Length (Days)	RMSE( $^{\circ}C$ )		$R^2$		Bias	
		ANFIS-GRID	RenFIS	ANFIS-GRID	RenFIS	ANFIS-GRID	RenFIS
Weekend Only (TD1)	2.5	3.35	0.25	0.21	0.97	0.29	0.008
Week Day Only (TD2)	2.5	0.87	0.21	0.67	0.98	0.23	0.01
Weekend and Week Day (TD3)	2.5	1.13	0.15	0.25	0.99	-0.23	-0.004

ANFIS-GRID shows better performance in each case when trained with data collected under three different conditions. For TD1, the least representative data set, RenFIS predictive performance is better as RMSE is 92% small, Bias is 97% small and  $R^2$  is 78% higher as compared with ANFIS-GRID. For RenFIS the difference between highest and lowest RMSE values is 40% compared with ANFIS-GRID where such difference is 74%. RenFIS, having an extra input by feeding back the output, compared with ANFIS-GRID is less sensitive to the differences in training data.

**Table 4.7:** Sensitivity Analysis for Forced Warm-Air Heating System.

Training Data Type	Training Data Length (Days)	RMSE( $^{\circ}C$ )		$R^2$		Bias	
		ANFIS-GRID	RenFIS	ANFIS-GRID	RenFIS	ANFIS-GRID	RenFIS
Weekend Only (TD1)	2.5	4.30	0.51	0.19	0.89	0.29	0.006
Week Day Only (TD2)	2.5	1.01	0.45	0.62	0.92	0.23	0.01
Weekend and Week Day (TD3)	2.5	1.19	0.19	0.22	0.97	-0.23	-0.01

The performance of the models trained with noisy data is also checked. The values for three performance indicators are given in Table 4.8 and Table 4.9 for hot-water heating system and forced warm-air heating system, respectively. ANFIS-GRID has the best performance when trained with training data set  $T\acute{D}2$  as RMSE and Bias have minimum and  $R^2$  has maximum value. RenFIS has the best performance when trained with training data set  $T\acute{D}3$  as RMSE and Bias have minimum and  $R^2$  has maximum value. Both the models have

poor performance when trained with training data set  $T\dot{D}1$  as RMSE and Bias have highest and  $R^2$  has lowest value.

**Table 4.8:** Robustness Analysis for Hot-Water Heating System.

Training Data Type	Training Data Length (Days)	RMSE( $^{\circ}C$ )		$R^2$		Bias	
		ANFIS-GRID	RenFIS	ANFIS-GRID	RenFIS	ANFIS-GRID	RenFIS
Weekend Only ( $T\dot{D}1$ )	2.5	10.61	0.52	0.08	0.95	1.92	-0.006
Week Day Only ( $T\dot{D}2$ )	2.5	0.88	0.21	0.69	0.98	0.24	0.01
Weekend and Week Day ( $T\dot{D}3$ )	2.5	2.10	0.16	0.25	0.99	-0.38	-0.004

As a result of analyzing the robustness, good predictive performance is seen for RenFIS when trained training data sets collected under three difference conditions. For  $T\dot{D}3$ , ANFIS-GRID has RMSE increased by 46% and RenFIS has RMSE increased by 6% when trained with noisy data. RenFIS has more robust behavior compared with ANFIS-GRID.

**Table 4.9:** Robustness Analysis for Forced Warm-Air Heating System.

Training Data Type	Training Data Length (Days)	RMSE( $^{\circ}C$ )		$R^2$		Bias	
		ANFIS-GRID	RenFIS	ANFIS-GRID	RenFIS	ANFIS-GRID	RenFIS
Weekend Only ( $T\dot{D}1$ )	2.5	9.46	0.72	0.10	0.87	2.02	0.007
Week Day Only ( $T\dot{D}2$ )	2.5	1.01	0.46	0.60	0.91	0.24	0.01
Weekend and Week Day ( $T\dot{D}3$ )	2.5	2.52	0.21	0.24	0.96	-0.32	-0.01

For sensitivity analysis training data sets are collected under three different conditions. The sensitivity analysis shows RenFIS shows better prediction ability for each condition as compared with ANFIS-GRID. The target application of the average air temperature estimator is in the development of robust control schemes for furnace or boiler operation.

The robustness of the developed models is analyzed by introducing the disturbance to the training data sets. WGN is used as a disturbance signal. This signal adds disturbance to the full length of data, which best replicates the errors in data measurement due to different reasons such as instrument errors and measurement errors. The results show that RenFIS gives better estimated results for the noisy signals due to the presence of feedback loop, which helps to model the dynamic behavior of the processes. ANFIS-GRID shows best estimation of the process if trained with data representing the repetitive quality of the whole universe of the data set, while RenFIS gives better performance as an estimator when trained with the data set representing overall qualities of the whole experimentally collected data.

# Chapter 5

## Control of Residential Forced Warm-Air Heating Systems

### 5.1 Control of Building Heating Systems

Residential heating, ventilating and air-conditioning (HVAC) systems are intended to produce a comfortable indoor environment for the occupants. A typical residential heating system consists of three major components: the generation plant that consumes primary source (e.g. electricity, oil, natural gas) to produce heating/cooling capacity, the distribution system that distributes the heating/cooling capacity to different parts of the building, and the terminal devices that deliver the heating/cooling capacity into the conditioned space. Table 5.1 shows the most common configuration of residential HVAC systems in North America.

Depending on the medium used to distribute the heating/cooling capacity throughout the building, residential heating systems can be categorized into two groups: hot-water and forced warm-air heating systems. In a hot-water heating system, primary energy, normally natural gas or oil, is consumed in a device called boiler to produce hot-water. The heating capacity stored in the hot-water is distributed to terminal devices, which transfer the heating energy from hot-water to the air inside the conditioned space to maintain the desired indoor thermal environment. In a forced warm-air heating system, indoor air is returned to a central equipment called furnace. The furnace consumes primary energy to warm up the return air. The warmed air is then distributed through an air-duct system to different

**Table 5.1:** Residential Heating and Cooling Systems (ASHRAE 2007).

	Central Forced Air	Central Hydronic	Zoned
Most Common Energy Sources	Gas	Gas	Gas
	Oil	Oil	Electricity
	Electricity	Electricity	
Distribution Medium	Air	Water	Air
		Steam	Water
			Refrigerant
Distribution Systems	Ducting	Piping	Ducting
			Piping or Free Delivery
Terminal Devices	Diffusers	Radiators	Included with product
	Registers	Radiant panels	
	Grilles	Fan-coil units	

parts of the building and delivered into individual rooms through terminal devices, normally air-registers.

In North America, the forced air system is most popular. In addition to technical reasons, this can also be justified by economic benefits: one air-duct system can be shared by both heating and cooling.

The basic components of a forced warm-air system include:

- The generation system: a furnace in which primary energy (e.g. natural gas) is consumed to generate heating capacity to be stored in warm-air. The furnace also includes a blower that drives air to circulate between the furnace and the building.
- The distribution system: air-duct system through which air is circulated between the furnace and the building.
- Terminal devices: air registers through which warm air from the furnace is diffused into the conditioned spaces. The amount of warm air is normally controlled manually.
- A control unit through which the users can change the desired room temperature.



## 5.2 Control of Forced Warm-Air Heating Systems

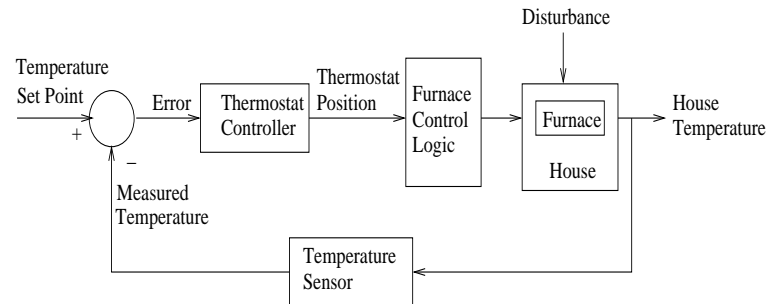
In Canada, forced warm-air ducted systems exist in most of the existing housing stock. Despite the evolution of houses over the last 30 years, forced-warm-air heating systems have changed very little [134].

Section 5.1 has highlighted the basic technology used to heat the residential buildings to yield the acceptable levels of comfort. In addition to the basic components a control scheme is required to deliver the proper amounts of heat.

In addition to the design, construction and commissioning, the performance of a residential heating system largely depends on how it is controlled. Traditionally, the control objective of the air temperature regulation has been to minimize the energy consumption while maintaining temperature within an acceptable comfort margin. This is sensible in commercial buildings, where occupancy patterns follow simple rules and where energy considerations dominate individual preferences. In residential buildings, however, the desires and schedules of occupants need to be weighted equally with energy considerations. In other words, the control of heating system should satisfy the requirement of the thermal comfort and energy efficiency. However, despite the research and implementation of the advanced control schemes, such as PI, PID, fuzzy, adaptive, feed forward etc., most of them are on the commercial buildings, laboratory or hospital.

Most residential buildings use thermostat based control which is simple but gives unsatisfactory performance. In this control method the occupant provides the thermostat with their preferred temperature set-points and the device performs all the tasks required to maintain the temperature the set-point temperature. This includes observing how the system is behaving, comparing that to the reference given by the occupant, determining the actions needed, and applying the input. The temperature observation is usually made by only one, local sensor. Depending on the energy system the thermostat is controlling the heating system, the thermostat will decide an action for the system based on the measured temperature being lower or higher than the set-point. If the system only has two modes of operation, ON and OFF, it is classified as a “two-position” system. The thermostat will trigger the heating

system to turn on if the sensor temperature is measured to be lower than the set-point. Fig. 5.1 visualizes this controlling scheme in a block diagram.

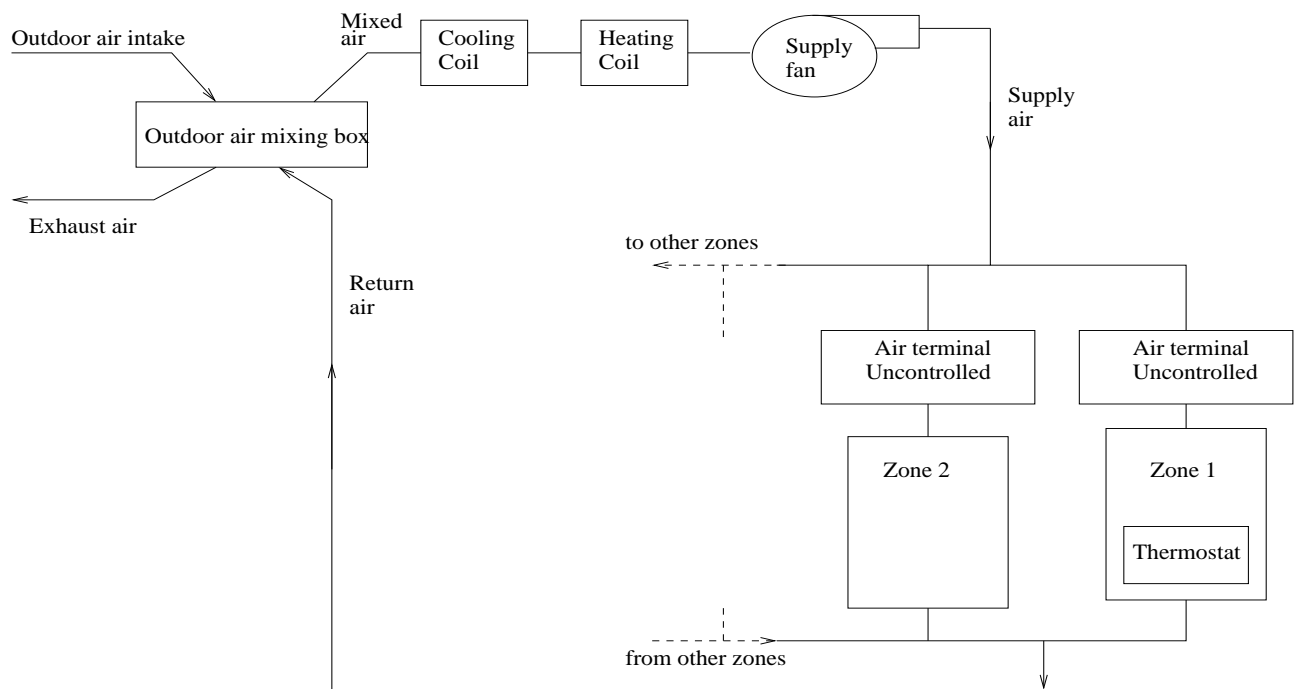


**Figure 5.1:** Block Diagram for Thermostat Controlled Heating System.

For the heating systems controlled by a single, centrally located thermostat, the temperatures in other rooms of the house are not actively controlled, so the temperatures in different rooms can differ significantly from the thermostat. As shown in Fig. 5.2, the thermostat senses the temperature in Zone 1 of the whole building. The building environment is controlled, only based on the reading of the sensor located in Zone 1. Manually adjusting the airflow to each room is the primary method available to control the temperature in the rooms without thermostat. The actual temperature depends on many factors such as the season, the outside temperature, radiation heating through windows, and the activities of people and equipment in the rooms. The desired temperature also depends on the activity of the occupant, for example lower temperatures for sleeping and higher temperatures for relaxing. Maintaining comfortable temperatures requires constant adjustment, or may not be possible.

The energy cost of heating system in a residential building is an important issue. To get a more energy efficient system, other control system shall be considered.

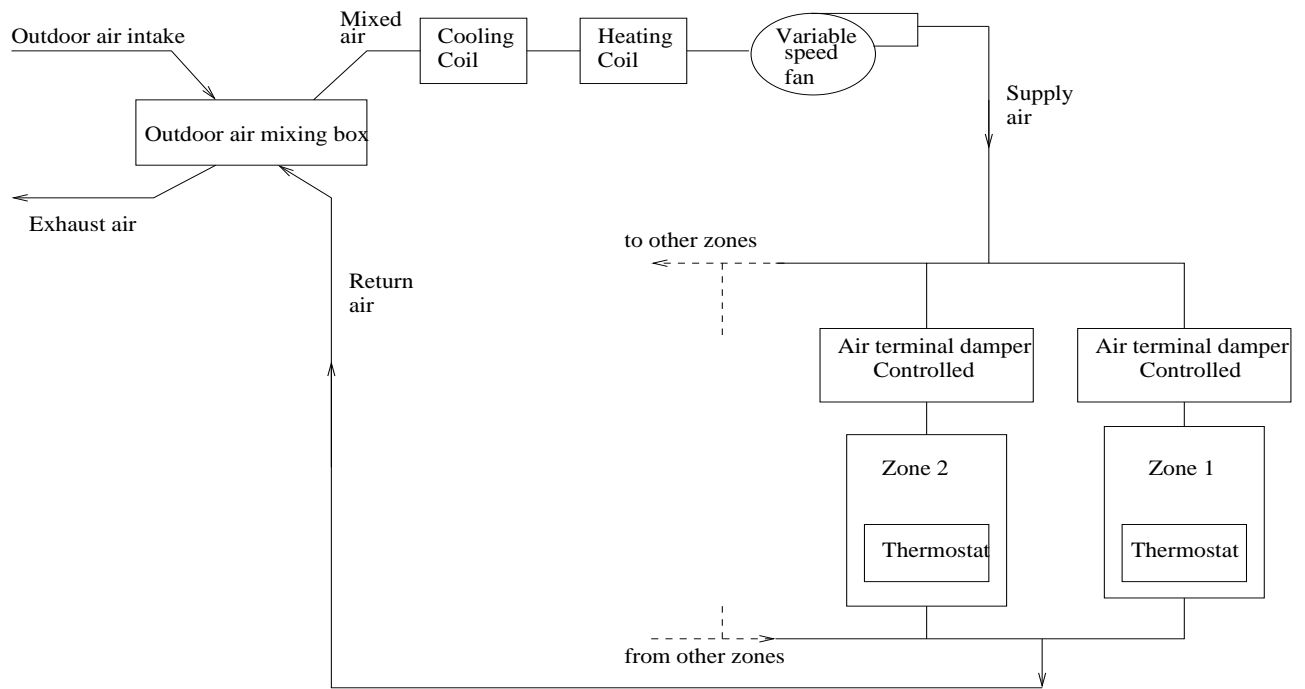
Zoned control systems have been developed to improve temperature control. As shown in Fig. 5.3, typically, a number of thermostats are located in different areas of the building, and a similar number of mechanized airflow dampers are placed in the air distribution ducts. A control unit dynamically controls the HVAC equipment such as the operation of the furnace and the airflow to simultaneously control the temperatures in each zone. These conventional



**Figure 5.2:** Single Zone Control System.

systems are difficult to retrofit as compared to the resulted benefits.

With only a few zones, there can still be significant temperature variations from room to room within a zone. A few systems have proposed thermostats for each room and airflow control devices for each air vent, but no practical solution for easy retrofit has been disclosed. As the number of independent zones increases it becomes more complex to specify appropriate setting for each zone while providing convenient centralized and remote control. Typical residential HVAC systems are designed to produce one fixed rate of heating and cooling. In residential buildings with only few zones, the zoned control methods have not been widely adopted because they are expensive, difficult and intrusive to install in most existing houses, and provide limited utility and benefit compared to their cost and inconvenience.



**Figure 5.3:** Zoned Control System.

## 5.3 Improving Energy Efficiency and Thermal Comfort

### 5.3.1 Energy Efficiency and Thermal Comfort: High Efficiency Heating Systems

Reducing the amount of energy of a heating system benefits both the occupants of the residence (lower energy bills) and energy providers (lower peak demands). Meyers *et al.* [135] reviewed modern energy conservation techniques and believe the most amount of energy can be conserved by preventing heating and cooling unoccupied spaces.

Smaller temperature differences between indoor and outdoor environments results in lower heat loss (winter) or heat gain (summer), lowering the thermostat set-point during heating season and raising it during cooling season reduce energy demand. Lowering the set-points has energy benefits, but may also result in lower level of comfort. Research has been conducted in a wide variety of climates throughout the world and over several decades to determine settings that optimize comfort and energy conservation. Nassif *et al.* [136] used a genetic algorithm search to optimize the monthly set-points for a building on the campus

of Ecole de technologie superieure in Montreal, Canada. An evolutionary programming technique was used by Fong *et al.* [137] to find the set-points for different weather conditions throughout a year in China.

Further reduction in energy consumption can be achieved by setting the comfort level in the house significantly lower during the nighttime and periods when the house is unoccupied. This practice is known as thermostat setback, and has been around for many years [138]. Programmable thermostats were designed to reap the benefits of thermostat setback while preventing daily periods of discomfort by scheduling the setback periods to end before the occupants would wake-up in the morning and return from work/school and extended periods away from the house. A large percentage of programmable thermostats are not used properly because of irregular occupant schedules and difficulty with the programming interface. Recent work at the University of Virginia with self-programming thermostats has yielded positive results [139]. These thermostats observe the occupancy patterns in a home and automatically optimize the heating and cooling schedule based on the statistical occupancy. The occupant has an opportunity to define the desired balance between energy and comfort using a single, intuitive knob. In the most cost efficient mode, the test researchers reported 15% energy savings on top of the Energy Star recommended setback schedule.

Other energy conservation methods include the use of high efficiency furnaces. The performance of a furnace is usually assessed by the following indices:

- Thermal Efficiency

$$\text{Thermal Efficiency} = \frac{\text{Fuel Energy Output}}{\text{Fuel Energy Input}} \quad (5.1)$$

- Annual Fuel Utilization Efficiency (AFUE)

AFUE measures the annual amount of heat actually delivered compared to the amount of fuel supplied to the furnace.

$$AFUE = \frac{\text{Annual Fuel Energy Output}}{\text{Annual Fuel Energy Input}} \quad (5.2)$$

- Steady State Efficiency (SSE)

SSE measures how efficiently a furnace converts fuel to heat, once the furnace has warmed up and is running steadily, and can be expressed as:

$$SSE = \frac{\text{Fuel Energy Input} - \text{Fuel Energy Loss}}{\text{Fuel Energy Input}} \quad (5.3)$$

*Condensing Furnace:* Because of the need to conserve energy and the escalating cost of the fuel, the heating industry developed a new generation of gas-fired furnaces, the condensing furnaces. These furnaces can achieve an overall operating efficiency of 90 to 97 percent. In a condensing furnace, an enlarged heat exchanger surface lowers the temperature of the exhaust gases, making the furnace more efficient. The exhaust gas temperature drops to the dew point of the water vapor in the gas, causing the vapor to condense to water and give up 2,255 KJ of energy for every kilogram of water condensed. Natural gas can yield more than 3.7 kgs of water per 105 MJ burned, giving up about 8,187 KJ of energy. Condensing furnaces achieve the AFUE of up to 94%.

*Pulse-combustion Furnace:* Another type of high-efficiency furnace is the pulse-combustion furnace. Whilst the heat in a condensing furnace results from a continuous burning of fuel, in a pulse combustion furnace, it results from 60 to 70 tiny explosions of a gas-air mixture per second in the combustion chamber [140]. These are expensive and hard to install as compared to the standard furnaces.

### 5.3.2 Energy Efficiency and Thermal Comfort: Control of Heating Systems

Improving the performance of individual components in the heating system can improve energy efficiency by requiring less electricity and running in shorter intervals. For central forced warm-air heating systems, energy can be saved by using a more efficient blower. Studies conducted at Lawrence Berkeley National Laboratory have investigated several high efficiency motors used for blowers [141, 142]. Two prominent technologies are permanent split capacitor (PSC) single-phase induction motor and brushless permanent magnet (BPM)

motor. Under field conditions, natural gas furnaces fitted with PSC motor driven blowers consumed 10% less electricity than the Department of Energy (DOE) test procedure results; BPM consumed 36% less electricity. Note that these percentages are only the electricity savings for the fan motors, actual energy savings would be less.

Multi-stage furnaces are used to improve the efficiency of the heating systems. A specific load can be met with the most efficiency by a particular capacity. The multi-stage systems are able to deliver different levels of capacity at the different stages. Lower load might only require a little bit of heat, so the low setting on a multi-stage furnace might only combust the natural gas at a fraction of what it is capable of. In 2006 both DOE and ASHRAE performed studies on residential two-stage furnaces with traditional fan motor technology; interestingly, each test arrived at a different conclusion. The Lawrence Berkeley National Laboratory compared the two tests and results reported by other field tests [143]. When combining the reduction in fuel with the increase in electricity consumption in two-stage furnaces, the DOE test yielded a 3% reduction in energy. The ASHRAE test showed almost no difference in the energy consumption compared to a single-stage system in the same efficiency class.

## **Multiple Sensor Control**

Another source of energy loss in residential heating systems is related to large temperature differences between the rooms. Individual rooms or area in a house can be several degrees warmer than the remainder of the house. Occupants can experience large periods of discomfort and again try to compensate by treating the thermostat like an ON-OFF switch by manually adjusting the set-point to a temperature higher than the comfort level desired. The system then heats constantly until the occupant achieves localized comfort; consequently, other areas of the house become uncomfortable and an unstable cycle of overshooting the set-points ensues. A solution that has been explored is to have multiple sensors distributed throughout the house so the controller will know about any large temperature differences. The ASHRAE Handbook recommends the use of multi-zone control when a sin-

gle thermostat is not able to properly characterize the house due to zone-to-zone temperature differences [144].

Simulations have been conducted that yield energy savings in central two-position systems with multiple sensors. Lin *et al.* [145] presented the multi-sensor, single-actuator control method. Ad hoc methods are used for averaging the information obtained from multiple sensors. The performance improvements occur primarily when the average load among all rooms is nearly zero, with some rooms requiring heating while others require cooling. Under these conditions, the single-sensor strategy either overcools or overheats, whereas the multi-sensor strategies use almost no energy. Though rare for residential use due to the price, multi-zoning and using a variable air volume (VAV) system to deliver variable amounts of air can save energy by only providing heating to the zones that need it. Oppenheim [146] developed several multi-zone setback strategies and compared them to a single zone strategy with an 8-hour  $-11^{\circ}\text{C}$  setback. He was able to see fuel savings of 12% with a multi-zone strategy using 22-hour setbacks in two zones, but observed 6% additional fuel consumption when the bedroom zone was set back 18 hours and the rest of the house only 12 hours. The study does not consider the effect of outdoor conditions in the analysis of comfort. The analysis could show the sources for the unexpected results of increasing the fuel consumption. Temple also showed that the energy consumption increased slightly in a home installed with a zoned system because rooms that were originally uncomfortable from not receiving enough cooling energy, were provided with it. This cost the system more energy but provided more overall and uniform comfort [147].

Constant air volume (CAV) systems can be operated as zoned systems by opening and closing the registers in different parts of the building. Manually adjusting the registers would be a demanding task if variable loads and demands were always being imposed on the building. The demand response enabling technology (DRETD) group at the University of California, Berkeley has performed some preliminary work in the field of automating dampers to multi-zone houses [148]. The research project divided a two-story home into four zones, and used an occupancy schedule to dictate the set-points for individual zones



during the heating season. The original thermostat was located in the zone that would always report the lowest temperatures, so without the multi-zone system this house experiences over heating in every other zone. This observation and the occupancy schedule are two important details that can explain the high 26% energy savings. Another factor that register automation systems need to be cautious of is restricting too much airflow. Walker determined that closing registers translates to increased energy usage by the fan. He concluded that closing 60% of the registers could lead to frozen evaporator coils and other pressure related issues [149].

Many research articles in the ASHRAE Journals discussed using wireless technology to develop sensor networks for building monitoring and control applications. Wireless sensor networks are becoming easier to install and operate while decreasing in price and operating power consumption [150,151]. Cost effectiveness in retrofit and new construction applications and the reliability of wireless sensor networks are still the issues to resolve [152].

Table 5.2 summarizes the latest energy efficient technologies.

**Table 5.2:** Residential Heating and Cooling Energy Efficient Techniques.

Energy Efficiency Technique	Investigative Authors	Benefits and Energy Savings
Set-point education	Vine 1986 [153]	Lower energy consumption
Thermostat setback	Ingersoll 1985 [138]	22% energy savings
Self-programmable thermostat with setback	Gao 2009 [139]	15% energy savings
Improved fan motor	Walker and Lurtz 2005 [141] Lurtz <i>et al.</i> 2006 [142]	10-36% less electricity
Ground source heat pump	Sanner <i>et al.</i> 2003 [154] Lund <i>et al.</i> 2004 [155] Omer 2008 [156]	Energy savings CO2 emissions
Evaporative cooled condensor	Hwang <i>et al.</i> 2001 [157]	Increased capacity
PID thermostat control	Kolokotsa <i>et al.</i> 2006 [158]	20% energy savings
2-Stage furnace	DOE 2003 ASHRAE 2003	3% energy savings No significant savings
Multi-zone sensing	Lin <i>et al.</i> 2002 [145]	17% energy savings
Multi-zone with VAV	Oppenheim 2002 [146] Temple 2004 [147]	12% energy savings Increased energy consumption Improved thermal comfort
Distributed wireless sensing and control algorithms	Ota <i>et al.</i> 2008 [159]	79% energy savings and thermal comfort improvements

## 5.4 Survey

Obtaining feedback from occupants is imperative in help identifying potential deficiencies in existing building systems. To identify potential energy savings and improvement to thermal comfort in residential buildings, the present situation is investigated first. The investigation is carried out in two parts, an on-site interview and a full scale survey.

As mentioned in Chapter 1, a survey is conducted by the author. For the first part, 15 respondents are interviewed at their dwellings in December, 2008. The objective of the interviews was to gather information on mechanical equipment specifications and to collect necessary information to help create a short-listed set of questions to be included in the full scale survey. Through the interviews, problems were identified with the lack of controllability and the occurrences of overheating in winter heating months. In order to generate a more holistic view on the problem and to provide a statistically supported analysis, it is necessary to conduct a full scale survey to evaluate the extent of the problems.

Following the interviews, a full-scale survey is conducted with Ontarians residing in low-rise residential buildings and a provincially representative sample was gathered. The questionnaire for full-scale survey is included in Appendix D. The main objectives of the full scale survey are as follows:

- To identify the potential deficiencies in the existing building stock by investigating the thermal comfort conditions of occupants.
- To analyze factors that impact on the thermal comfort levels.
- To investigate occupants' energy consumption behaviors specifically related to the control of their thermal environment.

### 5.4.1 Sample Population Distribution

The survey is planned to target low-rise residential buildings that would represent typical types and demographics of the Ontario population. A total of 519 surveys are collected from

April to June 2009 and 396 of which are completed and provided valid data. Only those that are completed are included as part of the statistical analysis. Summaries of the respondents are given in Table 5.3. It shows that the sample population distribution corresponds fairly well with Ontario population. Out of the 396 respondents, 67.4% resides in single-detached houses and 23.7% and 8.9% of which resides in double/row houses and low-rise apartment Buildings, respectively. This is comparable to the provincial distribution, according to the 2006 census, 66.9% of Ontarians reside in single-detached houses, 20.3% and 12.9% of which reside in double/row houses and low-rise apartment buildings, respectively. The median age of Ontario's population is 39, when compared to 33 for the survey population. The percentage of surveyed dwellings built before the year of 1986 was 61.2% and built after the year of 1986 was 38.8%. While 68.6% of Ontario dwellings were built before the year of 1986 and 31.4% were built after that year. The percentage of occupants owning the property is 81.2% and renting the property is 18.8%. Distributions between males and females have also been respected.

**Table 5.3:** Summary of Survey Sample.

			Dwelling Type		
			Single-Detached	Double/Row Houses	Low-Rise Apartment Buildings
Gender (%)	Male	50.6	52.4	48.4	44.1
	Female	49.4	47.6	51.6	55.9
Ownership (%)	Own	81.2	88.7	75.0	45.5
	Rent	18.8	11.3	25.0	54.5
Age (yr)	Mean	35.5	36.4	33.5	33.8
	SD	12.5	13.0	11.1	11.7
	Minimum	16.0	18.0	16.0	19.0
	Maximum	68.0	68.0	66.0	59.0
Year of Construction(yr)	Mean	1971	1971	1969	1979
	SD	31	30	34	33
	Minimum	1856	1856	1865	1865
	Maximum	2009	2009	2007	2008

## 5.4.2 Survey Findings and Analysis

As mentioned earlier in Chapter 1, ASHRAE Standard 55-2004 [160] specifies that for 80% of the time the building occupants are expected to express satisfaction over their thermal environment. Although the subjected building types are not expected to comply with the above comfort standard, the 80% acceptability threshold can act as a reference to this study. Results show 6.4% and 13.3% of the respondents are very dissatisfied and dissatisfied respectively with their room temperature in the winter. Satisfaction with room temperature was rated with a scale from 1 to 5 with 1 being very dissatisfied and 5 being very satisfied and results reveal that 5.9% and 14.5% of the female respondents and 6.3% and 12.0% of the male respondents feel very dissatisfied and dissatisfied, as shown in Table 5.4.

**Table 5.4:** Satisfaction with Room Temperature for Male and Female Respondents.

Satisfaction Level	Gender	
	Male	Female
Very dissatisfied	6.3%	5.9%
Dissatisfied	12.0%	14.5%
Neutral	31.8%	31.3%
Satisfied	42.6%	40.5%
Very satisfied	7.3%	7.8%

The method used to investigate thermal comfort acceptability in the present research is based on the maintenance of the temperature at the desired set-point. Yang and Zhang used three methods to investigate thermal acceptability and yields distinctive results [161]. Although it can be argued that employing different methods of assessing thermal comfort levels would produce diverse results, the method used in this study is chosen due to its simplicity.

It is believed that certain factors within the heating system would influence occupant's thermal sensation for example the number of programmable thermostats was selected to be evaluated with occupant thermal sensations: satisfaction level, controllability, and overheating and overcooling frequencies.

Most people have a programmable thermostat for temperature control in their dwelling.

Only 15.5% of the respondents stated they do not have a programmable thermostat at home. There are differences in the presence of programmable thermostats in the different house types (Table 5.5), for which 88.0% and 86.2% of the single-detached and double/row houses have programmable thermostats and only 55.1% of the low-rise apartment buildings have this equipment.

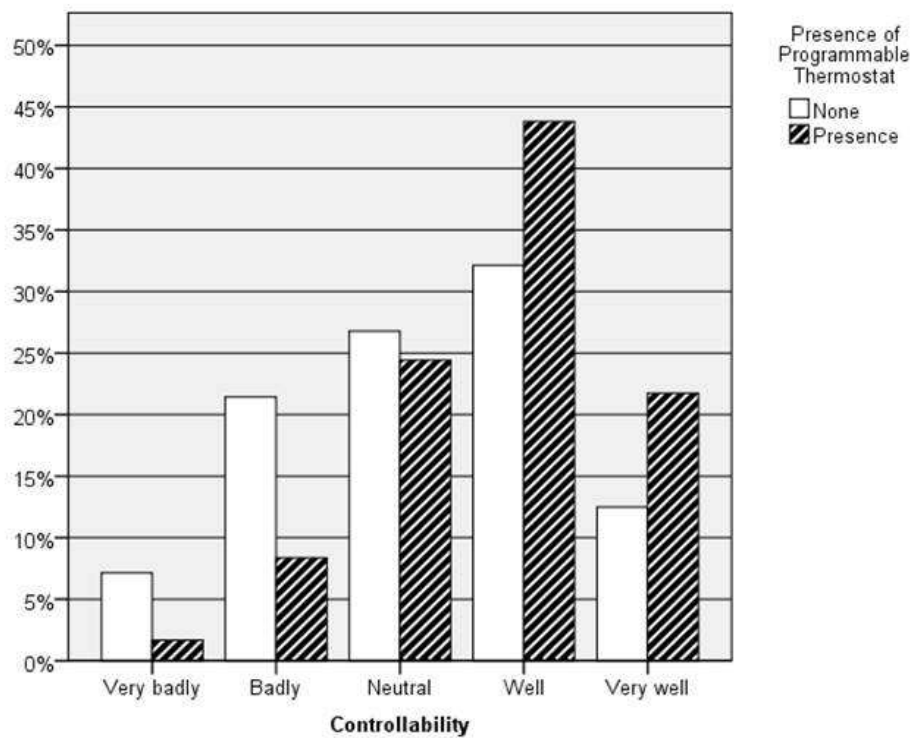
**Table 5.5:** Percentage of Different Types of Houses Equipped with Programmable Thermostats.

Presence of the Programmable Thermostat	Dwelling Type		
	Single-Detached	Double/Row Houses	Low-Rise Apartment Buildings
None	12.0%	13.8%	44.9%
Presence	88.0%	86.2%	55.1%

The respondents were asked to rate their controllability over room temperature in winter with a scale from 1 to 5 with 1 being very badly and 5 being very well. Controllability is the extent to which respondents feel they can control their room temperature. There exists a difference between male and female respondents, 12.0% and 14.2% of the female and male respondents rated their controllability as badly or very badly. Results also reveal that 12.1%, 11.3%, and 26.5% of the respondents rate their controllability over room temperature as badly or very badly in single-detached, double/row houses, and low-rise apartment buildings, respectively. The level of controllability is much lower in houses without programmable thermostats as shown in Fig. 5.4.

Respondents tend to have higher controllability over their room temperature if they have access to one or more programmable thermostats in the dwelling. Most people have a programmable thermostat for temperature control in their dwelling. Only 15.5% of the respondents stated they do not have a programmable thermostat at home. There are differences in the presence of programmable thermostats in the different house types, for which 88.0% and 86.2% of the single-detached and double/row houses have programmable thermostats and only 55.1% of the low-rise apartment buildings have this equipment. However, the level of satisfaction does not seem to be affected by the presence of programmable thermostats, as shown in Fig. 5.5.

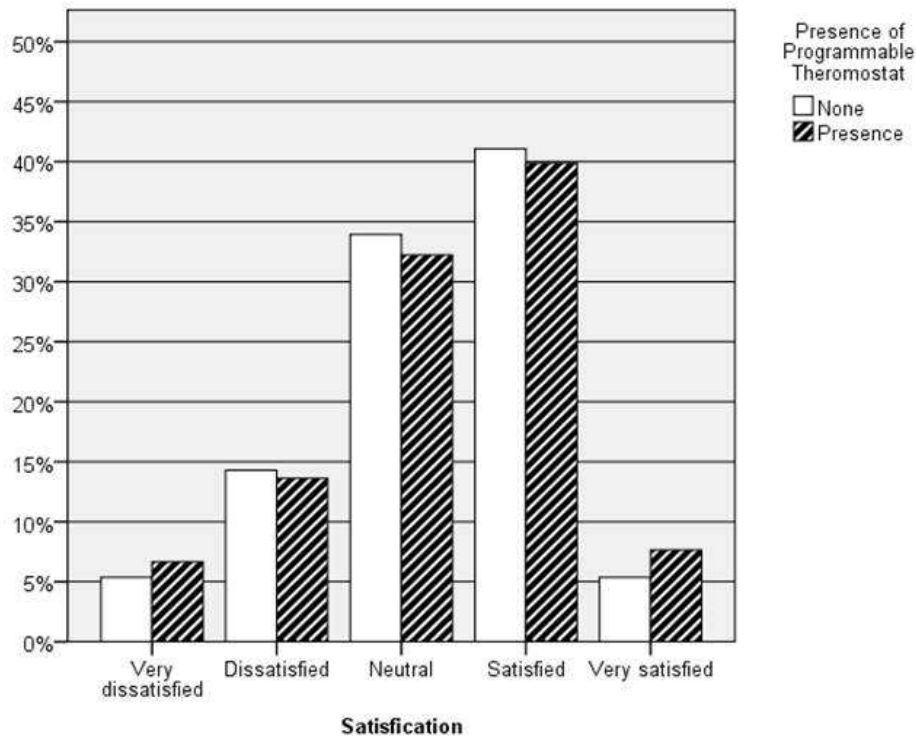
So as per the survey results the occupants feel more dissatisfied with the indoor environ-



**Figure 5.4:** Controllability Over Room Temperature in the Dwellings with and without Programmable Thermostats.

ment if their controllability over the heating systems is low and the presence of thermostats increases the level of controllability but does not improve the thermal comfort level. Survey results also indicate that only 41% and 52% of the respondents set temperature back during the day when the dwelling is unoccupied and during the night, respectively.

The presence of programmable thermostats does not encourage respondents in adjusting set-point temperature. This suggests that relying on the occupants' voluntary action in configuring the thermostat is problematic and therefore calls for further actions, such as introducing educational programs or incorporating more advanced technologies. The potential of existing equipment or technology in reducing energy consumption is hindered when they are not used in the way they were planned. Educating occupants in properly operating equipments already installed in their dwellings provide a big potential in reducing energy consumption without any capital investments and therefore should be considered and pur-



**Figure 5.5:** Satisfaction Rating in the Dwellings with and without Programmable Thermostats.

sued. Basically there is a need for a certain control algorithm which can control the operation of the heating equipment based on the estimation of the temperature representative for the whole building.

## 5.5 Inferential Sensor Technology

Retrofitting existing and new construction buildings with zoned control is expensive and invasive. This dissertation is intended to develop an inexpensive and efficient control technology for residential forced warm-air heating systems.

Inferential control technology can significantly improve control of the building automation systems. Recent research shows that control schemes based on inferential sensing techniques can significantly improve the overall performance of the built environment [9, 10, 72]. Liao has developed an inferential control scheme for optimal control of the boilers in multi-zone heating systems. The results show that, using an inferential sensor, the average room tem-

perature in a heating system can be estimated with satisfactory long-term accuracy according to the information available to the conventional boiler controllers. The inferential control scheme based on this inferential sensor can significantly improve the degree of thermal comfort and reduce fuel consumption by up to 20% in heating systems with poorly controlled radiators and 5% to 11% in heating systems with well controlled radiators. Wang developed a model-based control scheme to optimize the operation of an air-conditioning system. His simulation results show significant improvement on the overall control performance.

However, these inferential sensing and control schemes are very specific to certain types of building systems or devices. The topology of these schemes is difficult to understand and therefore their application in practice is considerably limited. Furthermore, the methods to commission these schemes are not well developed. As a result, it is difficult to implement these schemes in real building automation systems. This research employs neuro-fuzzy modelling methodology for the development of the inferential model to be used in the inferential control scheme. The training algorithm is also discussed in Chapter 3. Chapter 3 and Chapter 4 presented the development of the inferential model and data quality issues related to the performance of the inferential model. The results show that the model is structured and trained correctly and can be used for the development of the inferential control scheme.

## 5.6 Concluding Remarks

The performance of a residential heating system largely depends on how it is controlled. Traditionally, the control objective of the air temperature regulation has been to minimize the energy consumption while maintaining temperature within an acceptable comfort margin. Most residential buildings use “two-position” control which is simple but gives unsatisfactory performance. Zoned control systems have been developed to improve overall performance of buildings. These systems are difficult to retrofit as compared to the resulted benefits. The analysis shows that inferential control technology is an inexpensive and simple method to improve energy efficiency and thermal comfort for both the residential and commercial buildings.



# Chapter 6

## Advanced Control Scheme for Building Heating Systems

### 6.1 Introduction

The heating and cooling loads of the houses built today are generally overestimated, which is partly due to the common practice of utilizing single-zone CAV HVAC systems for conditioning the entire house [162]. This typically results in inefficient heating or cooling due to the differences in thermal load in different spaces throughout the house, and in turn reduces the thermal comfort of the occupants as well as wasting energy used for over-conditioning these spaces [163].

The ASHRAE Handbook of Systems and Equipment recommends the use of multi-zone control when a single thermostat is not representative of the temperature in all rooms [160]. Technologies do exist to provide multi-zone control in residential forced-warm-air heating systems. For instance, first is VAV delivery in conjunction with dampers in individual supply ducts is being used in new construction. However, to retrofit an existing single zone system into a multi-zone one would be expensive and invasive. While many users would prefer the multi-zone system over a single zone one, most would not opt to perform the extensive and invasive retrofit.

Second is a CAV system with a bypass loop. It has been tested and the research found that “the bypass damper reduced the system capacity as desired but the power consumption

remained relatively constant resulting in a reduction in efficiency with increasing amount of bypassed air” [147]. Therefore, such a system can provide zoned control, but it will not save energy.

A third multi-zone option is to use separate heating/cooling units for each zone. But, this method is expensive and the improvement in comfort level is often not worth the added cost to the homeowners.

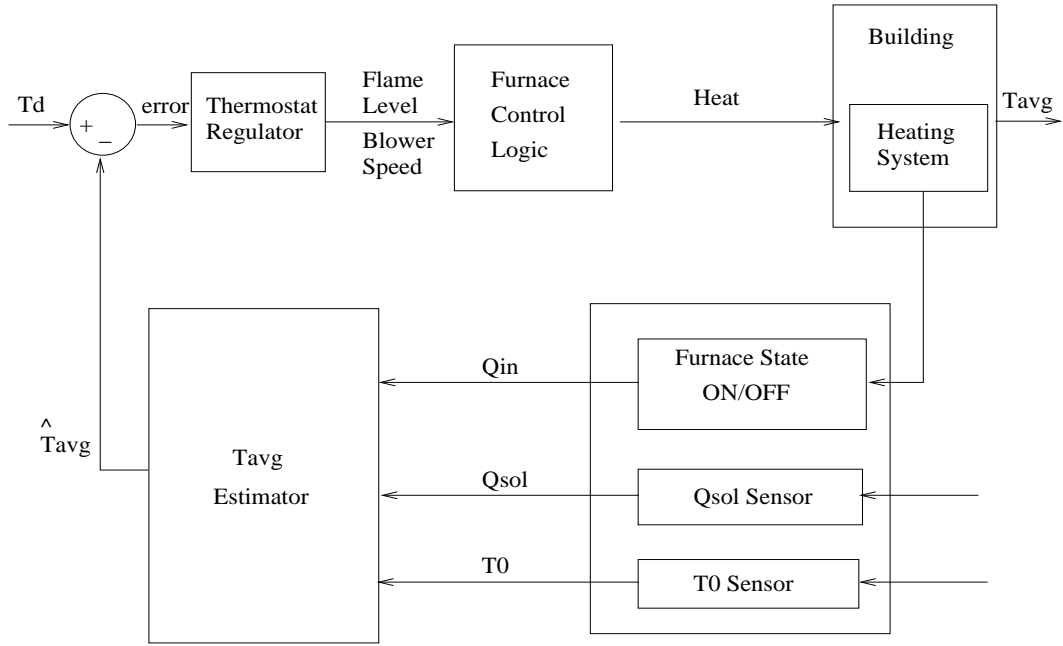
A cost effective control scheme is proposed in this research. In order to justify the need for this control method a survey is conducted. Survey findings presented in Chapter 5 shows a correlation between the occupants controllability over heating systems and the thermal comfort satisfaction level.

## 6.2 The Inferential Control Scheme

An average air temperature based control scheme has been described in Chapter 1. Compared with conventional control schemes, the inferential control scheme has a potential to save energy and improve thermal comfort by controlling the heating equipment according to the average air temperature in the building. In Chapter 3, an inferential model has been developed to estimate the average air temperature in the building based on the information available to the controller. Fig. 6.1 shows how the inferential model can be used in the development of an inferential control scheme.

As shown in Fig. 6.1, the inferential control scheme consists of the following major components:

- An inferential model, which is used to estimate the average air temperature in the building based on three inputs ( $Q_{in}$ ,  $Q_{sol}$ , and  $T_0$ ).
- An air temperature reset strategy, which determines the set-point of the supply air temperature according to the difference between the estimated and desired room temperature.



**Figure 6.1:** Furnace Control Scheme using Average Air Temperature Estimator.

- The furnace firing logic compares the supply warm-air temperature with the set-point to determine the status of the furnace. This is an ON-OFF or PID controller.

## 6.3 The Performance of the Inferential Control Scheme

### 6.3.1 The Simulated Building

The performance of the inferential control scheme has been investigated using a typical residential building simulator.

The building is an existing single-detached house located in Markham, Ontario. The building consists of three levels with two levels above ground. The total floor area is approximately  $290m^2$  ( $3120ft^2$ ). The envelope consists of walls constructed with a  $100mm$  brick layer on the exterior face, followed by layers of  $25.4mm$  air cavity,  $95mm$  expanded polystyrene insulation (EPS), plywood, and gypsum board on the inside. The below-grade wall consists of a layer of  $200mm$  concrete, followed by layers of  $25mm$  EPS insulation, plywood, and gypsum board on the inside. The roof is topped with a layer of roofing clay

tile and is insulated with a layer of 145mm EPS insulation. The windows are clear, double glazed units framed with polyvinylchloride (PVC).

The house is conditioned by a central heating system with a single-stage gas furnace with 84% efficiency. The living room area acts as a control zone as set out by the thermostatic control zone for unitary system with a programmable thermostat placed at the zone. The heating set-point was 19°C and the occupants do not set temperature back. The building is occupied by two adults and two children. The time occupants spent in each of the spaces were also inquired to design the desired temperature schedule and is given in Table 6.1.

**Table 6.1:** Occupancy Schedule.

	Weekdays	Weekends
Dining Room	07:00-09:00; 18:00-20:00	09:00-11:00;13:00-14:00;18:00-21:00
Kitchen	07:00-09:00; 17:00-19:00	09:00-11:00;13:00-14:00;17:00-20:00
Family Room	18:00-23:00	10:00-24:00
Living Room	18:00-23:00	10:00-24:00
Master Bedroom	24:00-09:00; 21:00-24:00	24:00-11:00;22:00-24:00
Bedroom 1	24:00-09:00; 21:00-24:00	24:00-11:00;22:00-24:00
Bedroom 2	24:00-09:00; 21:00-24:00	24:00-11:00;22:00-24:00
Study Room	21:00-23:00	-

### 6.3.2 Performance Evaluation Indices

The most important indices for evaluating the performance of the heating systems are: thermal comfort and energy consumption.

The thermal comfort refers to the effects of climatic impact on human response. Many researchers have been exploring ways to predict the thermal sensation of people in their environment based on the personal, environmental and physiological variables that influence thermal comfort. As stated earlier, Thermal comfort, as defined by ASHRAE 55-2004 is “that condition of mind which expresses satisfaction with the thermal environment”. Two of the most notable models used in predicting thermal comfort levels were developed by P.O. Fanger (the Fanger’s PMV Model) and the J. B. Pierce Foundation (the Pierce Two-Node Model).

The Fanger's PMV Model was developed from laboratory and climate chamber studies and is based on thermoregulation and heat balance theories [3]. According to these theories, the human body employs physiological processes, which includes sweating, shivering, and regulating blood flow to the skin to maintain a balance between the heat produced by metabolism and the heat lost from the body. Maintaining this heat balance is the first condition for achieving a neutral thermal sensation. Fanger determined that the only physiological processes influencing this heat balance in this context were sweat rate and mean skin temperature, and that these processes were a function of metabolic rate. Fanger then measured sweat rate and skin temperature on people who indicated they have achieved neutral thermal sensation. In these studies, participants were dressed in standardized clothing and completed standardized activities, while exposed to different thermal environments. Using the measured data, he then established a relationship between skin temperature and sweat secretion with metabolic rate using regression analysis with the measured data. Later, Fanger expanded on the relationship between physiological processes, thermal environment, and neutral thermal sensation by conducting laboratory and climate chamber studies with 1396 participants. The resulting equation combines four physical variables, namely air temperature, air velocity, mean radiant temperature, and relative humidity, and two personal variables, namely clothing insulation and metabolic rate into an index that can be used in predicting thermal comfort. Fanger's proposition was that, except for the four physical variables and two personal variables, other factors have no significant effects on the state of thermal comfort. This equation related thermal conditions to the seven-point ASHRAE thermal sensation scale (Table 6.2) and become known as the PMV index. The quality of the thermal environment may be expressed by the predicted percentage dissatisfied (PPD) index, which is related to the PMV value. For  $PMV = 0$ , PPD is equal to 5%, that is 5% of the occupants are dissatisfied with the thermal environment. A  $PMV = 0.5$  will correspond to 10% being dissatisfied. The ASHRAE Standard 55-1992 defines thermal comfort as satisfaction with the thermal environment by which 80% of the occupants would be thermally comfortable [6]. To meet this requirement, a PMV of -0.5 to +0.5 has been recommended.

**Table 6.2:** ASHRAE Thermal Sensation Scale.

-3	-2	-1	0	1	2	3
Cold	Cool	Slightly Cool	Neutral	Slightly Warm	Warm	Hot

Since the inception of the Fanger’s PMV thermal comfort model, the formulations and theories have been subjected to numerous peer reviews and validations. Within the context of this study, it is important to be aware of the limitations of the PMV model, specifically those related to the use of this model in predicting thermal comfort levels in thermal simulations. The formulation of the Fanger’s PMV model was based on data collected from laboratory and climate chamber studies, where the participants were subjected to constant conditions in a chamber for a period of time. This views occupants as passive recipients of thermal stimuli and overlooks the thermal adaption actions from occupants. Therefore, when interpreting the results generated by the Fanger’s PMV model, it has to be assumed that the adaptive natures of building occupants, including psychological adaption, physiological adaptations, and behavioral thermoregulation or adjustment, were not considered. The adaptive natures of building occupants, specifically those related to behavioral thermoregulation or adjustment, such as changing activities, clothing, putting on blankets, opening windows, and drinking cold or warm drinks, leads to dynamic conditions between the heat produced by metabolism and the heat lost from the body. However, the methodologies in determining comfort levels in simulation programs are based on steady-state conditions where predefined values have to be applied in the model and do not get updated with changing indoor environmental conditions. The prediction of PMV requires knowledge of the clothing insulation and the metabolic rate and the steady-state assumption cannot readily reflect the fluctuating metabolic rates and the clothing value with occupants. This problem is more severe in residential settings than office environments as there are more ways to adapt to the existing environment and have less predictable activities [164]. As well, the humidity ratio in the air might not be adequately accounted for in thermal simulation models. Moisture is generated in the house from domestic activities such as cooking and showering, and as part of the metabolic process of the occupants. Thermal simulation programs generally do

not account for these sources and therefore might not provide a representative prediction on the PMV. Although limitations occur in employing the Fanger's PMV model, it has served to be an indicator in providing thermal comfort evaluations for building spaces. With the limitations of the Fanger's PMV model acknowledged, in this research the performance is evaluated based on the temperature differences between desired and estimated values rather than looking into the actual measurement of thermal comfort index.

### **6.3.3 Analysis of Different Control Strategies**

The furnace control strategies are selected from: (1) Type 1: Single zone control with the supply air temperature set-point changing according to the thermostat reading. The thermostat installed in one room senses the temperature in that particular room not the entire building; (2) Type 2: Zoned control: one sensor is located in each room to sense the individual room temperatures. There are problems associated with the zoned control method. One is the smaller heated areas mean magnified heat loads such as an open window. This could potentially use more energy if the furnace is continually turned on. Fortunately, programming can re-adjust the set-point when it notices a problem until the room conditions are met and the system turns off. Currently, this type of complicated programming requires engineering fees which for some occupants may not seem reasonable (Goldshmidt, 2006). Another criticism of zoned control is pressure build-up. Static pressure is created in the duct work because of the increased resistance to airflow when more vents are closed. Variable speed fans again used with the system create less airflow resistance at the lower settings; (3) Type 3: Inferential control strategy.

The test conditions are as follows:

- Climatic data for Toronto, ON, Canada.
- The months included in heating season are January 1st to March 31st and October 1st to December 31st.
- The daily climatic conditions are described by the daily heating degree hours (HDH).

The daily HDH is calculated for the entire heating season and three days are accordingly identified to represent low, medium and high load, as shown in Table 6.3.

**Table 6.3:** HDH for Toronto for the Year of 2008.

Load	Month and Day of the Year 2008	HDH ( $^{\circ}\text{C.Hr}$ )
Low Load	October 13	21
Medium Load	March 21	484
High Load	February 28	773

- For multi-zone control, the building is divided into six zones as; Zone 1: Den/Kitchen, Zone 2: Family Room/Living Room, Zone 3: Master Bedroom, Zone 4: Bedroom 1, Zone 5: Bedroom 2, Zone 6: Study Room. Air temperature in each zone is used to assess the thermal comfort and the thermal comfort computation is based on the maintenance of the actual temperature at desired set-point. The desired temperature of each zone is calculated depending on the occupancy schedule as given in Fig. 6.2.
- The occupancy schedule for single-zone is given in Table 6.4. The desired temperature is determined based on the occupancy schedule and is shown in Fig. 6.3.

**Table 6.4:** Occupancy Schedule for Single-zone.

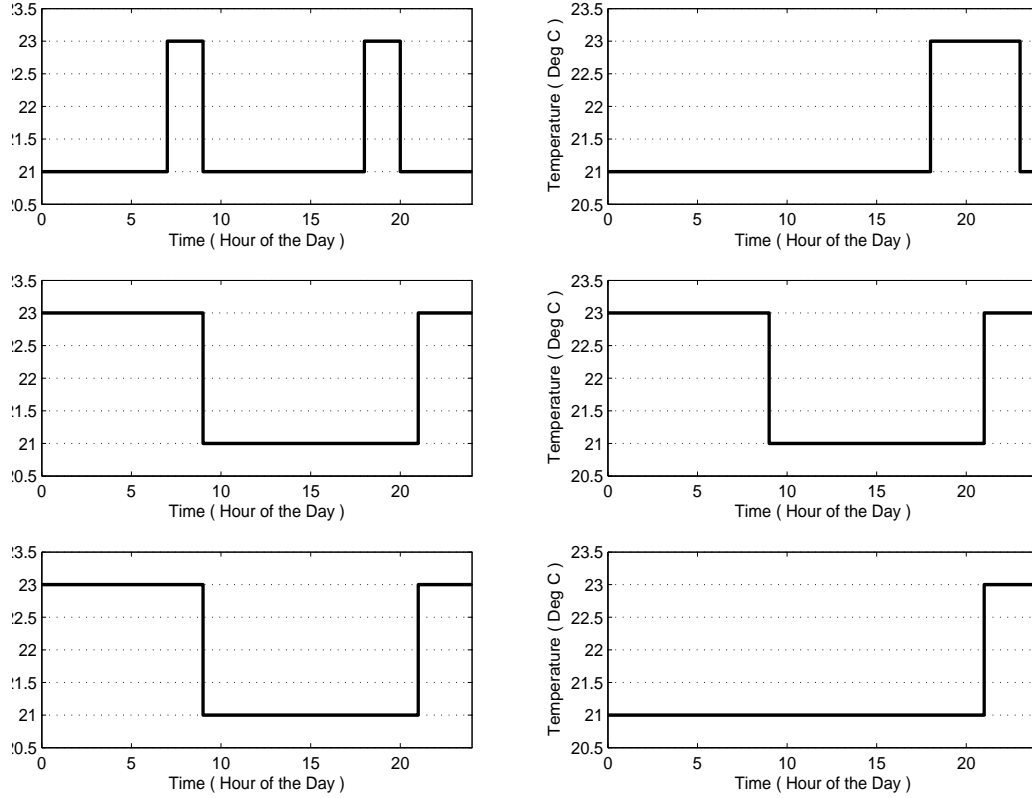
Weekday	05:00-09:00; 17:00-24:00
Weekend	07:00-24:00

- The furnace controller action depends upon the difference between the desired and the simulated temperatures. The thermostat set-point temperature and setback temperature are chosen to be  $23^{\circ}\text{C}$  and  $21^{\circ}\text{C}$ , respectively.
- Simulation environment: The three control schemes are simulated using SIMULINK modelling environment.

## 6.4 Performance of Different Control Strategies

The performance of three different types of control strategies is compared.

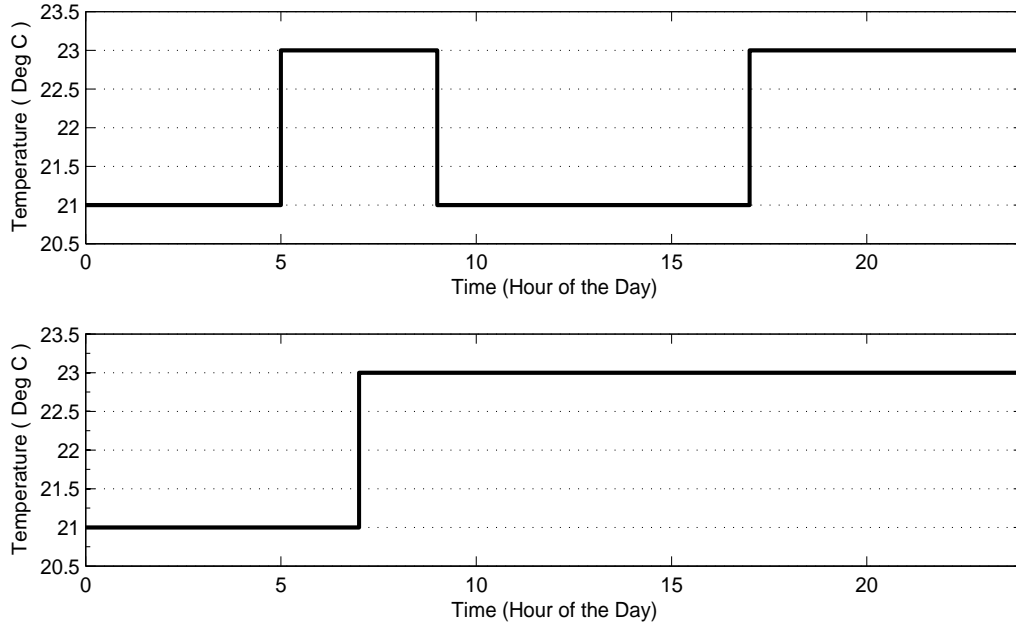




**Figure 6.2:** Desired Temperature Schedule for Multi-zone.

#### 6.4.1 The Performance of the Proposed Control Strategy in a Heating System with a Condensing and ON-OFF Control Furnace

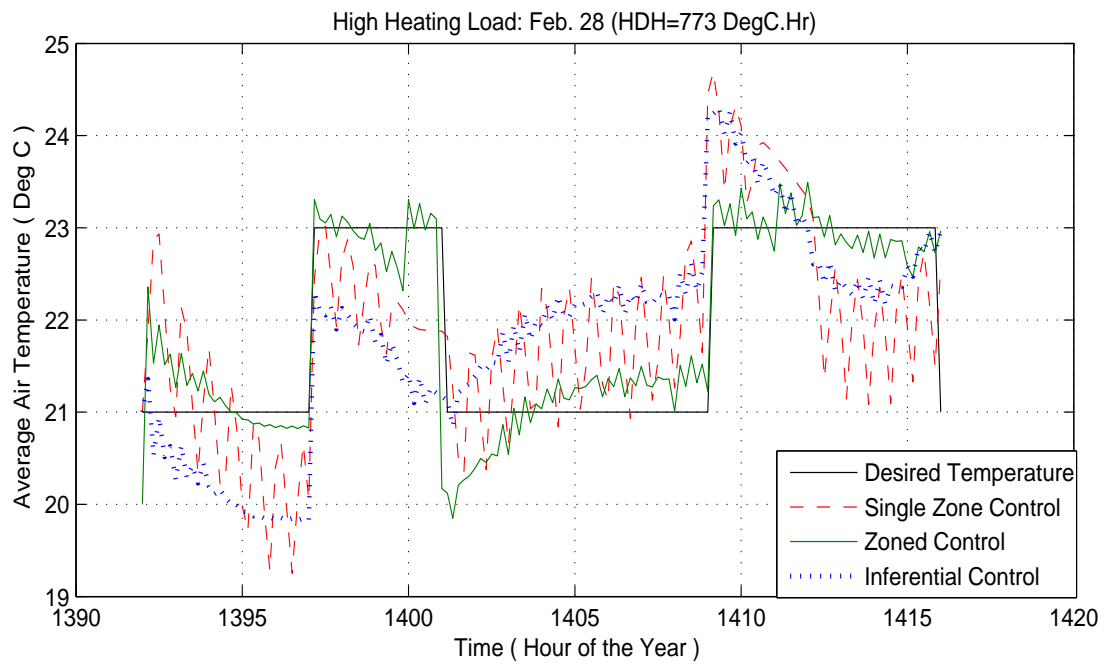
The inferential control strategy performance is compared with the single zone and zoned control strategies to demonstrate the improvement in the overall performance that the suggested scheme can make. Figs. 6.4, 6.5 and 6.6 compares daily profile of the average air temperature in a heating system with a single stage condensing furnace under different climatic conditions. The diverse climatic conditions are from low heating load through medium heating load to high heating load. The results show that the zoned control method is able to maintain the average air temperature within the desired range under all climatic conditions



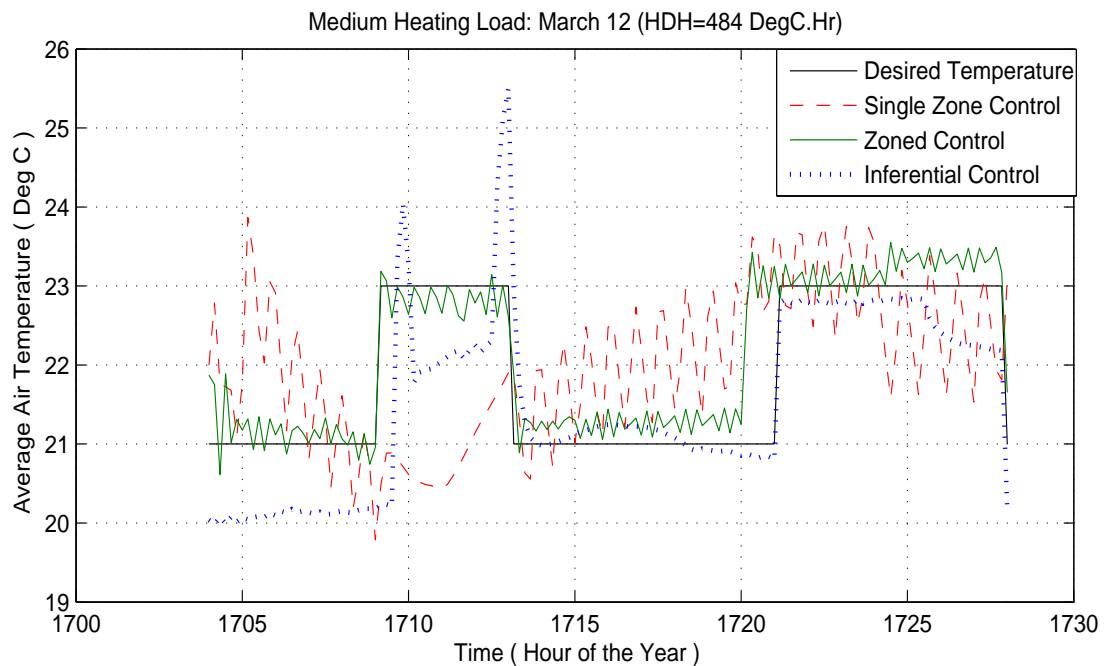
**Figure 6.3:** Desired Temperature Schedule for Single-zone.

while the conventional control methods fails to do so. The inferential control strategy also results into a temperature profile different from the desired temperature. The difference between actual and desired average air temperature is mainly because of the estimation error issues of the temperature estimator model. For example, the actual average room temperature will be maintained lower than the set-point if the estimator model estimates a higher average air temperature than the actual value.

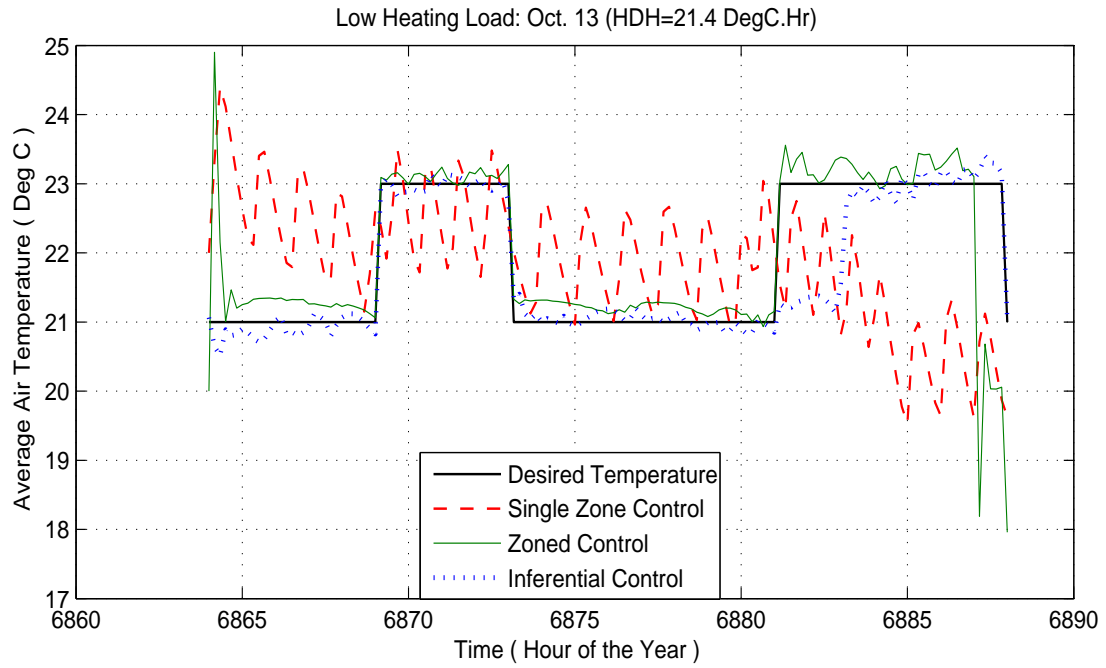
Table 6.5 compares the heating system energy consumption for the entire heating season. The heating system is controlled by three different types of control methods and is using a high efficiency furnace with ON-OFF control. The system consumes less energy when controlled with inferential control as compared to when controlled with single zone and zoned control methods. The system is more energy efficient when it is operated with zoned control as compared to when controlled with single zone control because extra energy is used to heat the unoccupied spaces as optimal scheduling is not applied to supply the heating capacity to individual zones. For the system employed with inferential control, the energy



**Figure 6.4:** Daily Average Air Temperature Profile in a Heating System with a Condensing Furnace and ON-OFF Control Logic for High Heating Load.



**Figure 6.5:** Daily Average Air Temperature Profile in a Heating System with a Condensing Furnace and ON-OFF Control Logic for Medium Heating Load.



**Figure 6.6:** Daily Average Air Temperature Profile in a Heating System with a Condensing Furnace and ON-OFF Control Logic for Low Heating Load.

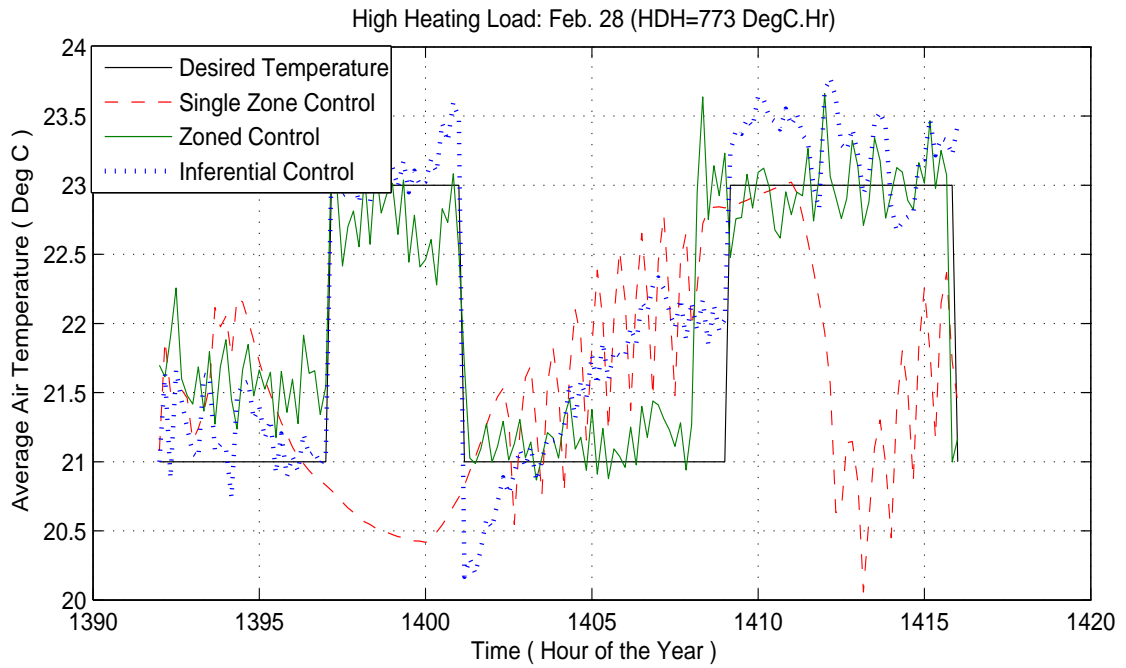
consumption is reduced by 11.7% and 15% when compared with the system incorporated with single zone and zoned control methods, respectively.

**Table 6.5:** Heating Energy Consumption in MJ for a Heating System with a Condensing Furnace and ON-OFF Control.

Month	Single Zone Control	Zoned Control	Inferential Control
January	19158	19600	16135
February	17122	17473	15891
March	15807	16337	14372
October	9692	10764	8232
November	12739	13451	10761
December	16363	16784	14862
Total	90881	94409	80253

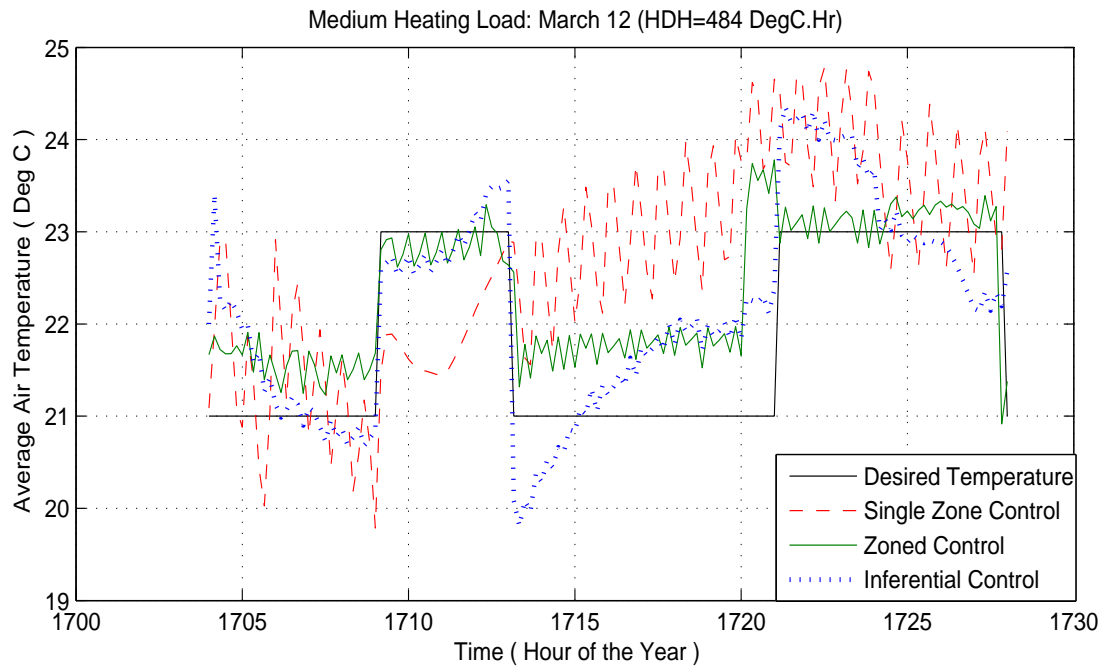
### 6.4.2 The Performance of the Proposed Control Strategy in a Heating Systems with a Non-Condensing and ON-OFF Control Furnace

Figs. 6.7, 6.8 and 6.9 compare the daily profile of the average air temperature in a heating system with a non-condensing furnace with ON-OFF control under diverse climatic conditions including low, medium and high heating loads. The results show that the zoned control method overestimates for medium load. For medium and high heating loads, the single zone control method underestimates and the estimation is higher than the set-point during setback periods. Inferential control scheme estimates higher than the set-point during setback periods for all the three climatic conditions.

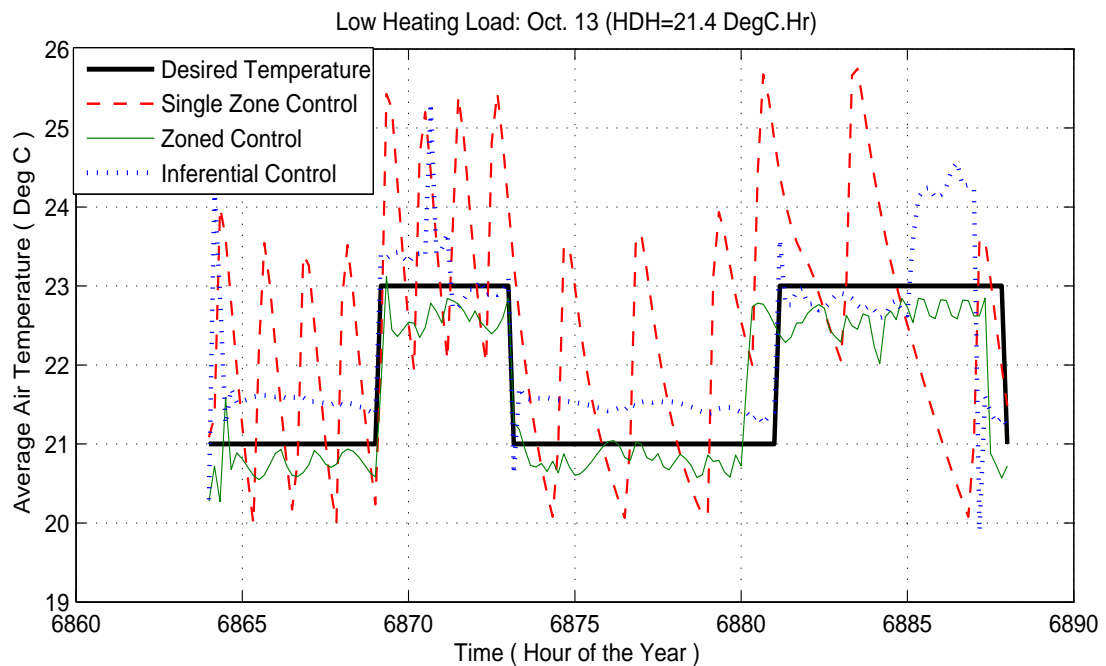


**Figure 6.7:** Daily Average Air Temperature Profile in a Heating System with a Non-Condensing Furnace and ON-OFF Control Logic for High Heating Load.

Table 6.6 compares the heating system energy consumption for the entire heating season. The heating system comprises a regular non-condensing furnace with ON-OFF control. Compared with single zone and zoned control methods, the inferential control strategy consumes less energy. When compared to the heating system with condensing furnace the



**Figure 6.8:** Daily Average Air Temperature Profile in a Heating System with a Non-Condensing Furnace and ON-OFF Control Logic for Medium Heating Load.



**Figure 6.9:** Daily Average Air Temperature Profile in a Heating System with a Non-Condensing Furnace and ON-OFF Control Logic for Low Heating Load.

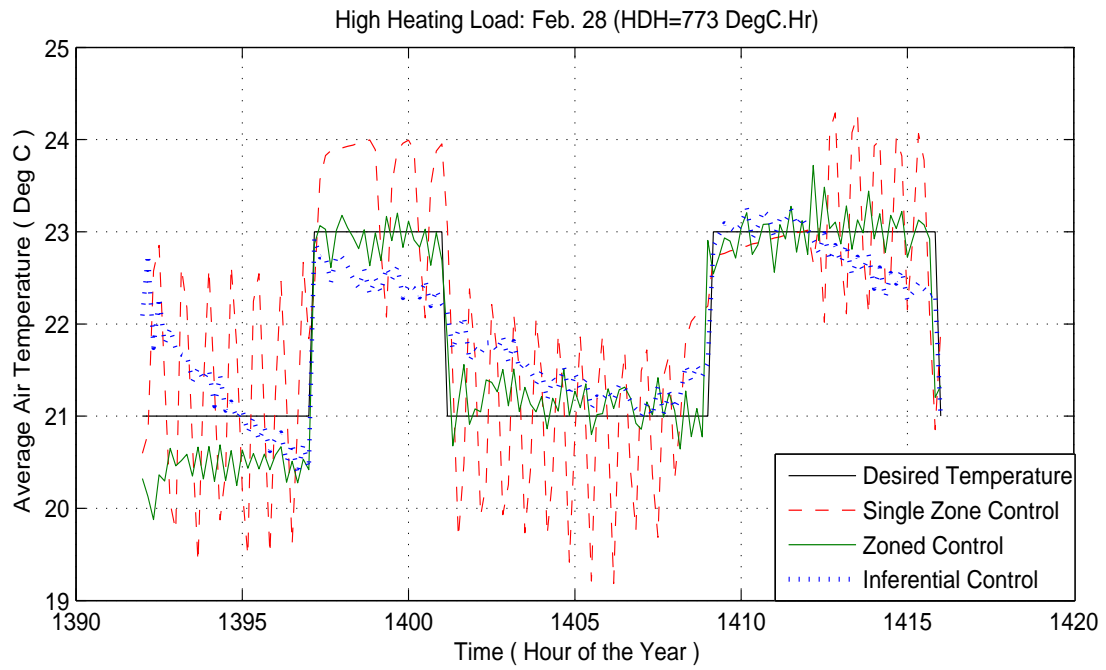
energy consumption is more with all the three control methods. The heating system with non-condensing furnace consumes 3%, 4% and 3.5% more energy compared to the heating system with condensing furnace when controlled with single zone, zoned and the inferential control methods, respectively. Energy consumption is 12% and 14.5% less when the heating system is controlled with inferential control method as compared to when controlled with single zone and zoned control methods, respectively.

**Table 6.6:** Heating Energy Consumption in MJ for a Heating System with a Non-Condensing Furnace and ON-OFF Control.

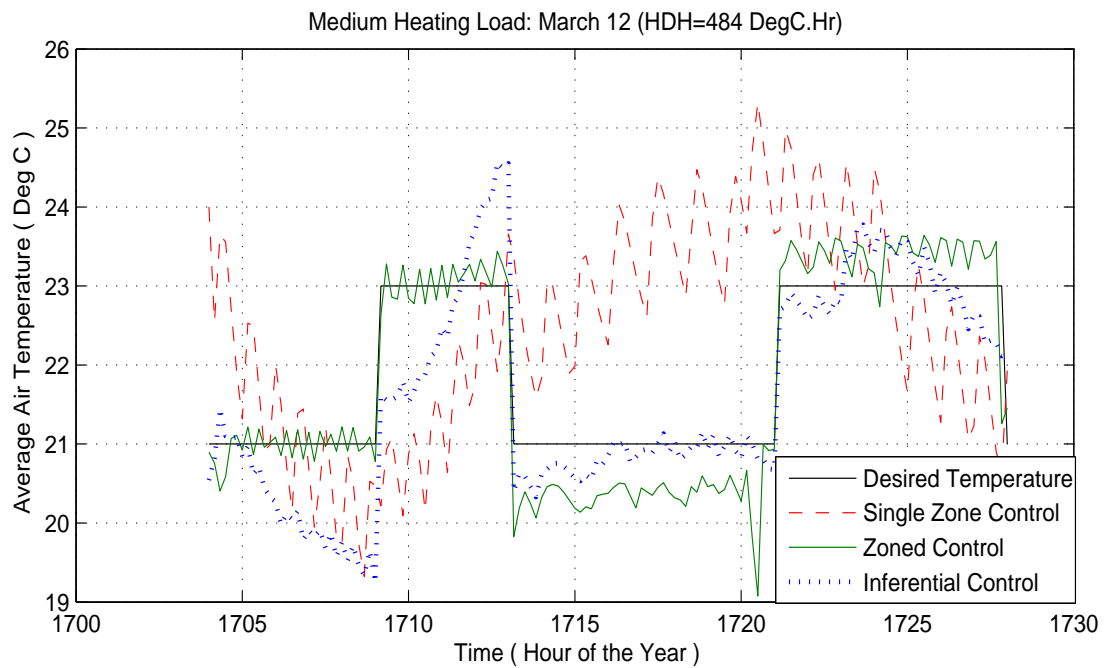
Month	Single Zone Control	Zoned Control	Inferential Control
January	19983	20186	17923
February	17212	17921	15402
March	16692	17012	14320
October	10127	11121	9010
November	13721	14000	12215
December	16724	16985	14220
Total	94459	97225	83090

### 6.4.3 The Performance of the Proposed Control Strategy in a Heating Systems with a Condensing Furnace, Controlled Blower Speed and Combustion Level

Figs. 6.10, 6.11 and 6.12 compare the daily profile of the average air temperature in a heating system with a condensing furnace. The blower speed and combustion level are controlled based on the output of the inferential sensor. The diverse climatic conditions are implemented ranging from low heating load through medium heating load to high heating load. The results show that, although single zone control method has large amplitude of the instantaneous temperature variations, for high load climatic conditions all the three methods are making similar estimations. The zoned control method estimates lower than the set-point temperatures during setback periods for medium load. For low and medium load conditions the zoned control strategy estimations are closer to the set-point temperature, while single zone and inferential control method estimations fail to follow the desired temperature values.

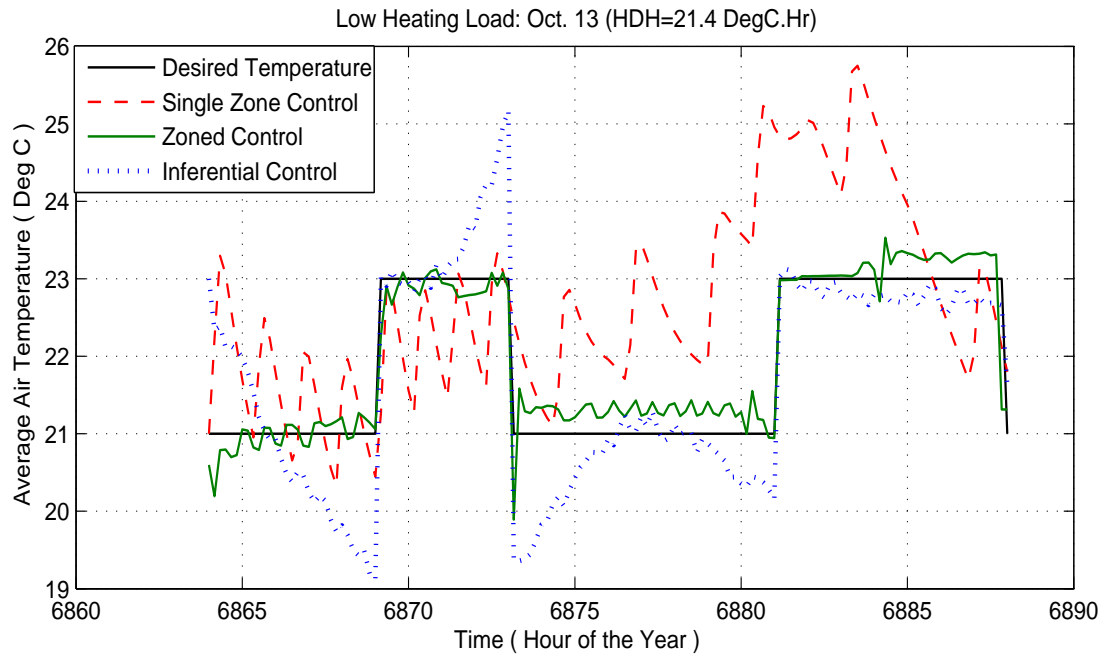


**Figure 6.10:** Daily Average Air Temperature Profile in a Heating System with a Condensing Furnace, Controlled Blower Speed and Combustion Level for High Heating Load.



**Figure 6.11:** Daily Average Air Temperature Profile in a Heating System with a Condensing Furnace, Controlled Blower Speed and Combustion Level for Medium Heating Load.





**Figure 6.12:** Daily Average Air Temperature Profile in a Heating System with a Condensing Furnace, Controlled Blower Speed and Combustion Level for Low Heating Load.

Table 6.7 shows the energy consumed by the heating system for the entire heating season, when controlled with three different control methods. The blower speed and flame level are controlled based on the output of the air temperature estimator. Again, the heating system controlled with an inferential control strategy consumes less energy as compared to the heating system controlled with single zone and zoned control methods. When compared to the heating system with a one stage condensing furnace the energy consumption is less, as the furnace is not running at full load under medium and low heating load climatic conditions. The heating system consumes 3%, 5.6% and 3.4% less energy compared with one stage furnace based heating system, when controlled with single zone, zoned and the inferential control methods, respectively. Energy consumption is improved by 12% and 13% when the heating system is controlled with the proposed control method as compared to when controlled with single zone and zoned control methods, respectively.

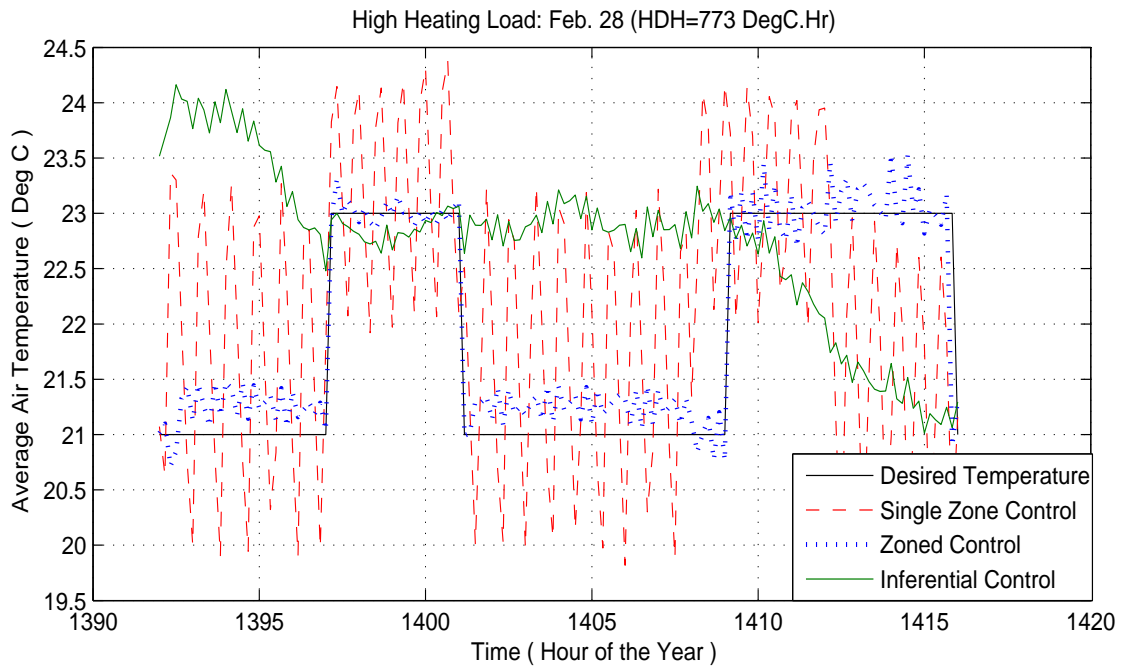
**Table 6.7:** Heating Energy Consumption in MJ for a Heating System with a Condensing Furnace, Controlled Blower Speed and Combustion Level.

Month	Single Zone Control	Zoned Control	Inferential Control
January	18234	18502	15921
February	16842	16920	15308
March	15222	15519	14102
October	9421	9612	7727
November	12514	12731	10214
December	15712	15824	14278
Total	87945	89108	77550

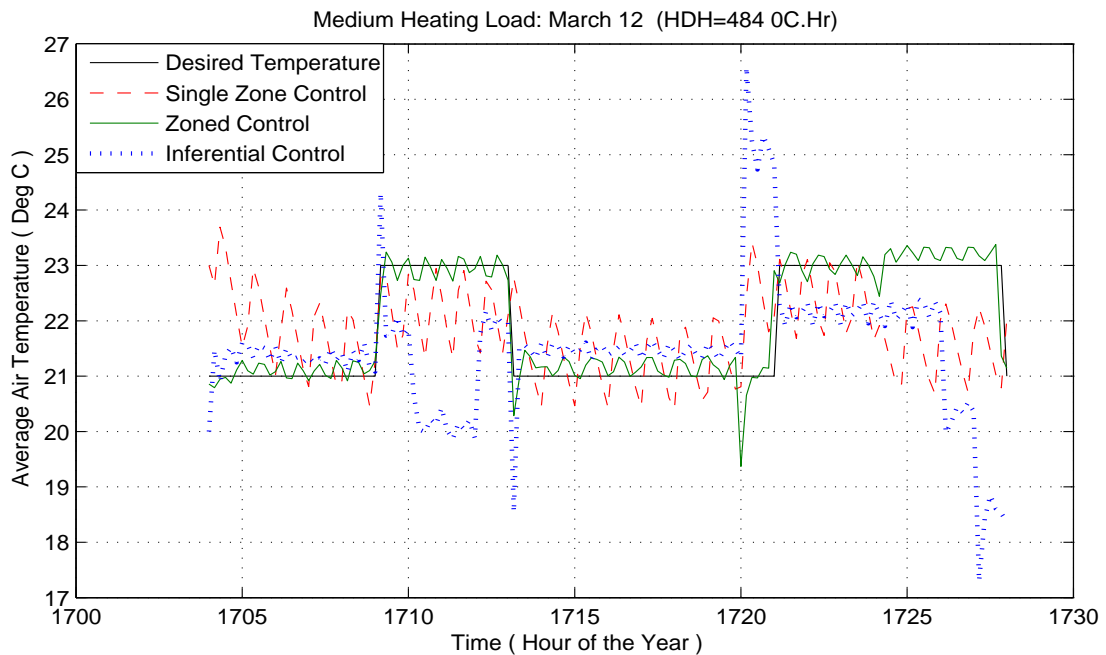
#### 6.4.4 The Performance of the Proposed Control Strategy in Heating Systems with a Non-Condensing Furnace, Controlled Blower Speed and Combustion Level

Figs. 6.13, 6.14 and 6.15 compare the daily profile of the average air temperature in a heating system with a non-condensing furnace under diverse climatic conditions including high, medium and low heating loads. The results show that the inferential controller estimation is higher than the desired temperature for high heating load conditions. Single zone control method follows the lower and higher temperature limits, but variations have very large amplitudes. The zoned control strategy estimations are very close to the desired temperature. For low and medium and high load conditions, the zoned control method is able to maintain the average air temperature within the desired range while the conventional and inferential control methods fails to do so.

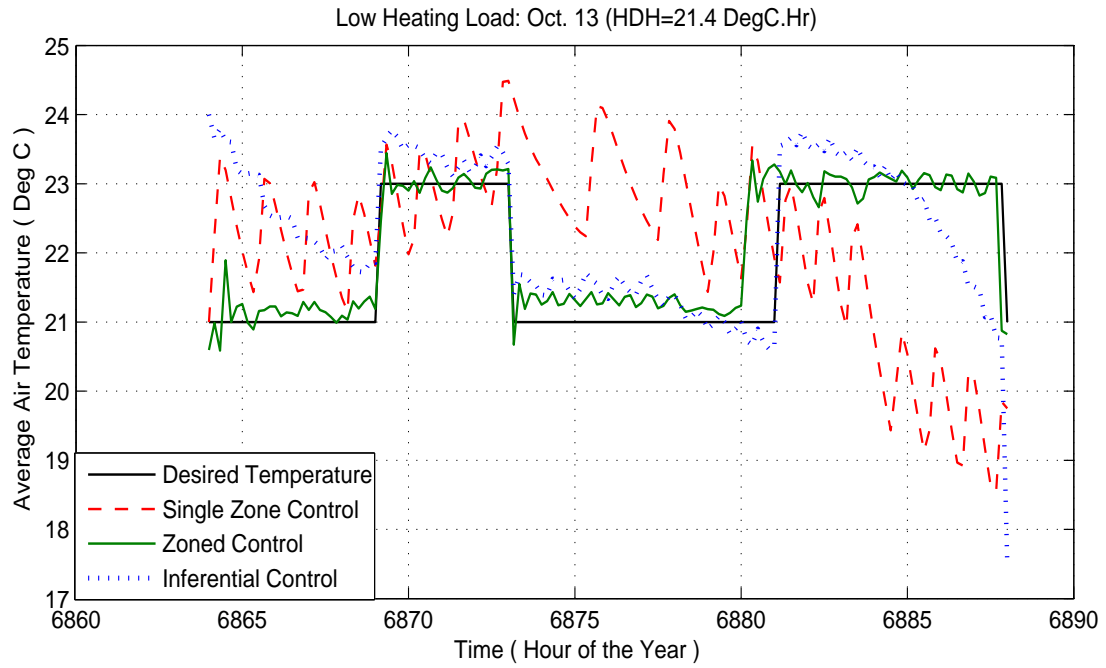
Table 6.8 presents the energy consumption for the heating system controlled with three different types of control methods. The analysis is carried out for the entire heating season. The heating system uses a non-condensing furnace to produce the desired heating capacity. For this scenario, the heating system controlled with the proposed control strategy consumes less energy when compared with the heating system controlled with single zone and zoned control methods. When compared to the heating system with non-condensing ON-OFF control furnace, the energy consumption is less for all the three types of control methods. The energy consumption is more compared with a heating system fed by a condensing fur-



**Figure 6.13:** Daily Average Air Temperature Profile in a Heating System with a Non-Condensing Furnace, Controlled Blower Speed and Combustion Level for High Heating Load.



**Figure 6.14:** Daily Average Air Temperature Profile in a Heating System with a Non-Condensing Furnace, Controlled Blower Speed and Combustion Level for Medium Heating Load.



**Figure 6.15:** Daily Average Air Temperature Profile in a Heating System with a Non-Condensing Furnace, Controlled Blower Speed and Combustion Level for Low Heating Load.

nance, validated by the fact that a condensing furnace is more efficient than a non-condensing furnace. The energy consumption for a heating system incorporated with an inferential controller is improved by an amount of 11% and 14%, compared to a heating system controlled with single zone and zoned control methods, respectively.

**Table 6.8:** Heating Energy Consumption in MJ for a Heating System with a Non-Condensing Furnace, Controlled Blower Speed and Combustion Level.

Month	Single Zone Control	Zoned Control	Inferential Control
January	18861	19715	16001
February	16991	17313	15514
March	15714	1103	14220
October	9520	9685	8654
November	12931	13110	10430
December	15936	16441	14917
Total	89953	92367	79736

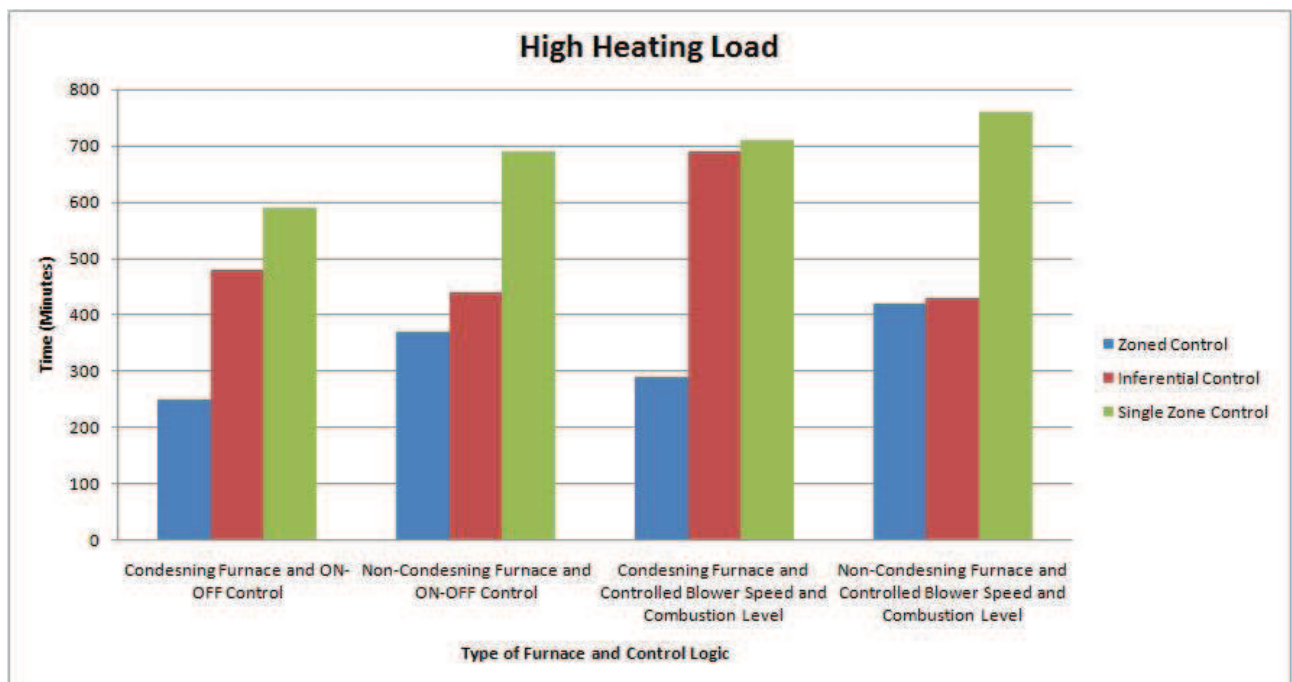
## 6.5 Concluding Remarks

The RenFIS based inferential sensor is applied to design an inferential control strategy that can improve the operation of residential heating systems. The performance of the inferential control strategy has been investigated through a simulation study. The indices used for evaluating the performance of the control methods are: the ability to maintain temperature at set-point and the energy consumption. The performance of heating system controlled with three different methods including, (1) Type 1: conventional single zone control, (2) Type 2: Zoned control, and (3) Type 3: inferential control, is compared. The diverse climatic conditions are implemented ranging from low heating load through medium heating load to high heating load.

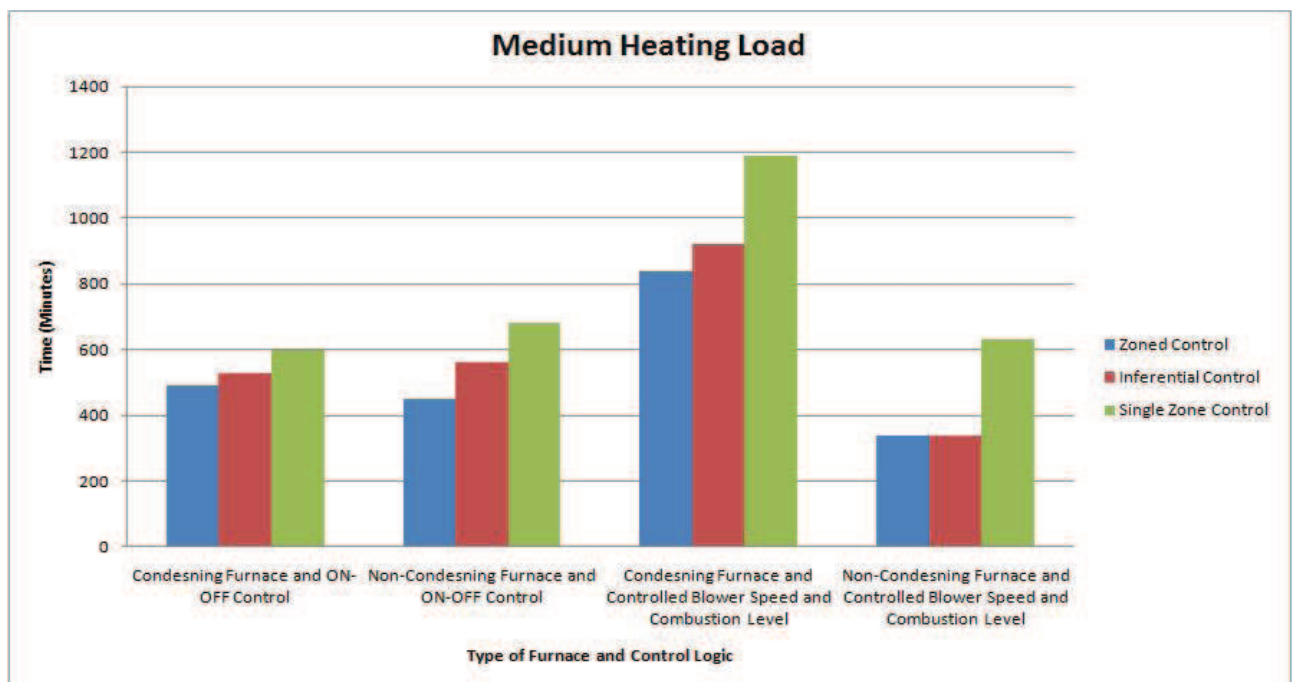
Figs. 6.16, 6.17, and 6.18 show the length of time for which the different schemes are not maintaining the temperature within desirable range on daily basis. Under all climatic conditions the zoned control method best maintains the temperature at set-point and inferential control method has median ability out of the three control methods and single zone control is least efficient to maintain the temperature within desirable range.

Results show that, for a heating system with a condensing furnace and ON-OFF control, the zoned control method is able to maintain the room temperature within the desired range under all climatic conditions while the conventional control method fails to do so. The inferential control strategy is also able to provide a temperature profile which follows the desired temperature. The heating system controlled with inferential control consumes less energy as compared to the heating system controlled with single zone and zoned control methods. The heating system controlled with zoned control scheme consumes more energy as compared to the heating system controlled with single zone control method. For the system employed with inferential control, the energy consumption is reduced by 11.7% and 15% when compared with the system incorporated with single zone and zoned control methods, respectively.

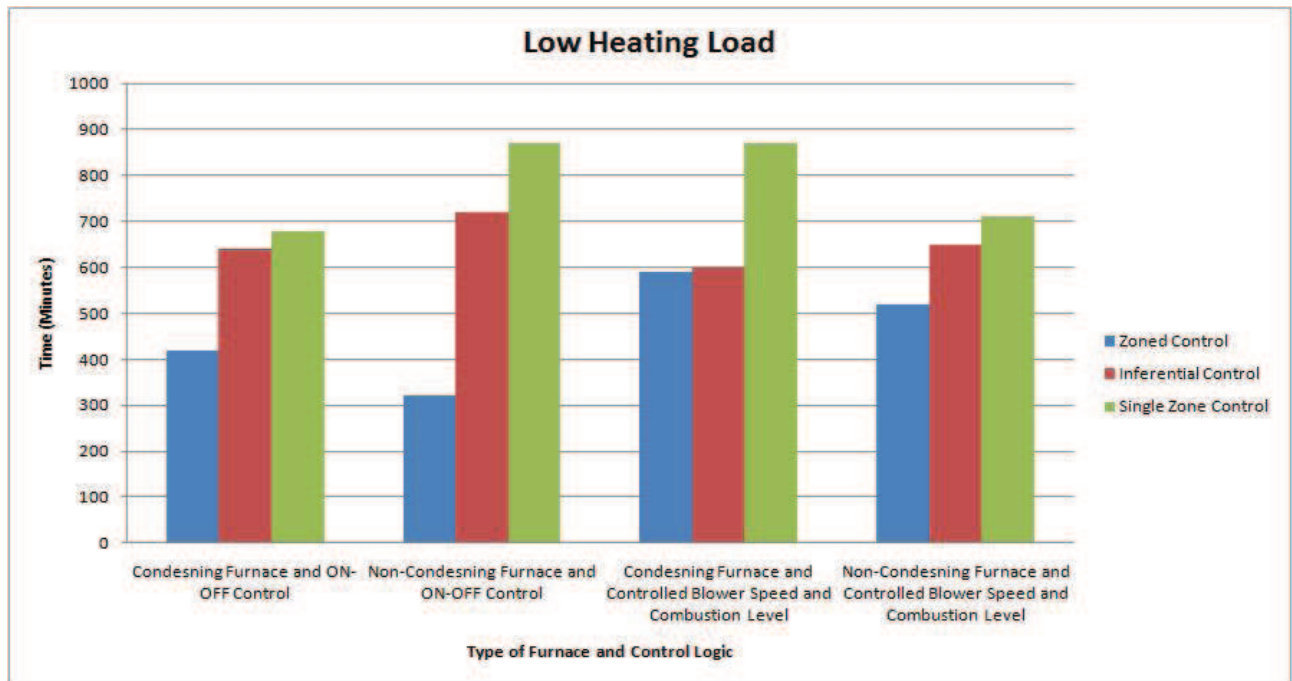
For a heating system with a non-condensing furnace and ON-OFF control logic, the zoned control method overestimates for medium load condition. For medium and high load



**Figure 6.16:** Daily Proportion of Time the Temperature is not Maintained within Desirable Range for High Heating Load Conditions.



**Figure 6.17:** Daily Proportion of Time the Temperature is not Maintained within Desirable Range for Medium Heating Load Conditions.



**Figure 6.18:** Daily Proportion of Time the Temperature is not Maintained within Desirable Range for Low Heating Load Conditions.

conditions, The single zone control method underestimates and estimates higher than the set-point during setback periods. Inferential control scheme estimates higher than the set-point during setback periods for all the three climatic conditions. As compared to the heating system with condensing furnace the energy consumption is more. The heating system with non-condensing furnace consumes 3%, 4% and 3.5% more energy compared to the heating system with condensing furnace when single zone, zoned and the inferential control methods are respectively applied.

Single zone control method, for a heating system with a condensing furnace and blower speed and combustion level are controlled based on the output of the temperature estimator, has large amplitude of the instantaneous temperature variations. The zoned control method estimates lower than the set-point temperatures during setback periods for medium load conditions. For low and medium load conditions the zoned control strategy estimations are closer to the set-point temperature, while single zone and inferential control methods estimations are away from the desired temperature values. Energy consumption is improved

by 12% and 13% when the heating system is controlled with an inferential control method as compared to when controlled with single zone and zoned control methods, respectively.

The results show that the inferential controller estimation is higher than the desired temperature for high load conditions when incorporated to a heating system fed from a non-condensing furnace and the blower speed is controlled based on the output of the temperature estimator. Single zone control method follows the lower and higher temperature limits, but variations have very large amplitudes. The zoned control strategy estimations are very close to the desired temperature. For low and medium and high load conditions, the zoned control method is able to maintain the room temperature within the desired range while the conventional and inferential control methods fails to do so. With incorporation of inferential control technology, the heating system energy consumption is 11% and 14% less compared to the heating system controlled with single zone and zoned control methods, respectively.

In addition to improved energy consumption and indoor environment, the inferential control technology is simple and inexpensive to implement. This can considerably reduce the initial cost and operating cost of building automation systems. The initial cost can be reduced because these schemes are self-adaptive and more robust to changing operating conditions than single zone. Consequently building automation systems become more affordable and more buildings can benefit from the automation techniques.



# Chapter 7

## Conclusion and Future Work

### 7.1 Conclusion of the Dissertation Work

This research investigated the design and development of RenFIS based modelling environment and explored its application in building automation. RenFIS has been further applied to develop an inferential sensor and the inferential control technology that can improve the control of the building heating systems, reducing energy consumption and improving the indoor environment quality.

The following conclusions can be drawn based on the research results presented in this dissertation:

- The theoretical studies show that RenFIS can significantly improve the performance of the technologies based on conventional ANFIS. The building heating system dynamics are effectively represented incorporating the feedback loop to the conventional ANFIS structure. The structure identification and learning algorithms are theoretically analyzed, to optimize RenFIS performance.
- RenFIS based average air temperature estimator model is structured and trained. The developed model can accurately estimate the temperature based on the information available to the heating equipment once it is trained using short-term monitoring data. The results show that the dynamical behavior of the system can be better represented

using feedback systems, which significantly improves the performance of the model. Results show that the inferential model performance, measured by RMSE, is significantly improved when structured and trained with RenFIS based modelling environment. The positive and negative peak for error are also reduced from  $11^{\circ}C$  to  $0.35^{\circ}C$  and from  $-13^{\circ}C$  to  $-0.75^{\circ}C$ , respectively. The experimental validation shows that the model is correctly structured and trained and is suitable for use as an estimator for incorporation of the inferential control scheme to the conventional control systems in the buildings.

- An experimental study is conducted to analyze the impact of data quality on the estimation ability of hybrid neuro-fuzzy based models. The analysis concludes that errors in the training data as well as in the testing data reduce the predictive accuracy of the model. Results show that the RMSE is improved by 37.5% and  $R^2$  is improved by 12% when trained with clean data. Therefore, this research highly recommends that the quality of the data set should be analyzed before they are applied in hybrid neuro-fuzzy based modelling algorithms.
- Sensitivity analysis shows that ANFIS and RenFIS are both sensitive to the selection of the training data. ANFIS gives better estimation of the process if trained with data representing most repetitive quality of the whole universe of the data set. While RenFIS, having a loop for feeding back the output, gives better performance if trained with the data set representing overall qualities of the whole experimentally collected data set. Robustness analysis is conducted before actual incorporation of the inferential control strategy to conventional control systems working under changing environment conditions. ANFIS has RMSE increased by 46% and RenFIS has RMSE increased by 6% when trained with noisy data. Results show that RenFIS has more robust behavior compared with ANFIS based model.
- A survey on the operation of these systems was conducted by the author. The survey results indicated a problem with the control of the conventional heating systems. The

analysis of results showed a correlation between the comfort level and the occupant's access to the control over room temperature. Respondents with higher heating system controllability tend to feel more thermally comfortable. In the multi-zone heating systems, the survey results show that 36% of the respondents experienced overheating in the winter.

- The current control practices for the forced warm-air heating systems has been reviewed. The analysis shows that the energy efficiency improvement methods are able to save energy and provide better indoor environment quality. For multi-zone buildings, multiple sensor control is the most efficient method for controlling the temperature of the indoor space, but very expensive and invasive to retrofit. Inferential control technology is an inexpensive and simple method to improve energy efficiency and thermal comfort for both the residential and commercial buildings.
- The RenFIS based inferential sensor is applied to design an inferential control algorithm that can improve the operation of residential heating systems. The performance of proposed control algorithm is compared with conventional and zoned control methods. The two performance evaluation indices are: (1) the ability to maintain temperature at set-point and (2) the energy consumption. The results show that temperature extremes, either higher or lower than the set point, are lesser as compared to the conventional system and energy consumption is low. Zoned control method when compared to proposed control technology, the temperature variations are less and energy consumption is high, but is very expensive and invasive to retrofit. The inferential control technology is simple and inexpensive. This can considerably reduce the initial cost and operating cost of the building automation systems. The initial cost can be reduced because these schemes are self-adaptive and more robust to changing operating conditions. Consequently building automation systems become more affordable and more buildings can benefit from the automation techniques.

## 7.2 Future Work

A number of future research directions naturally grow out of the work reported in this dissertation.

- *Prototyping:* The inferential sensor and inferential control schemes can work as standalone sensors and controllers that can be used in conventional control systems. They can also be implemented in BMS (Building Management Systems).

The developed control technology should be implemented in hardware prototypes for test in laboratory and through field trials in real installations to reduce the environmental impact of energy consumption in both residential and non-residential buildings. Further research can concentrate on the implementation of the RenFIS based inferential sensor in MEMS, so that it can be used in the industry for other applications and also for commercial built environments.

- *Applicability of the proposed scheme in other areas:* The applicability of this technique will be extended to different areas including building automation, intelligent automobile control, and structure diagnosis. For example, RenFIS can be employed to develop intelligent automobile control unit to improve the safety of driving. Based on the measured acceleration of certain parts of an automobile, a RenFIS based instrument can determine the moving performance of the vehicle thus provide safe driving support information. Furthermore, RenFIS can handle the complex mathematical relationship exhibited in structure monitoring systems, providing key information for the diagnosis of structure safety and service-ability.
- *Incorporation of the prior knowledge:* RenFIS performance is based on the structure identification and training process. With growing number of rules, both the structure identification and the training algorithms become complex. Number of rules can be reduced by incorporating expert knowledge. Further research can concentrate on the simplification of structure identification and the training process by augmenting the

model with prior-knowledge.

# Bibliography

- [1] L. Bernstein, *Climate Change 2007: Synthesis Report, Intergovernmental Panel on Climate Change (IPCC, 2007)*, 4th Assessment Report, Cambridge University Press, 2008.
- [2] J.T. Houghton, *Climate Change 2001: the Scientific Basics, Intergovernmental Panel on Climate Change (IPCC, 2001)*, 3rd Assessment Report, Cambridge University Press, 2001.
- [3] P.A. Fanger, *Thermal Comfort: Analysis and Applications in Environment Engineering*, McGraw-Hill, 1973.
- [4] J. A. Leech, K. Wilby, E. McMullen, and K. Laporte, “The canadian human activity pattern survey:report of methods and population surveyed,” *Chronic Diseases in Canada*, vol. 17, no. 3-4, pp. 118–123, 1996.
- [5] C.C. Federspiel, “Statistical analysis of unsolicited thermal sensation complaints in commercial buildings,” *ASHRAE Transactions*, vol. 104, no. 1, pp. 912–923, 1998.
- [6] ASHRAE, *Thermal Environment Conditions for Human Occupancy*, American Society of Heating, Refrigerating and Air-Conditioning Engineers, 2004.
- [7] Statistics Canada, “Report on energy supply-demand in Canada,” *1990-2005*, May 2007.
- [8] P. Gardner, *Energy Management Systems in Buildings*, Energy Publications, 1984.

- [9] Z. Liao and A.L. Dexter, “An inferential control scheme for optimizing the operation of boilers in multi-zone heating systems,” *Building Services Engineering Research and Technology*, vol. 24, no. 4, pp. 245–256, 2003.
- [10] Z. Liao and A.L. Dexter, “A simplified physical model for estimating the average air temperature in multi-zone heating systems,” *Building and Environment*, vol. 39, no. 9, pp. 1009–1018, 2004.
- [11] M.R. Kulkarni and F. Hong, “Energy optimal control of a residential space-conditioning system based on sensible heat transfer modelling,” *Building and Environment*, vol. 39, no. 1, pp. 31–38, 2004.
- [12] B. Sohlberg, “Grey box modelling for model predictive control of heating process,” *Journal of Process Control*, vol. 13, no. 3, pp. 225–238, 2003.
- [13] A. A. Argiriou, I. Bellas-Velidis, and C.A. Balara, “Development of a neural network heating controller for solar buildings,” *Neural Networks*, vol. 13, no. 7, pp. 811–820, 2000.
- [14] D.W. Pei, “On the strict logic foundation of fuzzy reasoning,” *Soft Computing*, vol. 8, no. 8, pp. 539–545, 2004.
- [15] J.S.R. Jang, C.T. Sun, and E. Mizutani, *Neuro-fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence*, Prentice-Hall, 1997.
- [16] H.B. Verbruggen and P.M. Bruijin, “Fuzzy control and conventional control: what is (and can be) the real contribution of fuzzy systems?,” *Fuzzy Sets and Systems*, vol. 90, no. 2, pp. 151–160, 1997.
- [17] Ning Wang, Zhong Zheng, and Honglu Chen, “Model-free PID controller with gain scheduling for turning processes,” *IEEE International Conference on Systems, Man and Cybernetics*, pp. 2424–2429, 2003.
- [18] L.X. Wang, *Adaptive Fuzzy Systems and Control*, Prentice-Hall, 1994.

- [19] C.T. Lin and C.S.G. Lee, *Neural Fuzzy Systems: A Neuro-Fuzzy Synergism to Intelligent Systems*, Prentice Hall, 1996.
- [20] J.S.R. Jang, C.T. Sun, and E. Mizutani, “ANFIS: Adaptive-network based fuzzy inference system,” *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 23, no. 3, pp. 665–685, 1993.
- [21] J.J.E. Slotine and W. Li, *Applied Nonlinear Control*, Prentice-Hall, 1991.
- [22] L. A. Zadeh, “Fuzzy sets,” *Information and Control*, vol. 8, no. 3, pp. 338–353, 1965.
- [23] C.T. Lin and C.S.G. Lee, “Neural-networks-based fuzzy logic control and decision system,” *IEEE Transactions on Computers*, vol. 40, no. 12, pp. 1320–1366, 1991.
- [24] L. A. Zadeh, *Toward a Theory of Fuzzy Systems*, Rinehart and Winston, 1971.
- [25] L. A. Zadeh, “Outline of a new approach to the analysis of complex systems and decision processes,” *IEEE Transactions on Systems, Man and Cybernetics*, vol. 3, no. 1, pp. 28–44, 1973.
- [26] E.H. Mamdani and S. Assilian, “An experiment in linguistic synthesis with a fuzzy logic controller,” *International Journal of Man-Machine Studies*, vol. 7, no. 1, pp. 1–13, 1975.
- [27] M. Sugeno and G.T. Kang, “Structure identification of fuzzy model,” *Fuzzy Sets and Systems*, vol. 28, no. 1, pp. 15–33, 1988.
- [28] T. Takagi and M. Sugeno, “Derivation of fuzzy control rules from human operator’s control action,” *IFAC Fuzzy Information*, 1983.
- [29] R.Hecht-Nielsen, *Neurocomputing*, Addison-Wesley, 1989.
- [30] T. Kohonen, K. Mäkelä, O. Simula, and J. Kangas, “Self-organizing maps: An optimization approach,” *Artificial Neural Networks, Amsterdam: North Holland*, vol. 2, no. 2, pp. 981–990, 1991.



- [31] S. Horikawa, T. Furuhashi, , and Y. Uchikawa, “On fuzzy modelling using fuzzy neural networks with the back-propagation algorithm,” *IEEE Transactions on Neural Networks*, vol. 3, no. 5, pp. 801–806, 1993.
- [32] C.C. Hung, “Building a neuro-fuzzy learning control system,” *AI Expert*, vol. 8, no. 11, pp. 40–49, 1993.
- [33] H. Ishibuchi, H. Tanaka, and H. Okada, “Interpolation of fuzzy If-Then rules by neural networks,” *International Journal of Approximate Reasoning*, vol. 10, no. 1, pp. 3–27, 1994.
- [34] J.J. Shann and H.C. Fu, “A fuzzy neural network for rule acquiring on fuzzy control system,” *Fuzzy Sets and Systems*, vol. 71, no. 3, pp. 345–357, 1995.
- [35] V.A. Constantin, *Fuzzy Logic and Neuro-Fuzzy Applications Explained*, Prentice Hall, 1995.
- [36] Y.C. Lee, C.H. Hwang, and Y.P. Shih, “A combined approach to fuzzy model identification,” *IEEE Transactions on Systems, Man and Cybernetics*, vol. 3, no. 2, pp. 190–198, 1994.
- [37] H.R. Berenji and P. Khedkar, “Learning and tuning fuzzy logic controllers through reinforcements,” *IEEE Transactions on Neural Networks*, vol. 3, no. 5, pp. 724–740, 1992.
- [38] C.F. Juang and C.T. Lin, “An online self constructing neural fuzzy inference network and its applications,” *IEEE Transactions on Fuzzy Sets*, vol. 6, no. 1, pp. 12–32, 1998.
- [39] D. Nauck, F. Klawonn, and R. Kruse, *Foundations of Neuro-Fuzzy Systems*, Wiley, 1997.
- [40] S. Tano, T. Oyama, and T. Arnould, “Deep combination of fuzzy inference and neural network in fuzzy inference,” *Fuzzy Sets and Systems*, vol. 82, no. 2, pp. 151–160, 1996.

- [41] J. Kim and N. Kasabov, “HyFIS: Adaptive neuro-fuzzy inference system and their application to nonlinear dynamical systems,” *Neural Networks*, vol. 12, no. 9, pp. 1301–1319, 1999.
- [42] N. Kasabov and Q. Song, “DENFIS: Dynamic evolving neural-fuzzy inference system and their application for time-series prediction,” *IEEE Transactions on Fuzzy Systems*, vol. 10, no. 2, pp. 144–154, 2002.
- [43] J.S.R. Jang and C. T. Sun, “Neuro-fuzzy modelling and control,” *Proceedings of IEEE*, vol. 83, no. 3, pp. 378–406, 1995.
- [44] A. Abraham and B. Nath, “Optimal design of neural nets using hybrid algorithms,” *Proceedings of 6th Pacific Rim International Conference on Artificial Intelligence*, pp. 510–520, 2000.
- [45] A. Valishevsky, “Adaptive learning algorithm for hybrid fuzzy system,” *Proceedings of the International Scientific Conference Traditions and Innovations in Sustainable Development of Society*, pp. 281–288, 2002.
- [46] S. Altug, M.Y. Chow, and H.J. Trussell, “Fuzzy inference systems implemented on neural architectures for motor,” *IEEE Transactions on Industrial Electronics*, vol. 46, no. 6, pp. 1069–1079, 1999.
- [47] E. Czogala and J. Leski, *Fuzzy and Neuro-Fuzzy Intelligent Systems*, Physica-Verlag, 2000.
- [48] M. Sugeno and T. Yasukawa, “A fuzzy-logic-based approach to qualitative modelling,” *IEEE Transactions on Fuzzy Systems*, vol. 1, no. 1, pp. 7–31, 1994.
- [49] Y. Lin, G.A. Cunningham, and S.V. Coggeshall, “Using fuzzy partitions to create fuzzy systems from input-output data and set initial weights in a fuzzy neural network,” *IEEE Transactions on Fuzzy Systems*, vol. 5, no. 4, pp. 614–621, 1997.

- [50] C.T. Sun, “Rule base structure identification in an adaptive-network-based fuzzy inference system,” *IEEE Transactions on Fuzzy Systems*, vol. 2, no. 1, pp. 64–73, 1994.
- [51] J.W. Miles and G.N. Minerbo, *The AIT Borehole Correction Implementation*, Technical Report, 1988.
- [52] J.C. Dunn, “A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters,” *Journal of Cybernetics*, vol. 3, no. 3, pp. 32–57, 1973.
- [53] R.R. Yager and D.P. Filev, “Approximate clustering via the mountain method,” *IEEE Transactions on Systems, Man and Cybernetics*, vol. 15, no. 1, pp. 116–132, 1994.
- [54] S. Chiu, “Fuzzy model identification based on cluster estimation,” *Journal of Intelligent and Fuzzy Systems*, vol. 2, no. 3, pp. 267–278, 1994.
- [55] X. Li and W. Yu, “Dynamic system identification via recurrent multilayer perceptrons,” *Information Science*, vol. 147, no. 1-4, pp. 45–63, 2002.
- [56] W. Yu, “Nonlinear system identification using discrete-time recurrent neural networks with stable learning algorithms,” *Information Science*, vol. 158, no. 1, pp. 131–147, 2004.
- [57] F.J. Lin, H.J. Shieh, P.H. Shieh, and P.H. Shen, “An adaptive recurrent-neural-network motion controller for X-Y table in CNC machine,” *IEEE Transactions on Systems, Man and Cybernetics-Part B: Cybernetics*, vol. 36, no. 2, pp. 286–299, 2006.
- [58] T.W.S. Chow and Y. Fang, “A recurrent neural-network-based real-time learning control strategy applying to nonlinear systems with unknown dynamics,” *IEEE Transactions on Industrial Electronics*, vol. 45, no. 1, pp. 151–161, 1998.
- [59] L.X. Wang and J.M. Mendel, “Generating fuzzy rules by learning from examples,” *IEEE Transactions on Systems, Man and Cybernetics*, vol. 22, no. 1, pp. 1414–1427, 1992.

- [60] C.F. Juang, “A TSK-type recurrent fuzzy network for dynamic systems processing by neural network and genetic algorithms,” *IEEE Transactions on Fuzzy Systems*, vol. 10, no. 2, pp. 155–170, 2002.
- [61] C.H. Lee and C.C. Teng, “Identification and control of dynamic systems using recurrent fuzzy neural networks,” *IEEE Transactions on Fuzzy Systems*, vol. 8, no. 6, pp. 349–366, 2000.
- [62] P.A. Mastorocostas and J.B. Theocharis, “A recurrent fuzzy-neural model for dynamic system identification,” *IEEE Transactions on Systems, Man and Cybernetics-Part B: Cybernetics*, vol. 32, no. 2, pp. 176–190, 2002.
- [63] J. Zhang, H. Chung, and W. Lo, “Chaotic time series prediction using a neuro-fuzzy system with time-delay coordinates,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 20, no. 7, pp. 956–964, 2008.
- [64] J. Zhang and A. J. Morris, “Recurrent neuro-fuzzy networks for nonlinear process modelling,” *IEEE Transactions on Neural Networks*, vol. 10, no. 2, pp. 313–326, 1999.
- [65] W.A. Farag, V.H. Quintana, and G.L. Torres, “A genetic-based neuro-fuzzy approach for modelling and control of dynamical systems,” *IEEE Transactions on Neural Networks*, vol. 9, no. 5, pp. 756–767, 1998.
- [66] Z. Liao and A.L. Dexter, “An experimental study on an inferential control scheme for optimizing the control of boilers in multi-zone heating systems,” *Energy and Buildings*, vol. 37, no. 1, pp. 55–63, 2005.
- [67] G. A. Montague, “Enhancing bioprocess operability with generic software sensors,” *Journal of Biotechnology*, vol. 25, no. 1-2, pp. 183–201, 1992.
- [68] M.T. Tham, G.A. Montague, and A.J. Morris, “Soft sensing: A solution to the problem of measurement delays,” *Chemical Engineering Research and Design*, vol. 67, no. 6, pp. 547–554, 1989.

- [69] K. A. Brodie, M.J. Willis, and M.T.Tham, “Adaptive inferential control,” *Proceedings IEE Colloquium Adaptive Controllers in Practice - Part Two*, pp. 9/1–9/3, 1996.
- [70] M.T. Guilandoust, A.J. Morris, and M.T. Tham, “Adaptive inferential control,” *Proceedings of IEE Pt. D*, vol. 134, no. 3, pp. 171–179, 1987.
- [71] M.T.Tham, G.A. Montague, A.J. Morris, and P.A. Lant, “Soft sensors for process estimation and inferential control,” *Journal of Process Control*, vol. 1, no. 1, pp. 3–14, 1991.
- [72] S. Wang and X. Jin, “Model-based optimal control of VAV air-conditioning system using genetic algorithm,” *Building and Environment*, vol. 53, no. 6, pp. 471–487, 2000.
- [73] J.A. Viera and A.M. Mota, “Combining first principles with grey-box approaches for modelling a water gas heater,” *Proceedings of the 13th Mediterranean Conference on Control and Automation*, pp. 1131–1136, 2005.
- [74] C. Ghiaus, A. Chicinas, and C. Inard, “Grey-box identification of air-handling unit elements,” *Control Engineering Practice*, vol. 15, no. 4, pp. 421–433, 2007.
- [75] D. Kolokosta, “Artificial intelligence in buildings: A review of the application of fuzzy logic,” *Advances in Building Energy Research*, pp. 29–54, 2007.
- [76] A.L. Dexter and D.W. Trewella, “Building control systems: Fuzzy rule-based approach to performance assessment,” *Building Services Research and Technology*, vol. 11, no. 4, pp. 115–124, 1990.
- [77] M. Santamouris, *Advances in Building Energy Research*, Earthscan Publishers, 2007.
- [78] A.I. Dounis, M.J. Santamouris, and C.C. Lefas, “Building visual comfort control with fuzzy reasoning,” *Energy Conservation and Manageme*, vol. 34, no. 1, pp. 17–28, 1993.
- [79] A.I. Dounis, M.J. Santamouris, C.C. Lefas, and A. Argiriou, “Design of a fuzzy set environment comfort system,” *Energy and Buildings*, vol. 22, no. 1, pp. 81–87, 1995.

- [80] A.I. Dounis, M.J. Santamouris, and C.C. Lefas, "Implementation of artificial intelligence techniques in thermal comfort control for passive solar buildings," *Energy Conservation and Management*, vol. 33, no. 3, pp. 175–182, 1992.
- [81] A.I. Dounis, C.C. Lefas, and A. Argiriou, "Knowledge-based versus classical control of solar-building designs," *Applied Energy*, vol. 50, no. 4, pp. 281–292, 1995.
- [82] P. Angelov, "An evolutionary approach to fuzzy rule-based model synthesis using indices for rules," *Fuzzy Sets and Systems*, vol. 137, no. 3, pp. 325–338, 2003.
- [83] P. Angelov, "A fuzzy approach to building thermal systems optimization," *Proceedings of the 8th IFSA World congress*, vol. 2, pp. 423–426, 1999.
- [84] V.I. Hanby, P. Angelov, and J.A. Wright, "HVAC systems simulation: A self-structuring fuzzy rule-based approach," *International Journal of Architectural Sciences*, vol. 1, no. 1, pp. 30–39, 2000.
- [85] B. Kosko, *Neural Networks and Fuzzy Systems*, Prentice-Hall, 1991.
- [86] J.F. Kreider, "Neural networks applied to building energy studies," *1995 Workshop on Parameter Identification*, pp. 243–251, 1995.
- [87] M. Ansett and J.F. Kreider, "Application of neural networking models to predict energy use," *ASHRAE Transactions*, vol. 99, no. 1, pp. 505–517, 1993.
- [88] M.R.B. Breckweg, P. Gruber, and O. Ahmed, "Development of generalized neural network model to detect faults in building energy performance-part I, part II," *ASHRAE Transactions: Research*, vol. 4372, pp. 61–93, 1993.
- [89] P.S. Curtiss, J.F. Kreider, and M.J. Brandemuehl, "Energy management in central HVAC plants using neural networks," *ASHRAE Transactions: Research*, vol. 99, no. 1, pp. 476–493, 1993.

- [90] D.A. Cohen and M. Krarti, “A neural network modelling approach applied to energy conservation retrofits,” *Proceedings of Fourth International Conference on Building Simulation*, vol. SA-96-3-1, pp. 405–411, 1996.
- [91] J.S. Haberl and S. Thamilseran, “Predicting hourly building energy use: the great energy predictor shootout II: Measuring retrofit savings - overview and discussion of results,” *ASHRAE Transactions*, vol. 102, no. 2, pp. 419–435, 1996.
- [92] S.J. Hepworth and A.L. Dexter, “Adaptive neural control with stable learning,” *Mathematics and Computers in Simulation*, vol. 41, no. 1, pp. 39–51, 1996.
- [93] M. Soleimani-Moheseni, B. Thomas, and Per Fahlen, “Estimation of operative temperature in buildings using artificial neural networks,” *Energy and Buildings*, vol. 38, no. 6, pp. 635–640, 2006.
- [94] M. Yalcintas and S. Akkurt, “Artificial neural networks applications in building energy prediction and a case study for tropical climates,” *International Journal of Energy Research*, vol. 29, no. 3, pp. 891–901, 2005.
- [95] A. D. Fuller, *Neural Fuzzy Systems*, Abo Akademi University, 1995.
- [96] S. Kurnaz, O. Kaynak, and E. Konakoglu, “Adaptive neuro-fuzzy inference system based autonomous flight control of unmanned air vehicle,” *Expert Systems with Applications*, vol. 37, no. 2, pp. 1229–1234, 2010.
- [97] R. Li and Z. Xiong, “Forecasting stock market with fuzzy neural networks,” *Proceedings of Fourth International Conference on Machine Learning and Cybernetics*, pp. 3475–3479, 2005.
- [98] G. Naadimuthu, D.M. Liu, and E.S. Lee, “Application of an adaptive neural fuzzy inference system to thermal comfort and group technology problems,” *Computer and Mathematics with Applications*, vol. 54, no. 11-12, pp. 1395–1402, 2007.

- [99] T.N. Singh, S. Sinha, and V.K. Singh, "Prediction of thermal conductivity of rock through physico-mechanical properties," *Building and Environment*, vol. 42, no. 1, pp. 146–155, 2007.
- [100] T. Ayata, E. Cam, and O. Yildiz, "Adaptive neuro-fuzzy inference system (ANFIS) application to investigate potential use of natural ventilation in new building designs in turkey," *Energy Conversion and Management*, vol. 48, no. 5, pp. 1472–1479, 2007.
- [101] X. Zhang, Y. Jin, and G. Guo, "ANFIS applied to a ship autopilot design," *Proceedings of fifth International Conference on Machine Learning and Cybernetics*, pp. 2233–2236, 2006.
- [102] A. Abdenmour, "Short-term MPEG-4 video traffic prediction using ANFIS," *International Journal of Network Management*, vol. 15, no. 6, pp. 377–392, 2005.
- [103] S. Abdulhamit, "Application of adaptive neuro-fuzzy inference system for epileptic seizure detection using feature extraction," *Computer in Biology and Medicine*, vol. 37, no. 2, pp. 227–244, 2007.
- [104] U. Caydas, A. Hascalik, and S. Ekici, "An adaptive neuro-fuzzy inference system (ANFIS) model for wire-EDM," *Expert Systems with Applications*, vol. 36, no. 8, pp. 6135–6139, 2009.
- [105] A. Mellit and S.A. Kalogirou, "Neuro-fuzzy based modelling for photovoltaic power supply system," *First International Power and Energy Conference*, pp. 89–93, 2006.
- [106] S. Akbulut, A.S. Hasiloglu, and S. Pamukcu, "Data generation for sheer modulus and damping ratio in reinforced sands using ANFIS," *Soil Dynamics and Earthquake Engineering*, vol. 24, no. 11, pp. 805–814, 2004.
- [107] Y. Lu, F. Yeh, C. Li, and M. Wu, "Study of using ANFIS to the prediction in bore-expanding process," *International Journal of Advanced Manufacturing Technology*, vol. 26, no. 5-6, pp. 544–551, 2005.



- [108] P.C. Nayak, K.P. Sudheer, D.M. Rangan, and K.S. Ramasastri, “A neuro-fuzzy computing technique for modelling hydrological time series,” *Journal of Hydrology*, vol. 291, no. 1-2, pp. 52–66, 2004.
- [109] M. Kakar, H. Nystrom, L. Aarup, T. Nottrup, and D. Olsen, “Respiratory motion prediction by using the adaptive neuro-fuzzy inference system (ANFIS),” *Physics in Medicine and Biology*, vol. 50, no. 19, pp. 4721–4728, 2005.
- [110] L.P. Maguire, B. Roche, T.M. McGinnity, and L.J. McDaid, “Predicting a chaotic time series using a fuzzy neural network,” *Information Sciences*, vol. 112, no. 1-4, pp. 125–136, 1998.
- [111] Y. Yildirim and M. Bayramoglu, “Adaptive neuro-fuzzy based modelling for prediction of air pollution daily leveles in the City of Zonguldak,” *Chemosphere*, vol. 63, no. 9, pp. 1575–1582, 2006.
- [112] P. Mitra, S. Maulik, S.P. Chowdhury, and S. Chowdhury, “ANFIS based automatic voltage regulator with hybrid learning algorithm,” *42nd International Universities Power Engineering Conference*, pp. 397–401, 2007.
- [113] A.R. Sedaghian and J.D. Lavers, “Neuro-fuzzy predictors for the approximate prediction of v-i characteristics of electric arc furnaces,” *Proceedings of 19th International Meeting of the North American Fuzzy Information Proceeding Society*, pp. 183–187, 2000.
- [114] A.R. Sedaghian and J.D. Lavers, “Application of feedforward neuro-fuzzy networks for current prediction in electric arc furnaces,” *Proceedings of International Joint Conference on Neural Networks*, pp. 420–425, 2000.
- [115] A.R. Sedaghian and J.D. Lavers, “Recurrent neuro-fuzzy networks predictor for the prediction of v-i characteristics of electric arc furnaces,” *Proceedings of 9th IEEE International Conference on Fuzzy Systems*, pp. 110–115, 2000.

- [116] ICITE BRE, “Controller efficiency improvement for commercial and industrial gas and oil fired boilers,” *A CRAFT project, Contract JOE-CT98-7010*, 1999-2001.
- [117] CNR1, “Net radiometer manual,” *Campbell Scientific Corp.*, pp. 1–30, 2008.
- [118] S. Jassar, Z. Liao, and L. Zhao, “Adaptive neuro-fuzzy based inferential sensor model for estimating the average air temperature in space heating systems,” *Building and Environment*, vol. 44, no. 8, pp. 1609–1616, 2009.
- [119] M. Kolehmainen, H. Martikainen, and J. Ruuskanen, “Neural networks and periodic components used in air quality forecasting,” *Atmospheric Environment*, vol. 35, no. 5, pp. 815–825, 2001.
- [120] S. Jassar, Z. Liao, L. Zhao, and K.L.R. Ng, “Parameter selection for training process of neuro-fuzzy systems for average air temperature estimation,” *Proceedings of IEEE International Conference on Mechatronics and Automation*, pp. 1218–1222, 2009.
- [121] J. Wang and B. Malakooti, “A feed forward neural network for multiple criteria decision making,” *Computers and Operations Research*, vol. 19, no. 2, pp. 151–167, 1992.
- [122] D. Ballou and H. Pazer, “Modeling data and process quality in multi-input, multi-output information system,” *Management Science*, vol. 31, no. 2, pp. 150–162, 1985.
- [123] Y. Huh, F. Keller, T. Redman, and A. Watkins, “Data quality,” *Information and Software Technology*, vol. 32, no. 8, pp. 559–565, 1990.
- [124] E. Masson and Y. Wang, “Introduction to computation and learning in artificial neural networks,” *European Journal of Operational Research*, vol. 47, no. 1, pp. 1–28, 1990.
- [125] C. Fox, A. Levitin, and T. Redman, “The notion of data and its quality dimensions,” *Information Processing and Management*, vol. 30, no. 1, pp. 9–19, 1994.
- [126] A. Bansal, R. Kauffman, and R. Weitz, “Comparing the modeling performance of regression and neural networks as data quality varies,” *Journal of Management Information Systems*, vol. 10, no. 1, pp. 11–32, 1993.

- [127] D. O’Leary, “The impact of data accuracy on system learning,” *Journal of Management Information Systems*, vol. 9, no. 4, pp. 83–98, 1993.
- [128] B.D. Klein and D.F. Rossin, “Data errors in neural network and linear regression models: an experimental comparison,” *Data Quality*, vol. 5, no. 1, pp. 33–43, 1999.
- [129] B.D. Klein and D.F. Rossin, “Data errors in linear regression models: Effect of errors in test data and errors in training data on predictive accuracy,” *Information Science*, vol. 2, no. 2, pp. 33–43, 1999.
- [130] M. Wei, B. Bai, A. Sung, Q. Liu, J. Wang, and M. Cather, “Predicting injection profiles using ANFIS,” *Information Sciences*, vol. 177, no. 20, pp. 4445–4461, 2007.
- [131] S.K. Ronald and J.J. Legendre, “Discovery of functional and approximate functional dependencies in relational databases,” *Journal of Applied Mathematics and Decision Sciences*, vol. 7, no. 1, pp. 49–59, 2003.
- [132] Y. Huhtala, J. Karkkainen, P. Porkka, and H. Toivonen, “TANE: An efficient algorithm for discovering functional and approximate dependencies,” *The Computer Journal*, vol. 42, no. 2, pp. 100–111, 1999.
- [133] S. Jassar, Z. Liao, and L. Zhao, “Impact of data quality on predictive accuracy of ANFIS based soft sensor models,” *Proceedings of the World Congress on Engineering and Computer Science*, vol. 2, pp. 1001–1006, 2009.
- [134] Canada Mortgage and Housing Corporation, “Optimizing residential forced-air HVAC systems,” *Technical Series*, pp. 96–218, 1996.
- [135] R.J. Meyers, E.D. Williams, and E.D. Mathews, “Scoping the potential of monitoring and control technologies to reduce energy use in homes,” *Energy and Buildings*, vol. 42, no. 5, pp. 563–569, 2009.

- [136] N. Nassif, S. Kajl, and R. Sabourin, “Evolutionary algorithms for multi-objective optimization in HVAC system control strategy,” *Proceedings of North American Fuzzy Information Processing Society, Alberta, Canada*, pp. 51–56, 2004.
- [137] K.F. Fong, V.I. Hanby, and T.T. Chow, “HVAC system optimization for energy management by evolutionary programming,” *Energy and Buildings*, vol. 38, no. 3, pp. 230–231, 2006.
- [138] J. Ingersoll and J. Huang, “Heating energy use management in residential buildings by temperature control,” *Energy and Buildings*, vol. 8, no. 1, pp. 27–35, 1985.
- [139] G. Gao and K. Whitehouse, “The self-programming thermostat: Optimizing setback schedules based on home occupancy patterns,” *First ACM Workshop on Embedded Sensing Systems for Energy -Efficiency in Buildings*, pp. 67–79, 2009.
- [140] A.A. Putnam, F.E.Belles, and J.A.C. Kentfield, “Pulse combustion,” *Progress in Energy and Combustion Science*, vol. 12, no. 1, pp. 43–79, 1986.
- [141] I. Walker and J. Lutz, “Laboratory evaluation of residential blower performance,” *Lawrence Berkeley National University: Berkeley California*, 2005.
- [142] J. Lutz, V. Franco, A. Lekov, and G. Wong-Parodi, “BPM motors in residential gas furnaces: What are the savings?,” *Lawrence Berkeley National University: Berkeley California*, 2006.
- [143] A. Lekov, V. Franco, and J. Lutz, “Residential two-stage gas furnaces: do they save energy?,” *Lawrence Berkeley National University: Berkeley California*, 2006.
- [144] ASHRAE, *ASHRAE Handbook: Fundamentals*, Atlanta, GA: American Society of Heating, Refrigerating, and Air Conditioning Engineers, Inc., 2009.
- [145] C. Lin, C.C. Federspiel, and D. M. Auslander, “Multi-sensor single-actuator control of HVAC systems,” *International Conference for Enhanced Building Operations*, pp. 14–25, 2002.

- [146] P. Oppenheim, “Energy saving potential of a zoned forced air heating system,” *ASHRAE Winter Meeting*, pp. 1247–1257, 2002.
- [147] K.A. Temple, *Field Performance of a Zoned Forced-Air Cooling System in an Energy Efficient House*, U.S. Department of Energy, 2004.
- [148] C. A. Brown, *Multizone Register Controlled Residential Heating: Optimized for Energy Use and Comfort*, M.S. Thesis, University of California, Berkeley, 2007.
- [149] I. Walker, “Register closing effects of forced air heating system performance,” *Lawrence Berkeley National University: Berkeley California*, 2003.
- [150] J. Wills, “Will HVAC control go wireless?,” *ASHRAE Journal*, vol. 46, no. 7, pp. 46–52, 2004.
- [151] W.M. Healy, “Lessons learned in wireless monitoring,” *ASHRAE Journal*, vol. 47, no. 10, pp. 54–58, 2005.
- [152] K. Roth, “Wireless for controls: An update,” *ASHRAE Journal*, vol. 50, no. 8, pp. 76–78, 2008.
- [153] E.L. Vine, “Saving energy the easy way: An analysis of thermostat management,” *Energy*, vol. 11, no. 8, pp. 811–820, 1986.
- [154] B. Sanner, C. Karytsas, D. Mendrinis, and L. Raybach, “Current status of ground source heat pumps and underground thermal energy storage in Europe,” *Geothermics*, vol. 32, no. 4-6, pp. 579–588, 2003.
- [155] J. Lund, B. Sanner, L. Ryback, R. Curtis, and G. Hellstorm, “Geothermal (ground-source) heat pumps - A world overview,” *Geo-heat Center Quaterly Bulletin*, vol. 25, no. 3, pp. 1–10, 2004.
- [156] A.M. Omer, “Ground-source heat pump systems and applications,” *Renewable and Sustainable Energy Reviews*, vol. 12, no. 2, pp. 344–371, 2008.

- [157] Y. Hwang, R. Radermacher, and W. Kopko, “An experimental evaluation of a residential-sized evaporatory cooled condenser,” *International Journal of Refrigeration*, vol. 24, no. 3, pp. 238–249, 2001.
- [158] D. Kolokosta, G. Saridakis, A. Pouliezios, and G. Stavrakakis, “Design and installation of an advanced EIB fuzzy indoor comfort controller using MATLAB,” *Energy and Buildings*, vol. 38, no. 9, pp. 1084–1092, 2006.
- [159] N. Ota, E. Arens, and P. Wright, “Energy efficient residential thermal comfort with wireless sensor networks: A case study for air conditioning in california,” *ASME International Mechanical Engineering Congress and Exposition Conference, Boston, MA*, pp. 43–52, 2008.
- [160] ASHRAE, *ASHRAE Handbook: HVAC Systems and Equipment*, Atlanta, GA: American Society of Heating, Refrigerating, and Air conditioning Engineers, Inc., 2004.
- [161] W. Yang and G. Zhang, “Thermal comfort in naturally ventilated and air-conditioned buildings in humid subtropical climate zone in china,” *Biometeorol*, vol. 52, no. 5, pp. 385–398, 2008.
- [162] A. Redfern, M. Koplou, and P. Wright, “Design architecture for multi-zone HVAC control systems from existing single-zone systems using wireless sensor networks,” *Proceedings of the SPIE - The International Society for Optical Engineering*, vol. 6414, pp. 64140Y.1–64140Y.8, 2007.
- [163] M.M. Ardehali and T.F. Smith, “Evaluation of variable volume and temperature HVAC system for commercial and residential buildings,” *Energy Conversion and Management*, vol. 37, no. 9, pp. 1469–1479, 1996.
- [164] L. Peeters, R. De Dear, J. Hensen, and W. D’haeseleer, “Thermal comfort in residential buildings: Comfort values and scales for building energy simulation,” *Journal of Applied Energy*, vol. 86, no. 5, pp. 772–780, 2009.

# Appendix A

## List of Publications

### *Journal Publications*

1. S. Jassar, Z. Liao, and L. Zhao, “A Recurrent Neuro-Fuzzy System and its Application in Inferential Sensing”, Applied Soft Computing, Accepted, 2010.
2. S. Jassar, Z. Liao and L. Zhao, “Data Quality in Hybrid Neuro-Fuzzy based Soft-Sensor Models: An Experimental Study”, IAENG International Journal of Computer Science, vol. 37, no. 1, 2010, pp. 64-75.
3. S. Jassar, Z. Liao and L. Zhao, “Adaptive Neuro-fuzzy based Inferential Sensor Model for Estimating the Average Air Temperature in Space Heating Systems”, Building and Environment, vol. 44, no. 8, 2009, pp. 1609-1616.

### *Book Chapter*

1. S. Jassar, Z. Liao, and L. Zhao, “Data Quality in ANFIS based Soft Sensors”, Machine Learning and System Engineering, pp 143-155, Springer, ISBN: 978-90-481-9418-6.

### *Conference Publications*

1. S. Jassar, Z. Liao and L. Zhao, “Hybrid Neuro-Fuzzy System Application to Inferential Sensing”, The 5th IEEE Conference on Industrial Electronics and Applications, 2010, pp. 1505-1510.

2. S. Jassar and Z. Liao and L. Zhao and K.L.R. Ng, “Parameter Selection for Training Process of Neuro-fuzzy Systems for Average Air Temperature Estimation”, Proceedings of IEEE International Conference on Mechatronics and Automation, 2009, pp. 1218-1222.
3. S. Jassar, T. Behan, L. Zhao and Z. Liao, “The Comparison of Neural Network and Hybrid Neuro-Fuzzy based Inferential sensor Models for Space Heating Systems”, 2009 IEEE International Conference on Systems, Man and Cybernetics, 2009, pp. 4299-4303.
4. S. Jassar, Z. Liao and L. Zhao, “Impact of Data Quality on Predictive Accuracy of ANFIS based Soft Sensor Models”, International Conference on Soft Computing and Applications, 2009, pp.1001-1006.
5. Ng, K.L. Ringo, Wong, Edmund S.W., Jassar S., Richman, R.C. and Liao Z., “Evaluating the Potentials of Variable Air Volume Systems in North American Houses”, 10th Asia pacific Conference on the Built Environment - Green Energy for Environment, 2009, pp. 29.
6. S.Jassar, Z.Liao, L.Zhao, “The Potential for Improving Space Heating Energy Efficiency & Reducing CO2 Emissions”, 14th International Conference on Environment, 30th June-3rd July, 2008, Edmonton, Alberta, Canada.

***Articles Submitted to refereed journals***

1. Jassar, S., Liao, Z. and Zhao, L., “On the training of Hybrid Neuro-Fuzzy models for Soft Sensors”, IEEE Transactions on Systems, Man and Cybernetics, under review, 2010.



# Appendix B

## Nomenclature

$x$	input to hybrid neuro-fuzzy model
$O$	output of different nodes in a hybrid neuro-fuzzy model
$y$	output of hybrid neuro-fuzzy model
$e$	error vector
$\theta$	unknown parameter to be optimized
$H$	height of the mountain function
$D$	density measure of data points
$v$	cluster center
$J$	cost function
$P$	binary membership matrix
$\omega_i$	firing strength or degree of certainty of $i$ th rule
$\bar{\omega}_i$	normalized firing strength of $i$ th rule
$a$	parameter determines the width of the generalized bell-shaped curve
$b$	parameter determines the width of the generalized bell-shaped curve
$c$	parameter determines the center of the generalized bell-shaped curve

$A_1, A_2, B, B_1, B_2$	linguistic values
$R_i$	$i$ th rule
$C_i$	output of $i$ th rule
$T_{avg}$	measured average air temperature in the building ( $^{\circ}C$ )
$\hat{T}_{avg}$	estimated average air temperature in the building ( $^{\circ}C$ )
$T_0$	exterior temperature ( $^{\circ}C$ )
$T_d$	desired temperature ( $^{\circ}C$ )
$Q_{sol}$	solar radiation ( $W$ )
$Q_{in}$	energy consumption by the heating plant ( $W$ )
$N$	order of the dynamic model
$E$	overall error function for the training algorithm
$m$	mean for Gaussian membership function
$\sigma$	variance for Gaussian membership function
$\Delta\tau$	sampling interval
$N_e$	total number of energy consumption samples
$K$	estimated data point
$L$	measured data point
$T$	total number of data points
$M$	number of training data points
$R^2$	coefficient of determination
$\bar{L}$	average of measured data points

$SS$	step size
$SSINC$	step size increase rate
$SSDEC$	step size decrease rate
$d$	unmodified variable
$\acute{d}$	modified variable after adding error signal
$TD1$	weekend only training data set
$TD2$	week day only training data set
$TD3$	weekend and week day training data set
$s$	noise signal strength
$T\acute{D}1$	weekend only training data set after adding noise signal
$T\acute{D}2$	week day only training data set after adding noise signal
$T\acute{D}3$	weekend and week day training data set after adding noise signal
$HDH$	heating degree hours ( $^{\circ}C.Hr$ )

# Appendix C

## Abbreviations

AFUE	Annual Fuel Utilization Efficiency
ANFIS	Adaptive Neuro-Fuzzy Inference System
ANN	Artificial Neural Network
ASHRAE	American Society of Heating, Refrigerating and Air-Conditioning Engineers
BP	Back Propagation
BPM	Brushless Permanent Magnet
CAV	Constant Air Volume
DOE	Department of Energy
DRETD	Demand Response Enabling Technology
EPS	Expanded Polystyrene Insulation
FCM	Fuzzy C-Means
FIS	Fuzzy Inference System
FLBC	Fuzzy-logic based Control
FLS	Fuzzy Logic System
GHG	Greenhouse Gases

HVAC	Heating, Ventilating and Air-Conditioning
IPCC	Intergovernmental Panel on Climate Change
LS	Least Squares
MF	Membership Function
MTE	Minimum Testing Error
PPD	Predicted Percentage Dissatisfied
PSC	Permanent Split Capacitor
PVC	Poly vinyl chloride
RMSE	Root Mean Square Error
RenFIS	Recurrent Neuro-Fuzzy Inference System
RNN	Recurrent Neural Network
SSE	Steady State Efficiency
VAV	Variable Air Volume
WGN	White Gaussian Noise

# Appendix D

## Survey Questionnaire

### *Entry Section*

1. Is this dwelling owned by a member of this household?

Ans. Yes / No / Don't know

2. Number of occupants.

3. On an average weekday is there someone at home all day?

4. Household total income per annum (participants can choose not to answer this question).

Ans. CAD 30,000-50,000 / CAD 50,000-70,000 / CAD 70,000-100,000 / CAD 100,000+

5. Annual consumption of gas? ( $m^3$ )

6. Annual expenditure on gas? (CAD)

7. Annual consumption of electricity? (kwh)

8. Annual expenditure on electricity? (CAD)

9. Type of the dwelling.

Ans. Single detached / Double or Row houses / Low-rise apartments / Mobile houses

10. Year of construction of the dwelling?
11. Number of stories?
12. Heated area of the dwelling (total floor space of a dwelling excluding garage)?

***Thermal Satisfaction***

13. How satisfied are you with room temperature in winter?

Ans. Very dissatisfied / Dissatisfied / Neutral / Satisfied / Very satisfied

14. What do you tend to do when you feel thermally uncomfortable (cold) in winter?  
(Please circle the most predominant action).

Ans. Put more clothes on / Adjust thermostat / Use supplementary heating equipment  
/ No action / Tolerates / Gets in contact with person in charge (e.g. home owner) /  
Do exercise

15. How well do you feel if you can personally control room temperature in winter?

Ans. Very badly / Badly / Neutral / Well / Very well

16. During winter time, do you experience overheating in certain parts of your dwelling?

Ans. Several times a day / Daily / Weekly / Monthly / Less frequently / Never

17. When experiencing overheating, how often would you then open the windows for better air circulation?

Ans. Several times a day / Daily / Weekly / Monthly / Less frequently / Never

18. During summer time, do you experience overcooling in certain locations of your dwelling?

If you don't have air conditioner at home, you can leave the next 3 questions blank.

Ans. Yes / No / Do not have an air conditioner

19. If your answer to the previous question is “Yes”, how often would you then open the windows for better air circulation?

Ans. Several times a day / Daily / Weekly / Monthly / Less frequently / Never

20. During the summer when the air conditioner is turned on, would you open the windows for more air circulation or a fresh breeze?

Ans. Yes / No

21. If your answer to the previous question is “Yes”, how frequently would you open the windows for more air circulation?

Ans. Several times a day / Daily / Weekly / Monthly / Less frequently / Never

### ***Heating***

22. What type of heating equipment provides most of the heat for the dwelling?

Ans. Furnace with forced air / Hydronic System / Electric baseboards / Heating stove / Electric radiant heating / Others - please specify / Do not know

23. Average age of the heating system?

Ans. 3 years old or less / 4-5 years / 6-10 years / 11-15 years / 16-20 years / 21-25 years / 26 or more years / Not Sure

24. Which months the heating system is operated for?

25. In addition to furnace or the main heating system, did your household use any other supplementary heating equipment?

26. What type of supplementary heating equipment did you use most often?

Ans. Electric baseboards / Portable electric heater / Wood stove / Furnace / Other - specify



27. How many programmable thermostats do you have in your dwelling?
28. Set-point temperature in winter - morning (e.g. 6a.m. to 10a.m.).
29. Set-point temperature in winter - during the day or not at home (e.g. 10a.m. to 5p.m.).
30. Set-point temperature in winter - after work / school or at home (e.g. 5p.m. to 11p.m.).
31. Set-point temperature in winter - bedtime (e.g. 11p.m. to 6a.m.).
32. Setback when not at home?
33. Setback during sleeping hours?
34. How often do you use your thermostat to regulate temperature in your dwelling in winter?
- Ans. Several times a day / Daily / Weekly / Monthly / Less frequently / Never
35. When was the last time your heating system is checked and maintained by a certified technician?
36. How often do you check or maintain your heating system by a certified technician?
37. Time spent in each part of the house?
- a) Kitchen
  - b) Living Room
  - c) Bedroom
  - d) Dining Room
  - e) Basement (if any)

### ***Window Opening Behavior***

38. On an average weekday, would you open your window blinds or shades and let the sun shine in for illumination when you are at home?
39. If your answer to the previous question is “Yes”, how much time on average do you rely on the natural sunlight for illumination? If you answer “No”, you can skip to the next question.
40. What is the most predominant usage of your windows? Rank from 1 to 3, with 1 being the most predominant.

Ans. Ventilation / Visual connection to the outside / Natural day-lighting