System Design and Dynamic Signature Identification for Intelligent Energy Management in Residential Buildings

by

Jaehwi Jang

B.S (Seoul National University, Korea) 2001

M.S (University of California, Berkeley) 2003

A dissertation submitted in partial satisfaction

of the requirements for the degree of

Doctor of Philosophy

in

Engineering - Mechanical Engineering

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, BERKELEY

Committee in charge:

Professor David M. Auslander, Chair

Professor Paul K. Wright

Professor Edward Arens

Fall 2008

The dissertation of Jaehwi Jang is approved.

Chair Date

Date

Date

University of California, Berkeley

Fall 2008

System Design and House Dynamic Signature Identification for Intelligent Energy Management in Residential Buildings

by

Jaehwi Jang

PhD Dissertation, Dept. of Engineering, University of California, Berkeley.

To my parents, wife, and son

Abstract

System Design and Dynamic Signature Identification for Intelligent Energy Management in Residential Buildings

by

Jaehwi Jang

Doctor of Philosophy in Engineering - Mechanical Engineering

University of California, Berkeley

Professor David M. Auslander, Chair

Increasing energy demand from residential buildings and evolving utility pricing policy to regulate energy use during peak times require a new paradigm for energy management in residential buildings. As a prototype for intelligent energy management systems of residential buildings, DREAM (Demand Responsive Electrical Appliance Manager), based on a wireless sensor network, was developed. This autonomous system consisting of wireless sensors and actuators, a graphical user interface, and a main control reduces peak electrical demand and ultimately optimizes energy management by identifying house dynamic signature as well as occupant thermal preference and patterns. In summer 2007, functionality and overall performance were evaluated with two field tests and showed promise for the DREAM system.

Due to significance of the house dynamic signature learning in an intelligent energy management system, three approaches were studied. Despite the simplicity of the model and success in identifying thermal characteristics of a house, the 1st order differential equation method, which considered thermal influences of five heat sources, showed limitations in representing actual temperature behavior delicately. The tabular method was suggested to capture house nonlinear behavior by learning temperature change rate with respect to

different events and periods. The prediction using the tabular method followed the actual measured temperature within a tolerable error range, except for a relatively long heater-on event. The last method, the ARX model fitting method, provided the best prediction result, but the performance was considerably influenced by the choice of sample data for parameter learning.

The multiple-model switching algorithm was proposed to minimize performance inconsistency in the ARX model fitting method. Instead of sticking to one model, it allows several candidates whose parameters are calculated from seven consecutive days, and selects one (multiple-model hard switching [MMHS]) or fuses all (multiple-model soft switching [MMSS]). Depending on the criterion to select or weight a candidate, the algorithm is divided into proximity-based model switching and applicability-based model switching. Overall, the MMSS showed better performance than the MMHS and, most of all, the applicability-based MMSS algorithm dramatically improved the prediction quality when anomalies in data were properly filtered.

All algorithms in this study were evaluated with the real data that were collected from more than 20 occupied houses in Northern California, Minnesota, and South Australia.

Professor David M. Auslander
Dissertation Committee Chair

Contents

List of Tables

List of Figures

Acknowledgments

This dissertation could not have been written without my research advisor, Prof. David Auslander, who not only encouraged me to see objects from various points of view, but also continuously showed what a good engineer should be. I would also like to express my deepest gratitude to him for his excellent guidance and support. I would also like to thank Prof. Edward Arens and Prof. Paul Wright for serving on my dissertation committee and for their valuable comments. I really admire their endless passion and humor as well as academic achievement.

This dissertation is based on the DRETD (Demand Response Enabling Technology Development) project I was involved for the last 5 years. I would thank the PIER (Public Interest Energy Research) program of California Energy Commission for providing me an opportunity to work with them and financial support.

In addition, I would like to acknowledge EcoFactor (John Steinberg, Scott Hublou, and Leo Cheung) for their financial support during my last semester and valuable house data that were used for most algorithm development and validation in my dissertation. They always simulated the inspiration on my research questions.

I thank my research partners, Xue Chen and Therese Peffer for their continuous feedback and great collaboration. I also thank Bill Burke, Massieh Najafi, Jonghak Kim, Shan Tang, and other members in Mechatronics Lab for making my campus life more interesting and all the people who helped me to reach this point.

Finally I am deeply indebted to my parents, Jintae Jang and Younghee Lee as well as my wife, Ahra Cho, and my son, Minsuh Jang, for their constant encouragement and wholehearted support.

Chapter 1

Introduction

1.1 Motivation

There has been an increasing interest in building control. This increasing interest can be considered natural in that a building is where humans spend most of their daily lives. As a building got equipped with heavy energy-consuming appliances such as air-conditioners, refrigerators, and water heaters, it became an object that should be under control. According to the Quarterly Fuel and Energy Report (QFER) data from California Energy Commission, 2006, electricity use by residential and commercial buildings accounted for 68% of total electricity use in California in 2005. Growing energy demand is pushing people to find a new methodology that requires less energy consumption without sacrificing comfort in the building.

Since the concept of intelligent buildings was introduced in the early 1980s, many commercial buildings have been labeled as intelligent buildings. The first generation intelligent buildings consisted of one or more independent intelligent devices such as a lighting control system, an HVAC system, and a security system. Since the individual system was controlled by its own controller, information flow between systems was fundamentally blocked. The second generation combined the disaggregated control systems by sharing information,

which allowed high-level control strategies. The third generation of intelligent buildings includes self-regulating capabilities that adapt to dynamic changes of human requirement and environment [1, 2, 3].

Although energy use by residential buildings is significantly increasing, most research in the area of building control has been concentrated on commercial buildings. The strong inclination to commercial buildings is rooted in relative effectiveness per system by a smart controller as well as by significant energy policy. A commercial building is a more convenient object than residential building from a control point of view since it has less uncertainty in occupant behavior and smaller variation in thermal properties over time. A time-of-day dependent utility pricing rate also stimulates building scientists' and engineers' interests in the intelligent building control for energy-saving or cost-saving.

The State of California began to recognize that peak electricity use mainly concentrated in hot summer periods can be redistributed by applying a dynamic pricing rate to residential buildings. In order to respond to the new pricing policy, a new energy management system for a residential building is inevitable. Accordingly, the thermostat control group led by Prof. David Auslander in Mechanical Engineering and Prof. Edward Arens in Architecture of University of California, Berkeley has developed a prototype of a smart energy management system specifically for residential buildings. This system, DREAM (Demand Responsive Electrical Appliance Manager), maximizes functional efficiency by disaggregating the elements of a conventional thermostat as well as intensifying autonomous control. This system is based on a new paradigm and shows promise in efficient energy management for residential buildings.

One of the fundamental tasks in autonomous building control must be system identification (or dynamic signature learning). A successful identification process allows a controller to predict house thermal behavior within a tolerable error and ultimately make an optimal decision from the prediction. It is challenging to accurately identify a house dynamic signature given limited data. Different approaches were tried to provide better prediction

quality. Both a tabular method and a numerical method based on physical models were initially suggested. An Interacting Multiple Model (IMM) method whose concept originated from a target tracking process was also introduced. Since these learning algorithms should be ready to be used in the controller of an autonomous energy management system, computational cost and algorithmic complexity should be considered.

Different from most previous studies, which dealt with simulation data, this study was based on the data from more than 25 real houses in various locations. While the real data contained significant amounts of noise and anomalies and made data analysis arduous, the uncertainty and diversity, which may never occur with simulated data, resulted in deeper analysis in various directions.

1.2 Overview

Chapter 2 covers the overall procedure in developing a new prototype of an intelligent energy management system for residential buildings. Increasing energy demand from residential buildings in California was analyzed with the actual electricity consumption data from the California Energy Commission. As a motivation to develop a demand responsive energy management system (DREAM) dynamic utility price rating is also introduced.

The first half of this chapter describes the hierarchical layered structure of DREAM and its infrastructure. Auxiliary elements such as the remote database and its interface, engineering GUI, and HVAC switch are briefly explained. The rest of this chapter is focused on optimization and control processes that the controller mainly performs to balance the energy demand and user comfort in given conditions. The performance of the DREAM controller was preliminarily evaluated through in-lab simulation, and the results are reported. The whole system was deployed in two occupied houses in Northern California during summer 2007, and the field test results are also discussed at the end of this chapter.

More technical approaches to identify the house dynamic signature (or house thermal

properties) are introduced in Chapter 3. As a background for system modeling and data analysis, general thermal properties of buildings and modeling for building energy are summarized. In order to develop identification algorithms and evaluate their performance, three categories of houses were used. These houses include four simulation houses from MZEST (Multi-Zone Energy Simulation Tool), two California houses in which the DREAM system was deployed, and 22 houses in Minnesota and Australia.

Three different algorithms, the 1st order physical model method, tabular method, and 4th order ARX (Auto-Regressive with Exogenous Input) model method, are suggested. The basic idea and the detailed description on each algorithm are explained in Chapter 3. The performance of three methods was analyzed in various ways. The drawbacks as well as the advantages of each algorithm are analyzed by comparing the best and worst results. The limitations of each method are also discussed at the end of each subsection.

In order to improve the ARX model method in Chapter 3, a multiple-model switching technique was applied. The multiple-model switching algorithm suggests one methodology to avoid poor prediction, which might be caused by parameter calculation based on biased or atypical data. As a criterion for selecting or weighting a model, proximity and applicability were separately used. Each method used two model-switching algorithms, hard-switching and soft-switching. The prediction results are compared with the 24-hour prediction results.

Chapter 5 discusses the constraints in the system and algorithm development and briefly summarizes the findings from each chapter.

Chapter 2

Demand Response Enabled Thermostat Development for Residential Buildings

2.1 Backgrounds

Every year in the United States after 1999, blackouts due to high demand on electricity have been reported. Most electricity shortage is caused by either the curtailment of supply, such as failures in the power generation or transmission systems and power market manipulation, or high usage during a peak time. In order to prevent frequent power shortage, intuitive remedies (for example, increasing supply or decreasing demand) may be suggested. Constructing new power plants is the simplest way to prevent electricity shortage. However, electricity suppliers are reluctant to construct power plants because electricity shortages occur mostly during hot peak periods, which are a relatively small portion of the year.

Alternatively, decreasing electricity demand can prevent electricity shortage as well. According to Tab.2.1, electrical energy consumption in California has increased by 62% since 1981. Although industrial electrical energy use in California is up to 16.3% of total consumption, this portion cannot be regulated merely to reduce energy demand. Therefore, electricity use from residential and commercial buildings, which account for more than 68%

of total consumption, would be a target for electrical energy demand control. Currently many efforts have been made to reduce energy use in commercial buildings. A central control unit that controls an HVAC system of an individual commercial building monitors energy use and regulates peak demand by applying various methods. Unfortunately, the same efforts have not been made for residential buildings due to difficulty in developing an energy management system.

Residential buildings are still major electricity consumers (31.0% of total consumption in 2005), and the proportion is increasing (jpredicted to be 34.4% by 2018). These statistics imply that more weight needs to be given to new energy-related policies and programs for residential buildings. Since the energy use of a single residential building is much smaller than that of a single commercial building, many studies have targeted energy management of commercial buildings [4]. Energy simulation tools such as EnergyPlus and California Non-residential Engine (CNE) were developed for commercial buildings only. The big difference between energy management systems of commercial buildings and residential buildings in control point of view is autonomy. To manage energy use of a commercial building, physical and thermal characteristics of the buildings would be analyzed. For analysis, complicated techniques or methods might be applied by a system engineer. Based on the analysis, the system can predict the thermal behavior of the building and determine temperature setpoint. Even when the system fails, the engineer would fix the problem. In addition, system operation is regular since the behavior of occupants in the commercial buildings is predictable. (For example, employees come to the building at 8am and leave at 5pm from Monday to Friday.) In contrast to commercial buildings, the energy management system for residential buildings should operate autonomously. There is no engineer who can analyze the characteristics of the house as well as fix the arising systematic problems. The patterns of occupants are relatively random. (For example, occupants may leave home for vacation or go to the office late due to unexpected events.) Temperature preference can vary every season. The uncertainty in residential buildings has been an obstacle for active

Year	Residential	Commercial	Industrial	Mining	Agriculture	TCU*	Street Lighting	Total Consumption
1980	52,082	47,600	40,771	4,104	13,737	7,956	1,685	167,935
1981	53,495	50,419	41,350	4,387	16,402	8,260	1,643	175,957
1982	52,574	50,297	37,784	6,061	14,507	8,759	1,706	171,688
1983	54,577	52,023	38,624	6,322	11,610	9,135	1,604	173,896
1984	57,564	55,092	40,411	6,978	15,320	9,766	1,535	186,666
1985	58,528	56,908	41,496	7,329	17,453	10,423	1,537	193,673
1986	58,452	59,306	42,232	6,514	15,940	10,084	1,512	194,038
1987	61,267	62,944	44,182	6,463	16,409	11,058	1,536	203,865
1988	64,033	65,958	46,421	6,535	17,995	11,465	1,494	213,902
1989	65,316	68,932	46,942	6,719	19,225	12,087	1,507	220,728
1990	67,667	72,753	47,384	6,786	20,774	12,430	1,580	229,375
1991	67,142	72,540	46,004	6,835	16,266	12,640	1,614	223,040
1992	69,225	76,018	45,928	6,600	15,471	12,967	1,652	227,862
1993	68,424	76,604	45,532	6,262	15,902	13,059	1,648	227,431
1994	69,774	76,687	45,388	6,160	16,948	12,842	1,649	229,448
1995	69,770	78,409	46,834	6,148	14,301	13,238	1,624	230,323
1996	72,164	80,709	47,207	6,202	16,874	13,293	1,660	238,108
1997	73,547	84,442	48,847	6,174	17,514	13,914	1,701	246,140
1998	75,387	86,330	47,294	5,794	13,485	13,608	1,758	243,657
1999	76,482	89,466	48,695	5,233	17,097	13,921	1,658	252,552
2000	80,612	95,148	49,801	5,713	17,532	14,486	1,730	265,021
2001	75,916	90,130	44,693	5,786	18,921	13,068	1,727	250,241
2002	77,731	92,978	45,557	5,724	21,057	13,152	1,715	257,914
2003	82,196	96,816	43,353	5,946	20,274	13,132	1,751	263,468
2004	83,774	99,467	43,617	6,446	22,246	13,589	1,787	270,927
2005	84,527	101,393	44,586	6,559	19,502	14,014	1,804	272,385

Source: QFER, California Energy Commission. September 2006.
* Transportation, Communication & Utility

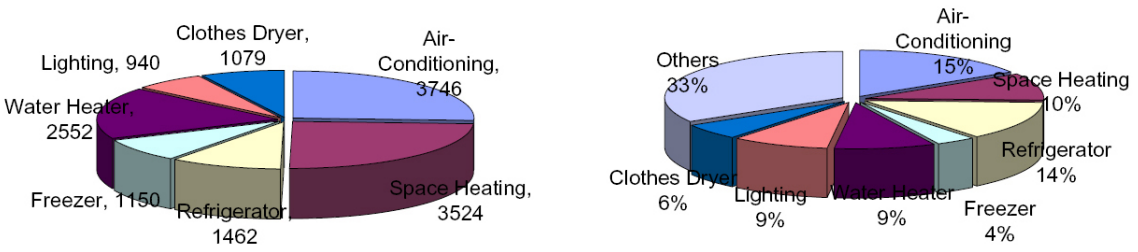
Table 2.1: California Electricity Consumption by Sector (million kWh)

studies on them. Fortunately, the California Energy Commission (CEC) began to recognize the importance of energy management in residential buildings and began to suggest new policies and programs.

One of the programs that the State of California has been emphasizing is demand response. The term *demand response* covers a variety of programs, including traditional direct control (interruptible) programs and new price-responsive demand programs. A key distinction is whether the program is dispatchable. Dispatchable programs, such as direct control, interruptible tariffs, or demand bidding programs, have triggering conditions that cannot be anticipated by the customer. Nondispatchable programs are not activated using a predetermined threshold condition, and allow the customer to make the economic choice whether to modify its usage in response to ongoing price signals [5].

In the price-responsive demand program, a time-of-use (TOU) price rate plan that the State of California has provided to reduce the demand for electricity during peak periods would be replaced by a dynamic price rate plan. In time-of-use price rate, electricity price varies on the pre-specified time periods in which the energy is consumed. In this structure, higher prices are charged during utility peak-load times, which can provide an incentive for consumers to curb power use during peak times. However, fixed periods and prices of the time-of-use plan are not flexible enough to regulate high electricity demand during a peak period or emergencies such as a power breakdown. However, dynamic price rates differ from the time-of-use (TOU) price rates in that price and period are not pre-specified. Instead, they would be determined by real-time electricity supply condition.

In order to meet the needs of a new program, an electrical energy management system for residential buildings has been developed with a new paradigm. In the system development, the specificity of residential buildings had to be considered, and the system had to properly respond to the dynamic price rate without degrading the consumer's life. Fig.2.1 summarizes annual electricity consumption by various home appliances. More than 40% of electricity in a residential house is consumed by air-conditioners, heaters, and refrigerators.



(a) Anual Consumption per Household [kWh (b) Total Anual Electricity Consumption

Sources: Energy Information Administration, Office of Energy Markets and End Use, Forms EIA-457A-C, E, and H of the 1997 RECS; Energy Information Administration, Office of Energy Markets and End Use, 1993 and 2001 Residential Energy Consumption Surveys.

Figure 2.1: Residential Consumption of Electricity by End Use, 2001

A residential thermostat, therefore, would be an adequate start point as a hub for energy management in residential buildings since major electricity consuming devices are easily controlled by it and air-conditioning mainly causes a peak demand in hot summer periods.

2.2 System Structure and Control Hierarchy

A multidisciplinary research group at University of California, Berkeley that had strong interest in developing infrastructure as well as providing a new paradigm for demand responsive systems started their research in June 2003. Four sub-groups determined their own research topics: 1.power metering (Prof. Richard White), 2.energy scavenging (Prof. Paul Wright), 3.communication (Prof. Paul Wright and Prof. Jan Rabaey), and 4.control (Prof. David Auslander and Prof. Edward Arens). Their outcomes are contributing to demand response enabling technology development. The power metering group is developing power metering infrastructure, and the energy scavenging group is finding a new energy scavenging methodology. The communication group is working on the infrastructure of wireless communications to increase reliability of wireless communication inside a building. The control group is developing controls that can provide autonomous energy management to residential buildings. Although the research area of each group seems to be independent, the work is related in that constraints or properties of one group's infrastructure are often

determined by another group's infrastructure.

In this section, the outcomes from the control group are described. As a control hub that enables residential buildings to be responsive to a dynamic price, a low-cost energy management system, DREAM (Demand Responsive Electrical Appliance Manager), has been developed. The DREAM system based on a wireless sensor network basically is able to do what a conventional residential thermostat, such as an HVAC system operation control, does. However, the main control that can be characterized by system identification and learning algorithms provides more autonomy to the system, which would ultimately increase the efficiency of energy management for residential buildings.

2.2.1 Overall Structure of DREAM (Demand Responsive Electrical Appliance Manager)

In order to aggressively respond to dynamic utility price, the system should have authority over access to the information on real-time electricity use and control of electrical appliances, including HVAC equipment. This implies that the system requires both a main control unit that can make a decision regarding optimization between electricity use and user comfort and infrastructure that supports data acquisition and communication. Subjects of the communication involve sensors, a main control unit, occupants, and possibly an electricity supplier. As mentioned, design of the DREAM started from conventional thermostats. However, it increased functionality of a controller, a thermometer, and power switches, compatibility, and extensibility by disaggregating components of conventional thermostats. The disaggregation has been achieved mainly by wireless sensor network techniques.

Fig.2.2 represents a schematic of the DREAM in a typical residential building. Basic information such as temperature and humidity are measured both inside and outside the house. Occupancy status also can be obtained by motion sensors. These sensors make it possible to specify a control area, a *target zone*, and apply different strategies in different

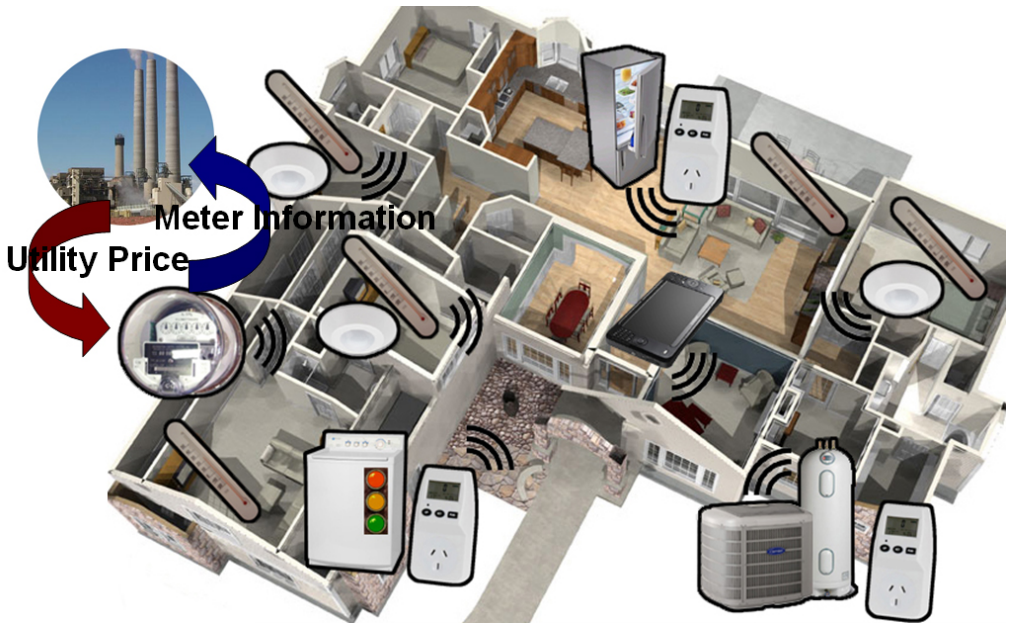


Figure 2.2: Schematic of the Demand Response Electrical Appliance Manager

situations. Electricity use by electrical appliances is monitored by a power meter that measures each individual electrical device as well as entire electrical consumption. An electricity supplier (utility company) would collect the meter information of each house and provide a dynamic utility price rate in advance by two-way communication. The control unit, a small computer in Fig.2.2 combines all the available information to optimize cost and comfort.

The DREAM system has been designed and developed based on the following protocol and functionality.

- ◇ Utility price information from the utility is transmitted to a household. In a normal situation, this price information is assumed to be available approximately 24 hours ahead. A sudden price change may occur due to an unexpected event such as a natural disaster or stoppage of a power station.

- ◇ The main controller receives data from the wireless sensors and controls electrical appliances such as the air-conditioner and water heater via wireless actuators. An additional sensor or actuator can be put into an existing system without any structural

constraint.

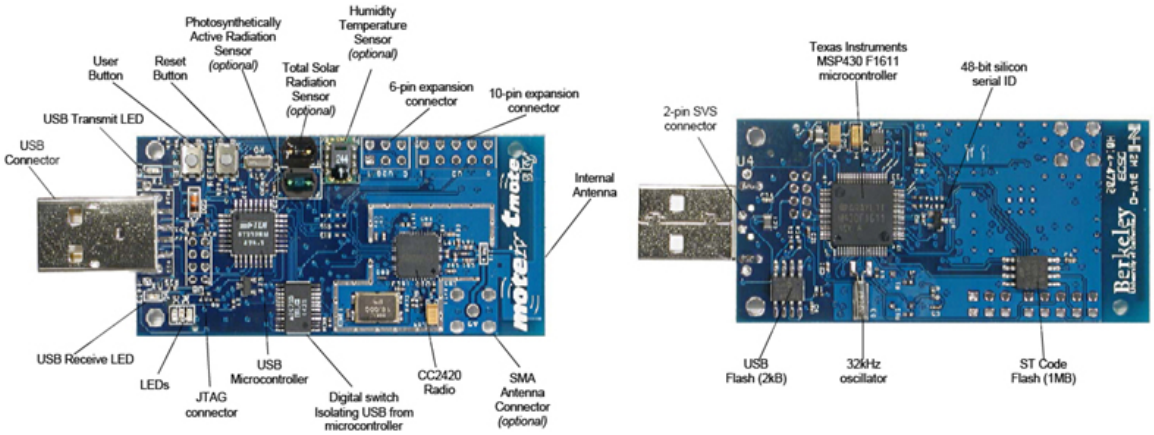
- ◇ The existing thermostat should be easily replaced by the main controller by simple wiring.
- ◇ A graphical interface helps an occupant interact with the system. Energy-related information such as utility price, individual/total electrical usage, and real-time notice or suggestion is displayed on the screen.
- ◇ Electrical interval consumption data is relayed back to the utility.

Wireless Communication

The more information regarding temperature distribution inside the house, outdoor weather conditions, occupancy, and appliance power use is available, the more optimized decision-making can be achieved by the system. The system design, therefore, is driven by the need for distributed sensing and actuating with low cost. While wired data transfer is reliable, the cost of wiring can be the deciding factor especially in large size or multi-story houses. Installing an additional sensor or changing the sensor location would also be accompanied with time-consuming wiring tasks.

To avoid complexity of sensor installation and provide flexibility in total sensor number and location, a computing concept, *motes* (literally means a small particle) have been used for the system. The core of a mote is a small, low-cost, low-power computer. The computer monitors one or more sensors that may detect temperature, light, sound, position, acceleration, vibration, stress, weight, pressure, or humidity depending on mote applications. The computer connects to the outside world with a radio link. The most common radio links allow a mote to transmit at a distance approximately from 10 to 200 feet. Since a fundamental concept with motes is tiny size (and associated tiny cost), small and low-power radios are normal.

The DREAM system found its sensor networks based on Tmote Sky shown in Fig.2.3, a



Source: http://www.sentilla.com/pdf/eol/tmote-sky-datasheet.pdf

Figure 2.3: Tmote Sky from Moteiv

mote platform for extremely low power, high data-rate sensor network applications. It has integrated sensors, radio, antenna, microcontroller, and programming capabilities. The ultra low power microcontroller on the circuit makes the low power operation possible. This 16-bit RISC processor features extremely low active and sleep current consumption. In order to minimize power consumption, it is in sleep mode the majority of the time, wakes up as quickly as possible to process, then returns to sleep mode again. It uses a USB controller from FTDI to communicate with the host computer and features the Chipcon CC2420 radio, which is an IEEE 802.15.4 compliant radio providing reliable wireless communication, for wireless communications. The radio provides fast data rate and robust signal[1].

Even though the radio transmitter and microprocessor in each node requires a very small amount of power, signal quality between two nodes should be high enough to guarantee robust operation of the controller and electrical appliances. To minimize signal overlapping and dropping, a communication protocol specified for the DREAM system was introduced by modifying TinyOS, which is an operating system for motes. Each sensing node is programmed to spread signals with specified frequency while actuating nodes catch signals during wake-up periods. The protocol defines how often and how long the signal should

[1]Specification of the Tmote Sky can be found from http://www.sentilla.com/pdf/eol/tmote-sky-datasheet.pdf

be sent to the actuating node as well as which signal is valid or should be ignored. The base station, a central node directly connected to the controller through USB, receives and sends signals to and from other nodes.

Due to a relatively short communication range of motes (the covering range can considerably shrink when radio is blocked by walls), a mesh network is proposed for the DREAM system. Under the condition that motes are well distributed around a target building, the covering range is theoretically infinite. However, deploying in mesh architecture may be expensive since all motes should be awake to transmit routing information without problems. To minimize information redundancy and power use, the system is currently configured to communicate in a star network with message forwarding used in the repeater mote. Depending on the building size and structure, the total number of the repeater motes increases. In an extreme case, the star network with the repeater motes can be shaped into a partial mesh network. For most typical residential houses, the star network with a single repeater mote usually is sufficient.

Sensors, Actuators, and Price Indicators

The Tmote Sky supports up to six analog and four digital inputs. An auxiliary circuit board with four mono audio jacks that provide power to a sensor, and connection straight to an ADC was designed and mounted on the mote. This mote, referred to as a *generic mote*, is not plug and play. The power and signal pins have to be set according to the specific sensor or actuator that is connected to the input jacks. Currently three different types of sensors (temperature sensor, solar radiation sensor, and motion sensor) are available for the generic mote. Fig.2.4 shows the generic mote and occupancy switch. The temperature sensor measures air temperature shielded from radiation and globe temperature using RTD (resistance temperature detector) and the solar radiation sensor measures global radiation. The motion sensor module is comprised of a Panasonic passive infrared motion sensor (AMN41121) and a 3V watch battery (CR2032). The sensor detects changes in infrared

(a) Generic Mote and Sensors (b) Occupancy Switch

Figure 2.4: Generic Mote (Temperature, Radiation, and Motion) and Occupancy Switch

radiation that occur when there is movement by a person or object whose temperature differs from the surroundings [6].

Two additional sensors that have their own motes instead of using the generic mote were also developed. A CT (Current Transformer) sensor is developed to measure real-time power usage by an individual device as well as a whole house. Tab.2.2 describes characteristics of individual sensing information.

Main Control Unit

A main control unit or a base station is a brain of the DREAM system. It consists of a base mote, a computing platform, and control algorithms. It basically supervises the whole wireless sensor network of the system and optimizes cost from electricity consumption and user comfort by applying various strategies. Since all data are processed through this unit, the control process should be well organized to prevent system failure.

The base mote is essentially the same as other motes used for sensing and actuating. The only difference is that it is particularly programmed to send a command to other motes

Group	Sensor	Sensing Object	Location
Weather Station	anemometer wind vane pyranometers	wind speed wind direction total horizontal radiation / direct normal radiation	on the roof
Outdoor Mote	on-board RH sensor air temperature sensor battery voltage sensor	relative humidity of outside air temperature of outside 2-1.5V battery voltage residual	under the eave of the roof
Generic Mote	air temperature sensor globe temperature sensor IR temperature sensor battery voltage sensor	air temperature globe temperature motion of occupants 2-1.5V battery voltage residual	wall / inside AC / near air vent
Power Mote	current / volage sensor	current in two main circuits / voltage in the main circuit breaker panel	inside circuit breaker

Table 2.2: Sensing Information

and receive signals from sensing motes. Since it is powered through a USB connection, it can respond to input data immediately.

The computing platform should have a CPU and memory for data processing. For a user to monitor system information and interact with the system, an output terminal should also be required. During system development, a personal computer with a monitor may be a reasonable computing platform. A tablet PC or a mobile PC was used for the actual system deployment in the test houses.

Control algorithms contain both statements to make the system follow pre-defined procedure and mathematical calculations for data process and decision making. The DREAM control unit uses Java, one of high level programming languages, to realize control algorithms. Java is known to be inadequate for real-time control in most cases due to its relatively slow response. Fortunately, the DREAM system does not require millisecond accuracy, and this level of accuracy does not necessarily improve the performance of the system at all. The fact that Java is a unique platform-independent language and exuberant libraries for wireless communication is crucial for the wireless network based system.

Figure 2.5: Occupant Interface of the DREAM

Graphical User Interface

The basic occupant interface design (the main user interface of DREAM) was developed in 2005 by Therese Peffer et al [7]. On the left side, which is modeled after the Honeywell Round thermostat to provide better readability, the basic information (temperature, humidity, and HVAC status) and thermostat control switches (mode and HVAC control) are shown. More interaction between an occupant and DREAM is achieved through the right side tabbed panels. They not only show cost information, electrical usage, and program schedule, but also alert the price changes and upcoming events such as precooling. Occupants can adjust the economic index according to their monthly budget. The updated information from the occupant interface is directly used for the controller to make an optimal decision.

An engineering interface was designed to check all functions and algorithms of the controller by showing internal information flow on the screen. Therefore, this interface can be considered a temporary interface in development for a system engineer. After validating the functionality of the controller through the engineering interface, only useful functions and information for occupants were put on the occupant interface screen. In

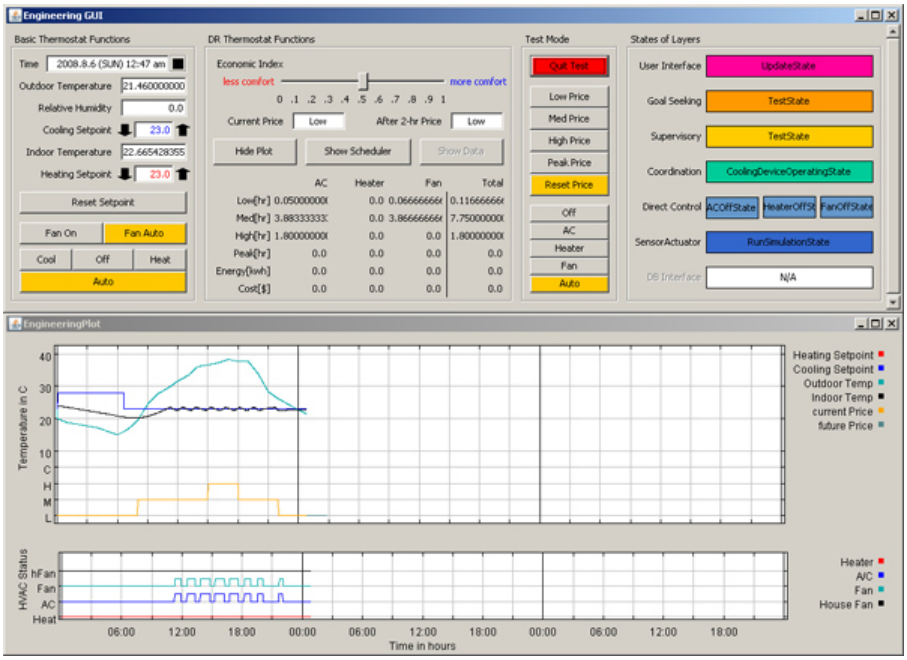


Figure 2.6: Engineering Interface of the DREAM

the sense that the engineering interface is relevant to debugging processes, the interface is less graphical and more numerical than the occupant interface. The screen shot of the engineering interface is shown in Fig.2.6.

The engineering interface consists of three individual windows: basic control window, plot window, and sensor network window. Basic information and functions that a conventional thermostat has such as time, temperature, setpoint, and mode switch are shown in the basic control window. In addition, a user can change the economic index that indicates the occupant's balance point regarding energy cost and comfort and check real-time energy use of electrical appliances, monthly accumulated cost, and current state of each control layer. In order to test wireless relay and price indicator, it has several buttons as well. Through the plow window, seven basic data points (indoor temperature, outside temperature, cooling setpoint, heating setpoint, current utility price, forecasting utility price, and HVAC status) are plotted. The plot helps the engineer to check system functionality at a glance as well as see the response to various control strategies directly. The sensor network window is used for wireless network diagnostics. Since the total number of sensors

and relays using wireless communication is unlimited and the control strategy is mainly depending on the sensing information, it is important to detect malfunctioning sensors and relays immediately to prevent further systematic problems. Assigned ID numbers of both motes that are activating within a wireless communication range and their sensors/relays are automatically recorded with their time stamps.

2.2.2 Layered Structure of Controller

The addition of demand responsiveness and whole house control to basic thermostat functionality leads to a control system with considerable complexity. In order to handle that complexity, we have adopted a layered design for the control system software. In a layered design, each layer (in theory) interacts only with the layers above and below it. This provides for modularization of function and semi-independent design of each layer. This form of layering has been a major factor in the success of the Internet. Networking was theoretically characterized by a seven layer design [8]. In practice, a four layer design was used to implement TCP/IP (Transmission Control Protocol/ Internet Protocol) – strict adherence to the layered design is why computers from all manufacturers, using network interfaces from many manufacturers, with signals going over wires, wirelessly, over fiber optics, all can connect seamlessly.

The layered structure of controller is easily realized into an executable programming code by using a Java scheduling package, TranRunJ that supports Tast/State design method

We are using the layered hierarchy shown in Fig.2.7 to implement the control. The lower part of the hierarchy describes basic control functions used to maintain temperature in the house and other functions the controller may be responsible for. The lowest layer, the Mote Interface Layer, maintains communication between the controller and the motes with sensors and actuators. Other than that, the function of the lowest levels is similar to a conventional thermostat – manipulate the HVAC system to maintain temperature in the

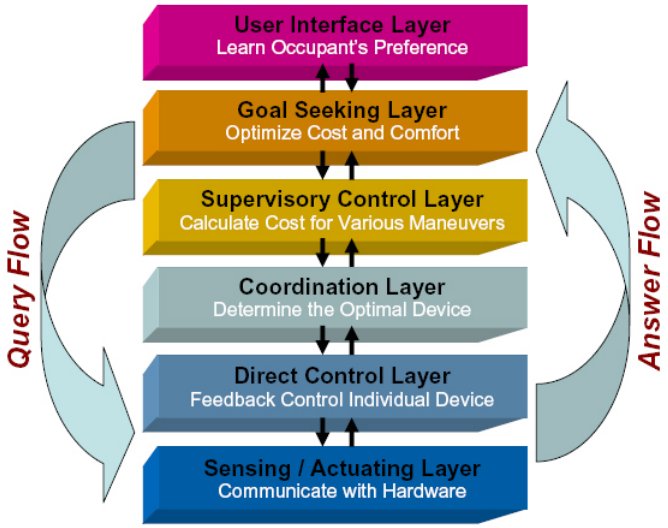


Figure 2.7: Hierarchy of Layered Control Structure

house. Layers above that, however, fulfill the demand responsiveness needs of the system. Of these, the most interesting is the Goal Seeking Layer. It makes decisions about how to best balance comfort and cost. In the middle, choices must be made as to how to meet the compromise decided on by the Goal Seeker. In many cases, there are choices as to how to achieve this, such as with whole house fan, air conditioning, ceiling fans, etc.

GoalSeeking Layer

The most complex layer is the Goal Seeking Layer. It receives pricing information and must make decisions about how to best balance comfort and cost. In the middle layers, choices must be made as to how to meet the compromise decided on by the Goal Seeker. In many cases, there are choices as to how to achieve the goal; for example, for cooling, one might use a whole house fan, air conditioning, and/or ceiling fans.

Supervisory Control Layer

While the goal seeking layer optimizes energy cost and user thermal comfort in response to a dynamic utility price and environmental circumstances, the four lower layers (Supervisory

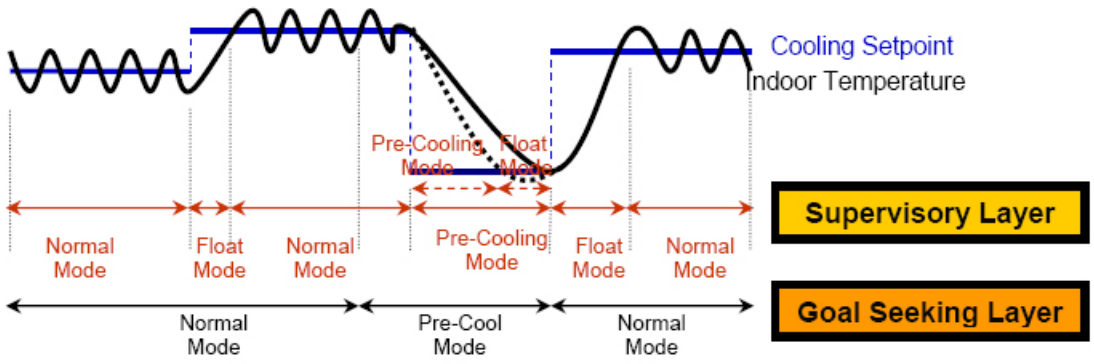


Figure 2.8: Mode Definition in Supervisory Control Layer and Goal Seeking Layer

Control layer, Coordination layer, Direct Control layer, and Sensing and Actuating layer) provide the information the Goal Seeker requires for its optimizing process. Generally speaking, the Supervisory Controller updates power and energy requirements for different modes, the Coordinator selects the least power consuming device in a given condition, the Direct Controller determines actual on/off timing with diverse control strategies, and the Sensor and Actuator interface predicts house behavior and communicates with hardware. The following describes more details of the roles of each layer.

The Supervisory Controller answers the questions, "How much energy will be required to keep indoor temperature at a specific setpoint?" and "If we have decided to precool the house, what time should the precooling start?" To answer these questions, we defined several modes based on the necessity of air-conditioning or heating operations. Precooling modes in the Goal Seeker and Supervisory Controller may not be identical depending on when an air-conditioner starts operating or stops operating. Modes defined in the two layers are compared in Fig.2.8.

- ◇ Normal Mode: keep a constant setpoint
- ◇ Precooling Mode: turn on a cooling device before the utility price increases
- ◇ Precooling with Normal Mode: keep the indoor temperature at the precooling set-point

- ◇ Preheating Mode: turn on a heating device before the utility price increases
- ◇ Preheating with Normal Mode: keep the indoor temperature at the preheating setpoint
- ◇ Float Mode: turn off HVAC device and let the indoor temperature float

The Supervisory Controller determines the current and future modes based on the current indoor temperature, a setpoint, and the predicted precooling start time, and stores the information on the power or energy requirement for individual modes.

Coordination Layer

The Coordinator manages several HVAC devices. The main role is to determine the least power consuming device in a given condition. For example, if the setpoint goal is to keep the indoor temperature at 74F when the outdoor temperature is 67F, the Coordinator has two options. An air conditioner can reduce the indoor temperature rapidly, but represents an expensive operation cost. The Coordinator will therefore choose the whole house fan instead of the air conditioner in this situation. The choice, however, may be different if the goal is to reach a setpoint of 74F within 30 minutes.

To determine the least power consuming device, the Coordinator needs information on the power consumption of individual devices. This information is organized with respect to both setpoint and outdoor temperature. After one device is chosen, the Coordinator activates the corresponding direct controllers and deactivates the other direct controllers.

Direct Control Layer

The Direct Controller determines the actual operation of HVAC devices. As mentioned above, the Coordinator activates some of the direct controllers. The Coordinator gives activation commands to related direct controllers at the same time. For example, air conditioning is provided by the combination of the compressor and blower fan. Some direct

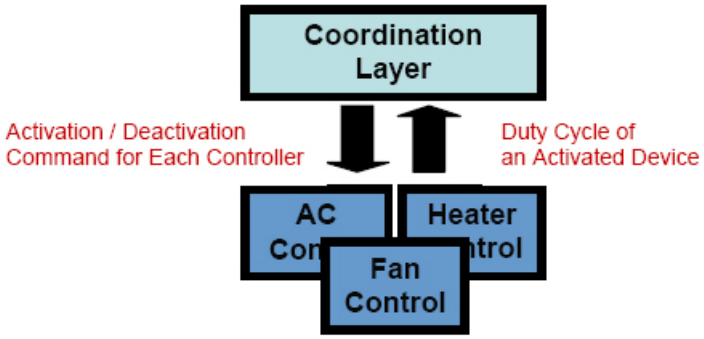


Figure 2.9: Interaction between Coordination Layer and Direct Control Layer

controllers have a minimum off time to protect the device from damage due to frequent on/off switches. In the current system, a four minute off time is applied for all direct controllers.

The output from direct controllers is an on/off command on a device. We considered three methods of direct control: on-off control with hysteresis, on-off control with asymmetric hysteresis (so-called anticipation), and cycle rate control. We are currently using on-off control with hysteresis since it is the simplest and is testable with a house simulation model. For on-off control with hysteresis, setpoint $\pm$1F is used as a default. For example, for a cooling setpoint of 74F, the AC controller sends an on signal when the indoor temperature is equal to or higher than 75F and an off signal when the indoor temperature is equal to or lower than 73F.

The duty cycle of devices is one of the important pieces of information that is determined in the Direct Control layer. A duty cycle is the ratio of on to total time. If an air-conditioner is turned on for 6 minutes and off for 4 minutes repetitively, this is a 60% or 0.6 duty cycle. The duty cycle under certain conditions is used to predict the energy consumption of the device in the future. The Direct Control layer calculates the duty cycle of each device and updates the previous one with respect to a specific setpoint and outdoor temperature.

Sensing and Actuating Layer

The lowest layer is the Sensing and Actuating layer. The main function of this layer is to communicate with wireless sensors, LEDs, and power relays. In simulation mode, communication can be tested by using a simulation house model, which will be explained in a later section. Since this layer is the gateway to the controller from the outside world, the information from outside (sensor data) and inside (command, setpoint) can be shared. For this reason, house identification and learning are performed in this layer. The next chapter describes this procedure in detail.

2.2.3 Auxiliary Structures for a Functionality Test and Algorithm Analysis

The goal of the DREAM project is definitely to develop a well-functioning energy management system for residential buildings. In this sense, an actual residential building will be an ultimate test bed. However, before a residential building can be used as a test bed, proper hardware functioning should be guaranteed and control algorithms in the system should be verified to minimize damage due to system failure. Another difficulty in dealing with an actual building as a test bed is an occupant. At a developmental stage, various control scenarios should be applied to the test bed, which inevitably disturbs occupants. Auxiliary structures of the DREAM are designed to minimize the problems that might occur during development by providing the system flexibility.

Flexible Control Code for Different Physical Environments

The layered control structure allows the system to be applied to different physical environments with minimal code changes. Since the Sensing and Actuating layer, the lowest layer in the control hierarchy, is the only layer that directly communicates with the physical devices, simply adding new states that deal with a corresponding physical environment

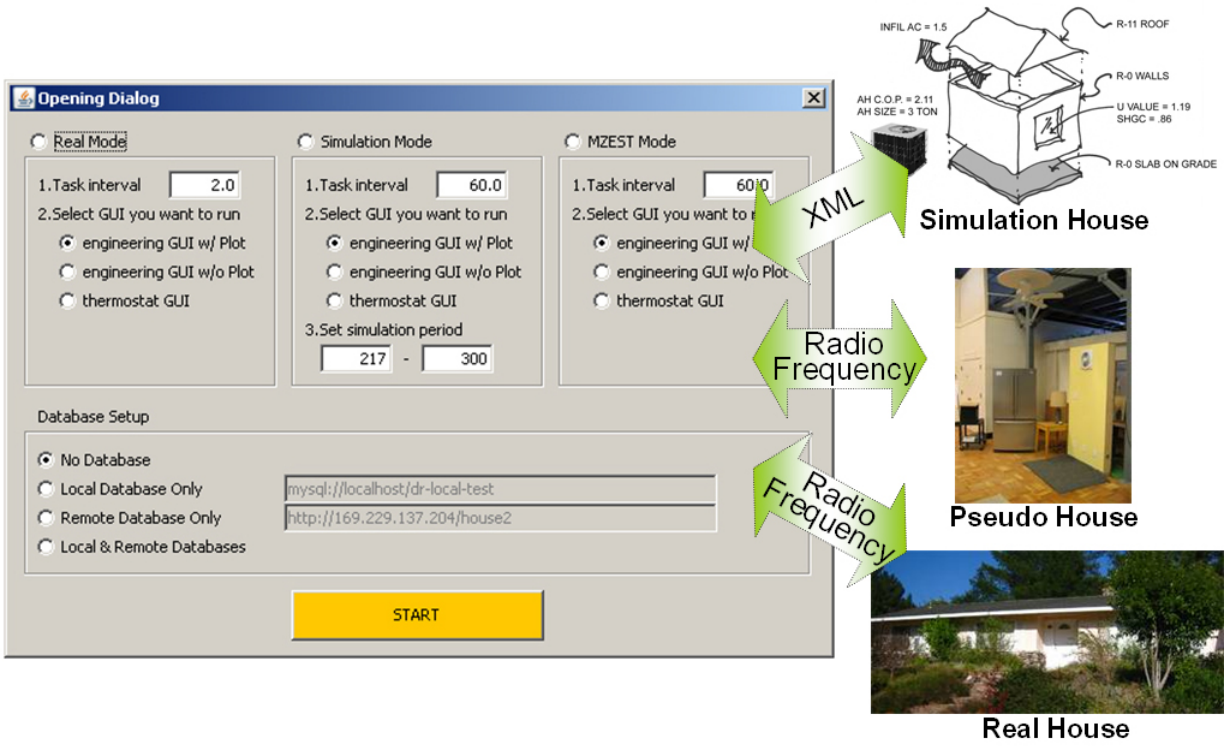


Figure 2.10: Different Physical Environments and Opening GUI

could provide enough flexibility to the system. An opening dialog GUI shown in Fig.2.10 helps an engineer to chose a suitable mode.

For testing purposes, three different physical environments (simulation house, pseudo house, and real house) have been used. Simulation house models are convenient to change their physical characteristics such as house size, house foundation, house location, and air-conditioner size. Therefore, the simulation model is useful to test basic functionality and find problems in the control algorithms. Compatibility with existing hardware (thermostat, HVAC equipment, and electrical appliances) cannot be tested with simulation houses. For hardware testing, a wall-type pseudo house was designed. The wall contains a fan, a power meter, a power outlet, and a set of sensors on it. By using the pseudo house, unexpected hardware-related issues and radio communication problems were identified and addressed. After sufficient tests, the DREAM system was finally applied to a real house. A real house has many more uncertainties and complexities in both house thermal behavior and

occupant's behavior.

Local and Remote Databases

In the DREAM structure, data collection is important in that historical data are a fundamental source to characterize the system's dynamic signature and the occupant preference or pattern as well as experimental trace to check the system's functional problems, find operational anomalies, and validate the control algorithms. For this reason, a persistent and stable database is imperative for the DREAM. The DREAM system has two different databases for its different purposes.

Besides sensor (measured) data, control (command) information and system information are stored in the central control unit (currently a small computer with MS Windows). Depending on the engineer's need, data type and data acquisition timing can be flexibly adjusted. The database that resides in the central control unit is called a *local database*. The local database is internally used for system or occupant learning. Therefore, it stores only necessary data for learning among all historical data and later deletes them if they are no longer useful. Since this recursive and selective data storing in the local database can limit the disk and memory usage by the database, it prevents the system from slowing down or abrupt stopping.

In contrast to the local database, the remote database is set up outside of the central control unit. The data, which can be different from that for the local database, is transmitted through the Internet to a server in the lab. Since the disk size of the server in the lab is much bigger than that of the local computer, all historical data can be saved. This database design is flexible, such that different sensors, motes, and measurement units can be added independently of one another. A change of hardware in the field requires only the modification of the proper ID numbers in the database. By assigning a unique mote and sensor ID for non-sensor data such as setpoint, utility price, and other commands, any type of information can be saved in the database. Since every single information is saved

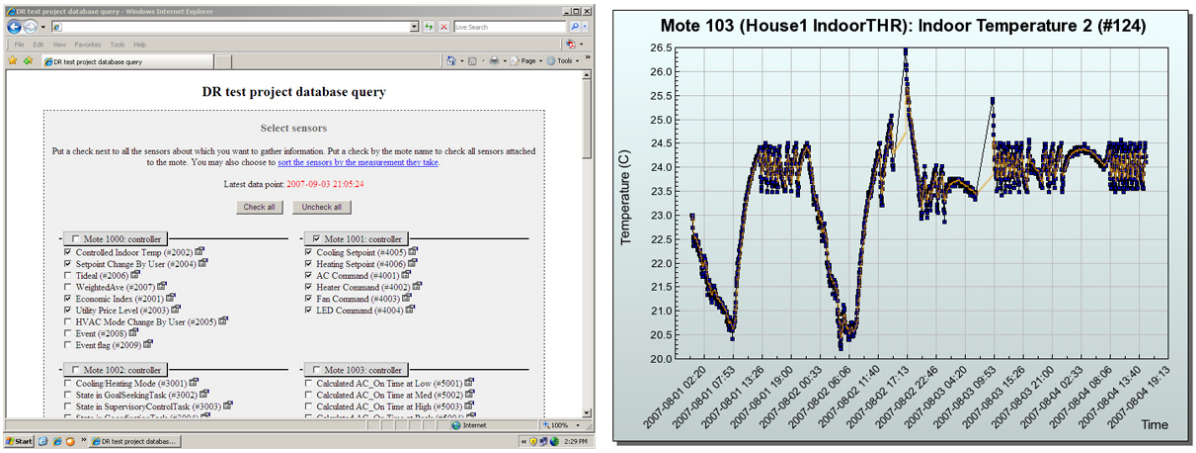


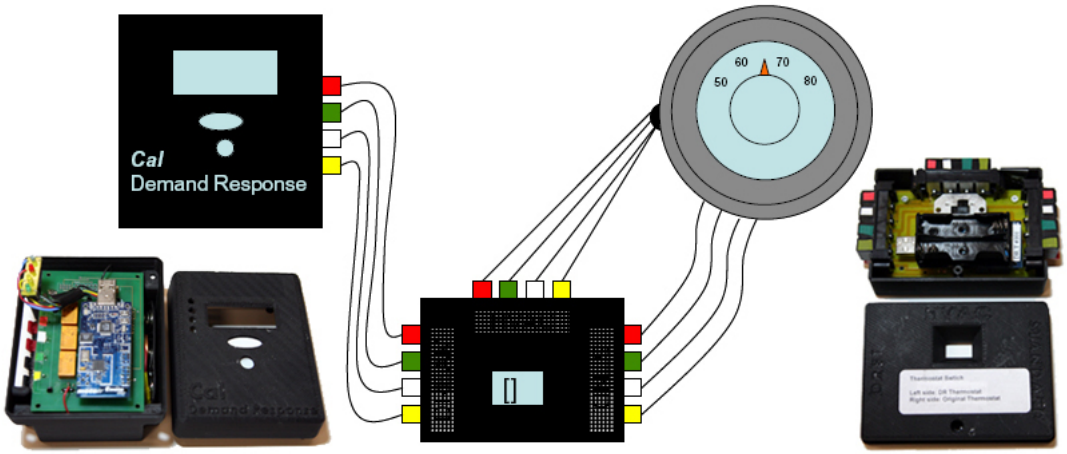
Figure 2.11: Interface of the Remote Database

with ID, this structure allows a local database to easily refer to the remote database in case the local database requires old data for learning.

To review an aspect of the operation and keep track of the problems of the system in real-time, an interface for the remote database was developed. This interface allows users (system engineer or occupants) to see the data selectively by choosing a data period and mote/sensor IDs. It shows a plot as well as an actual number to help users to catch the data profile easily. In addition, it automatically monitors the latest data uploading time and notifies the system's malfunctioning to the user with a red colored font when data uploading has not been performed for more than 5 minutes. This function allows the engineer to check the system status in real-time without traveling to the implementation site, and it also allows occupants to obtain useful information by themselves. Fig.2.11 shows the first web page of the remote database for data selection and the resultant plot for one of the selected sensors.

Switch Relay

As mentioned, the system failure during field test has serious results. Therefore, a preventive device to prepare for urgent system failure is necessary. A thermostat-switch and sensor mote is one of the device sets for safety as well as for experimental procedures. This



Source: Arens, Edward., et. al.[6]

Figure 2.12: Wiring between a Thermostat-switch(center), a DREAM HVAC Relay(left), and a Household Thermostat(right)

switch allows the user to completely disengage the aforementioned HVAC actuation mote and reconnect the household's original thermostat. The connection between the household's thermostat, the DREAM HVAC relay, and the thermostat-switch is diagrammed in Fig.2.12. In addition this device monitors the operation of the HVAC control wires, the output of both the experimental HVAC mote and extant house thermostat, and sends these feedback signals to the controller to verify operation of the relays.

The thermostat-switch contains a physical switch that physically shorts the four wires from the wall with two thermostats also attached to it. The second function of this device is to determine if the heater, fan, or AC is on. With two sets of sensors, the mote is able to determine which thermostat is shorting which signal, regardless of which position the switch is in.

The switch together with the sensors enables us to perform three tasks:

- ◇ Monitor the actual electrical connection of the HVAC relay mote before it is hooked up to the house's HVAC system for testing purposes.
- ◇ Monitor the usage pattern of the existing thermostat.
- ◇ Once the testing phase is over, we can still monitor which signal the HVAC relay

mote is sending to the HVAC system. The switch enables a fallback to the existing house thermostat in case something goes wrong.

2.3 Optimization and Control Strategies

In order for the DREAM to accomplish the goal of reducing electricity load during hours of peak demand while minimizing occupant thermal discomfort caused by the load reduction, the DREAM controller must be able to intelligently respond given inputs from outside. Since it is assumed that the electricity price is reasonably determined by utility companies based on either current or predicted electricity demand as well as unexpected urgent events, the only driving factor on which decision making in the controller is based will be the price.

Although the DREAM controller can control electrical appliances including HVAC equipment in an aggressive way through a compulsive shutdown, rather modest methods are applied in current control. Both electricity use from an air-conditioner and user comfort are controlled exclusively by setpoint changes. In this sense, a setpoint is the most important output that the DREAM controller should determine. Another method is informing and warning. Providing occupants with as much cost-related information as possible helps them to get more interest in their energy use and control their behavior in a smart way via interaction with the system.

Now the goal of the DREAM controller is narrowed down from electricity load reduction and occupant comfort preservation during peak periods to optimal setpoint determination. The next three subsections focus on how to define the optimal setpoint from a thermal comfort perspective and how to scale the user comfort. Based on the derived methodology, detailed optimization process and control strategies will be described.

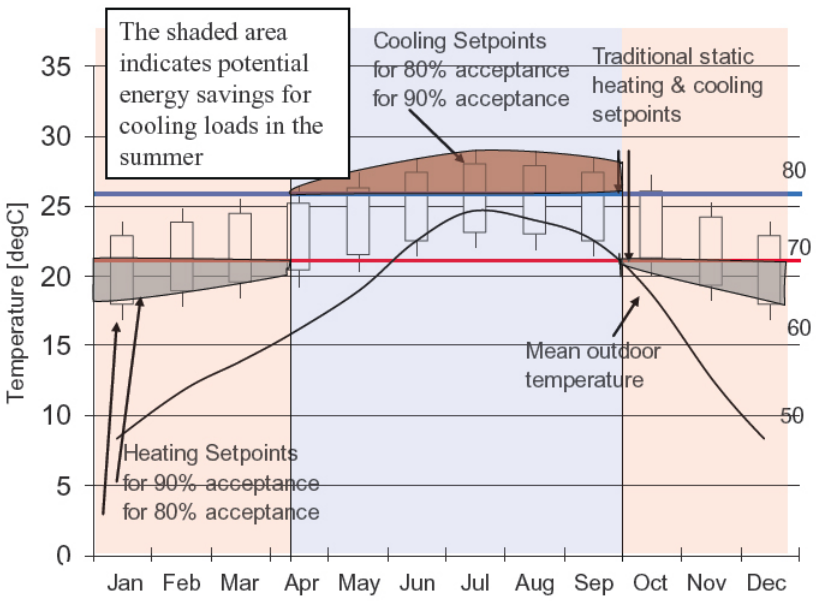
2.3.1 Optimal Setpoint in Residential Buildings

The typical programmable thermostat on the market controls indoor air temperature based on two different setpoints corresponding to comfort temperature for occupied condition and setup/setback for either unoccupied or night time condition. While these two setpoints in programmable thermostats are manually set by users, the DREAM controller should autonomously determine more than two setpoints (the DREAM controller uses dynamic setpoints instead of pre-determined static setpoints) from any kind of a thermal comfort standard. One study shows that dynamic setpoint can avoid over-heating or over-cooling that frequently occurs for some months under static setpoint setup [9]. This implies that people may decrease energy use without sacrificing their comfort by adopting a dynamic setpoint. Therefore, the key is how to scale thermal comfort by relating temperature to comfort level.

Although no thermal comfort standard has been established for the residential sector, the Adaptive Comfort Standard (ACS) from ASHRAE 55-2004 has been the most appropriate alternative standard in that it is based on naturally ventilated buildings [10, 11]. According to the ACS and two studies in the residential sector, people adapt to indoor temperature based on the average outdoor temperature for the previous month and there exists a seasonal variation in temperature [12, 13]. Based on the previous studies, standard comfort distribution[2] with respect to each month has been established.

Fig. 2.13 represents the difference between a static setpoint for typical thermostats and q dynamic setpoint based on monthly average outdoor temperature in Sacramento, California. The following discussion is mainly focused on the cooling setpoint rather than the heating setpoint since the ACS was developed with respect to cooling load, and energy demand is more concentrated during summer period. 78F (25.5C) in blue solid line is the default cooling setpoint for occupied condition most often used in an EnergyStar programmable

[2]Standard comfort distribution indicates how much satisfied ordinary human is with a given temperature. In Fig.2.13, a different term, "acceptance", was used instead of "satisfaction".



Source: Arens, Edward., et. al.[6]

Figure 2.13: Adaptive Temperature Setpoints for the Sacramento, CA Climate

thermostat. The standard comfort distribution indicated with rectangles shows that actual comfort range in summer periods (from May to October) is located at higher values. While the outside temperature keeps increasing and goes up to 100F (37.8C) in Sacramento in July and August, people adapt to this high temperature in any rate. The adaptation includes both psychological and behavioral changes. As a result, a higher setpoint than 78F (25.5C) during hot weather periods does not degrade their thermal comfort. In the adaptive thermostat temperature setpoint, the setpoint is 82F (27.8C).

The summer nighttime setback is another issue in determining an optimal setpoint. Several studies on the effect of temperature during a sleep period have been conducted. Tsuzuki et al found that the person feels too uncomfortable to fall asleep without adequate air flow if the air temperature is above 26C (78.8F) or the relative humidity is high, or person is clothed [14]. Another study shows that 26-27C (78.8-80.6F) compared to 31 or 36C (87.8 or 96.8F) provides better sleep [15]. From these results, 26C (78.8F) is used as a static nighttime setpoint from 10 pm to 6 am in our controller.

2.3.2 Optimization Schemes

In the DREAM, optimization is the process in which an optimal setpoint is calculated in a given environmental condition. As noted, the optimal setpoint is defined as a temperature value at which the cost and user comfort is well-balanced. Although the level of comfort that a person recognizes is difficult to be scaled, the adaptive comfort model based on the standard comfort distribution can be used for it.

The Goal Seeking layer primarily deals with the optimization process. However, the User Interface layer provides a scaled thermal preference from the adaptive comfort model, and the Supervisory Control layer helps the Goal Seeking layer to predict the cost that is required to maintain the condition. The process consists of three steps:

- ◇ Step 1 (Mode Determination): Heating or cooling mode is automatically determined based on recent weather condition.
- ◇ Step 2 (Control Strategy Selection): Based on house occupancy and a forecasted price change, the current state is categorized into 8 different states. Departure and arrival of an occupant is predicted by a learning method.
- ◇ Step 3 (Temperature Setpoint Calculation): By combining the information on user thermal preference and the cost variance to perform the selected control strategy (from the previous step), an optimal setpoint is calculated.

A typical residential thermostat has heating, cooling, and auto modes. From the current mode setup, a thermostat decides a setpoint type (heating setpoint or cooling setpoint) as well as a target HVAC device. The *auto* mode in a typical thermostat should be interpreted as a mixed mode rather than an automatic mode since both heating and cooling can be performed instead of choosing one of these mode. In contrast to the typical thermostat, the DREAM controller actually determines an appropriate mode based on the average outdoor temperature for the previous 7 days. The residential Alternative Calculation Method

State Name	Control Objective
Normal	Optimize energy cost and user comfort under no forecasted price and occupancy changes
Precooling / Preheating	Minimize energy cost by cooling or heating the house before price increases
Departure / Arrival Preparation	If the house is likely to be occupied, keep the indoor temperature at a desired value. If the house is likely to be unoccupied, do not
Combination	Precooling / Preheating state + Departure / Arrival Preparation state

Table 2.3: Categorized States and Their Control Objectives

(ACM) approval manual by the California Energy Commission (CEC) says, "the state of the building's conditioning mode is dependent upon the outdoor temperature averaged over hours 1 through 24 of day 8 through day 2 prior to the current day (e.g., if the current day is June 21, the mode is based on the average temperature for June 13 through 20). The ACM shall calculate and update daily this 7-day running average of outdoor air temperature. When this running average temperature is equal to or less than 60 °F the building shall be set in a heating mode and all the thermostat setpoints for the heating mode shall apply. When the running average is greater than 60 °F the building shall be set to be in a cooling mode and the cooling mode setpoints shall apply." [16]

House occupancy and utility price are two main criteria to determine an upcoming state among 8 states. Each state has its own control (optimization) objective. Tab.2.3 summarizes categorized states and their control objectives. Since price changes can be broadcasted by a utility company, the only issue is how to predict house occupancy as accurately as possible. It is evident that the consistency of an occupant behavior (schedule) is very important. The family size, day of week, and occupation can affect prediction quality as well. The DREAM controller uses an occupancy prediction model. The model calculates occupant's departure and arrival times by analyzing past occupancy data distribution.

The final output of the optimization process is a setpoint. Due to different objectives of 8 states, a procedure for setpoint calculation should be specified with respect to each

state.

Normal State

If there are no upcoming price increases and no house occupancy changes, the controller considers a current condition as a normal state. In the normal state, an optimal setpoint T^{opt}_{normal} is calculated to minimize the following utility function $U_{normal}(T)$[6].

$$\begin{aligned} U_{normal}(T) &= (1-e) \times Cost + e \times Discomfort \\ &= (1-e) \times Cost + e \times (1 - Comfort) \end{aligned} \tag{2.1}$$

$$T^{opt}_{normal} = \{T \mid min(U_{normal}(T))\} \tag{2.2}$$

where T and e[3] are temperature and the economic index.

Electricity cost is obtained by multiplying electricity price rate (dollar/kW) by measured power consumption (kW). Thermal comfort for a given user is calculated by the adaptive comfort model in the User Interface layer. In order for the Eq.2.1 to be meaningful, $Cost$ and $Comfort$ values are normalized. Economic index e can be directly input through the user interface by a user.

Precooling/Preheating (Pre-Price Change) State

In either precooling or preheating state, the most important issue is how much beneficial overcooling or overheating[4] the house during a relatively lower utility price period is in a total cost point of view. While the thermal comfort in a normal state is one of main factors for setpoint optimization, it merely constraints the lowest limit for precooling or

[3] Economic index e explains both sensitivity to thermal comfort and affordability in electricity cost. The value ranges from 0 to 1. The person who is more sensitive to his or her comfort and affords to pay more electricity cost has an economic index close to 1. In case that the economic index is 1, the optimal setpoint always maximizes user's comfort.

[4] Since both precooling and preheating setpoints are relatively far from the comfortable temperature, they can be considered as overcooling and overheating.

the highest limit for preheating. The optimal setpoint for the precooling state minimizes the total cost that consists of the cost before an actual cooling starts, the cost during an actual cooling period, and the cost during a high price period. Eq.2.3 shows three sub-cost calculation in the utility function $U_{precool}(T)$[6].

$$\begin{aligned} U_{precool}(T) &= Cost_{before.precool} + Cost_{during.precool} + Cost_{after.precool} \\ &= Price_{low} \times Power_{device} \times (t_{begin.of.precool} - t_o) \times DC_{before.precool} \\ &+ Price_{low} \times Power_{device} \times (t_{end.of.precool} - t_{begin.of.precool}) \\ &+ Price_{high} \times Power_{device} \times (t_{end.of.high.price} - t_{end.of.float.period}) \\ &\times DC_{after.precool} \end{aligned} \tag{2.3}$$

$$T^{opt}_{precool} = \{T | min(U_{precool}(T))\} \tag{2.4}$$

where t and DC are time and duty cycle of a cooling device.

If a soaking effect[5] is not considered, determining an optimal setpoint for precooling $T^{opt}_{precool}$is the same as determining a precooling start time $t_{begin.of.precool}$. $t_{begin.of.precool}$ differs from the state transition time t_o(the precooling state in the Goal Seeking layer begins at t_o). If $t_{begin.of.precool}$ is given, the final indoor temperature at the end of precooling would be $T^{opt}_{precool}$. When the utility price is increased at the end of precooling, cool indoor temperature tends to increase under no cooling device operation. This period is called a *float period*. In Eq.2.3, $(t_{end.of.high.price} - t_{end.of.float.period})$ is the total cooling device operation time in a high price, which should be as short as possible for effective precooling. More details are described in the next subsection.

[5]If two identical houses have been precooled down to the same precooling setpoint, the house that has stayed at the setpoint for longer time is warmed up more slowly than another house. In a heat transfer point of view, the long exposure to cool temperature allows objects (walls and furniture) to be cooled besides air.

Departure/Arrival Preparation State

The controller optimizes energy cost and thermal comfort by either avoiding an unnecessary heating and cooling or preconditioning the house during the departure/arrival preparation state. The following two scenarios provide how this state works.

Scenario 1 At 8:43am on a hot summer day, the outside temperature has reached 85F. In order to keep the indoor temperature at 73F for occupant's comfort, the air-conditioner has been run for 1 hour. However, the DREAM schedules to turn off the air-conditioner at 8:45am even though the house will be still occupied. Since the DREAM controller knows with high probability that the house will be unoccupied around 9:00am on most weekdays, turning off the air-conditioner at 8:45am will save the occupant money without affecting his or her thermal comfort.

Scenario 2 At 5:00pm on the same day, the occupant has done his work and was thinking of the cold beer that he put in the refrigerator that morning. When he opens the door as soon as he arrives home around 5:30pm, the awful heat inside the house is waiting for him. Although the air-conditioner begins to run, the thermometer still indicates 82F. He recognizes that he needs a smarter thermostat more than cold beer.

The optimal setpoint in this state is determined based on the prediction probability. The utility function $U_{arrival/departure}(T)$ in Eq.2.5 combines the utility functions, $U_s(T)$ and $U_f(T)$ for two probabilistic cases.

$$U_{arrival/departure}(T) = P_s \times U_s(T) + (1 - P_s) \times U_f(T) \tag{2.5}$$

$$T^{opt}_{arrival/departure} = \{T | min(U_{arrival/departure}(T))\} \tag{2.6}$$

where P_s: Probability that prediction succeeds

$P_f = 1 - P_s$: Probability that prediction fails

$U_s(T)$: Utility function for the case that prediction succeeds

$U_f(T)$: Utility function for the case that prediction fails

Combination State

In some cases, precooling/preheating events and house occupancy change occur at the same time. If the directions of the desired setpoints for both events do not agree, the controller makes an optimal decision. The decision mainly depends on the prediction probability of arrival or departure.

2.3.3 Control Strategies

Most control and optimization strategies by the DREAM controller are based on the house identification (or dynamic signature learning). Under the assumption that the indoor temperature can be predicted from the given data within an acceptable error range, the DREAM controller includes the following control strategies. The house identification procedure including theoretical backgrounds is discussed in Chapter 3.

Minimum Off-Time Constraint

A minimum-off time control is used in most air-conditioners to protect the compressor motor from frequent on/off switching. It intentionally delays the restart of a compressor motor for a specified minimum time. Since the minimum-off time is not controlled by a thermostat but by an AC unit, this control algorithm seems to be unnecessary for the DREAM controller. Initially the minimum off-time control algorithm was implemented to mimic the behavior of a real thermostat and an air-conditioner in simulation.

In contrast to a typical residential thermostat that mainly relies on the on-off control with respect to the setpoint, the DREAM controller primarily depends on the behavior prediction of the house and HVAC equipment. For example, if a typical programmable

thermostat sends an *On* signal to the AC unit before a specified minimum off-time is elapsed, the minimum off-time device connected to the compressor motor will automatically cut off the power to the motor. This intentional blocking does not affect the control goal and quality of the thermostat. However, the blocking can affect the quality of optimization by the DREAM controller since several-minute-long air-conditioning in a peak time is totally different from that in a low price period.

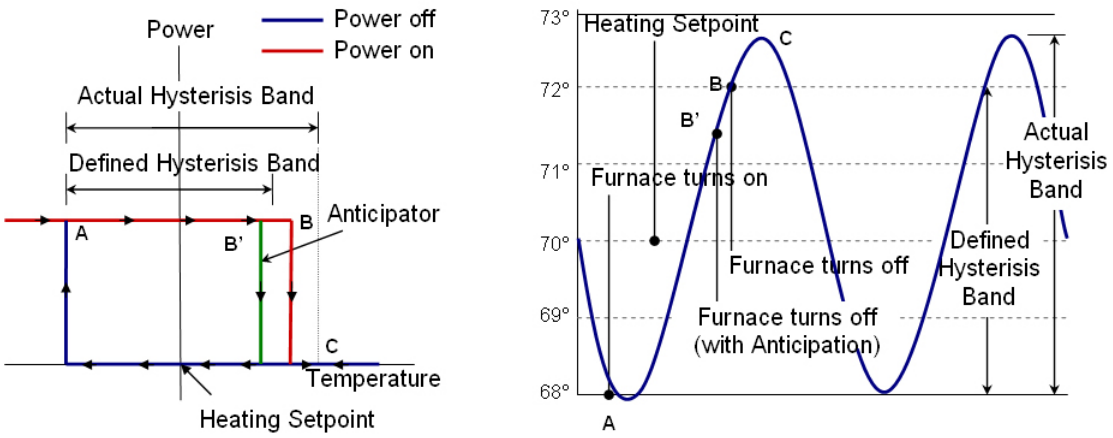
Besides the control quality problem, there is another reason why the minimum-off control should be performed in the controller level. The DREAM is based on a wireless sensor network. When the controller sends an *On* signal to the HVAC relay, the HVAC relay follows the command and sends back the updated HVAC status. If the feedback signal and command signal are different, the controller will send the command repeatedly until the feedback and command agrees. Currently the DREAM controller uses a 4-minute minimum off-time control in the Direct Control layer. The 4-minute interval was selected by considering minimum requirements for hardware protection and efficiency in control.

Adaptive Hysteresis Band Control

A temperature setpoint is much more important in the DREAM than in typical programmable thermostats since one criterion for performance valuation is an occupant's thermal comfort level (assumed to be a function of indoor temperature). When people set their temperature setpoint at certain values, they naturally expect that their thermostats will keep the indoor temperature exactly at those values. For example, if the setpoint is 73F, they may think the indoor temperature will be 73F sooner or later.

However, all programmable thermostats use on-off control with their own fixed hysteresis bands[6]. The on-off control with hysteresis helps a furnace or an air-conditioner not to turn on and off frequently by allowing fluctuation around the setpoint. Some thermostats also use an anticipator to make the original (defined) hysteresis asymmetric with respect to

[6] A hysteresis band is also called a "temperature swing" in some thermostat manuals



Sources: Auslander, David, M., "Feedback Control Characteristics of Residential, Single-State Thermostats.", 2006[17].

Figure 2.14: Power-Temperature Characteristic with Hysteresis and Anticipator

the setpoint. Fig.2.14 describes the on-off control with hysteresis and anticipation in the heating mode. Without an anticipator, the original hysteresis is symmetric. (A (68F) and B (72F) are symmetric with respect to the setpoint 70F). However, the actual temperature swing is asymmetric (A (68F) and C (73F) are asymmetric with respect to the setpoint 70F). The anticipator turns the furnace off earlier than would have happened without the anticipator, and residual heat around the furnace keeps increasing the indoor temperature.

In the DREAM controller, the size of a hysteresis band is adaptive to the sizes of the furnace and air-conditioner as well as to the house. When the air-conditioner size is relatively large in comparison to the house size, a larger band will avoid frequent on/off cycles. The timing of turning off a furnace or an air-conditioner for anticipation is also determined by a house learning.

Preconditioning

One of the ways that the DREAM controller provides occupants comfort is preconditioning. With a traditional non-programmable thermostat, an occupant should wake up in uncomfortable temperature (too cold in the winter or too hot in the summer) if he or she uses a night-time setback setpoint. A programmable thermostat seems to solve this

problem by scheduling the heater or AC operation in the morning. For example, an occupant usually wakes up around 7:00 am, he or she may schedule the thermostat to turn the air-conditioner on at 6:30am. However, the time within which the indoor temperature can reach the desired setpoint strongly depends on the weather condition. If the outside temperature is very hot, 30-minute-long air-conditioning may not be enough to cool the house down by 7:00 am. If on the other hand the outside temperature is relatively mild, this much air-conditioning may not be necessary.

In order to minimize the expense for comfort, it is important to turn on the air-conditioner or the furnace at a proper time. The DREAM controller determines the time by predicting the indoor temperature behavior based on the given weather conditions and house information. According to one study (EcoFactor, 2008), appropriate preconditioning benefits the least efficient house since the less able a given structure is to store energy, the more its energy usage will be determined by the length of time.

Precooling

The concept of precooling originates from night precooling. Night precooling cools the structure during the night time by the circulation of cool air within a building. The cooled structure is then able to serve as a heat sink during the daytime hours. Therefore, the quality of night time weather such as temperature, humidity, and pollution and the thermal coupling of the circulated air to the building mass are crucial for precooling [18].

There are two variations on night precooling. One is night ventilation precooling and another is mechanical precooling. This involves the circulation of outdoor air into the space during the naturally cooler nighttime hours. This can be considered a passive technique except for any fan power requirement needed to circulate the outdoor air through the space. In mechanical precooling, the building mechanical cooling system is operated during the nighttime hours to precool the building space to a setpoint, which is usually lower than that of normal daytime hours. The electric utility rate for peak and off-peak loads is important

to determine the cost-effectiveness, in particular for a mechanical precooling scheme. One study result that precooling enables significant energy cost savings of 10% to 50% and peak power requirements of 10% to 35% over a traditional nighttime setup control strategy is very encouraging [19].

The precooling strategy in the DREAM is closer to the mechanical precooling than to the night ventilation precooling. Since the night ventilation precooling requires whole house fan for air circulation, the mechanical precooling is less limited. The big difference between the DREAM's precooling and the conventional mechanical precooling is a precooling time. While the conventional precooling cools the building structure during the night, the DREAM's is performed within several hours (or possibly several minutes) before the peak price comes. Therefore, the DREAM's precooling is less sensitive to the thermal energy storage in the building mass in that even precooling without soaking can allow significant cost saving. But it requires more delicate control for a fine performance.

Fig.2.15 illustrates how the electricity load originally assigned in a peak price period can be shifted to the earlier lower price period by an appropriate precooling. A peak price period ($t3+$ $t4$) is shaded in light red. The black oscillating curve stands for an indoor temperature profile under a peak-time setback. Without precooling, the air-conditioner still consumes a lot of energy during a peak time. If precooling starts from $t2$ (the dotted curve in dark gray), it can shift the load, but not enough. When it begins at the middle of $t1$(the dotted curve in dark red), no air-conditioning is required during a peak time.

It is certain that a longer peak price period requires a lower precooling setpoint. Unfortunately, the lowest cooling setpoint is limited by the comfort allowance, outdoor temperature, and the efficiency of a cooling device. In addition, a total precooling time is not linearly proportional to a temperature drop. The following steps show how the DREAM controller determines an optimal precooling start point.

1. Check if the price increases within a predefined period. (3 hours in the DREAM. This limits the maximum precooling period)

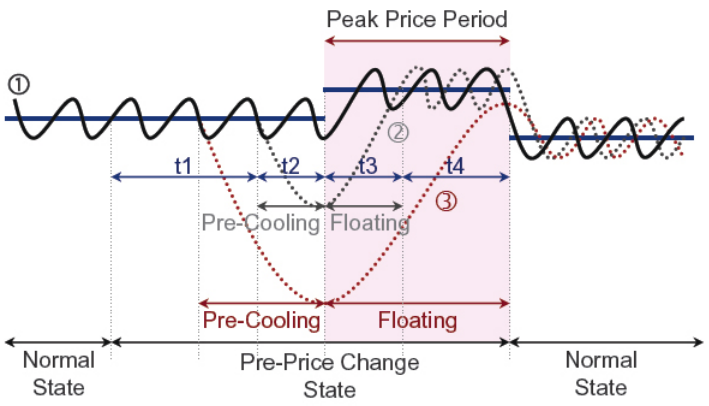


Figure 2.15: Schematics of Precooling Strategy in the DREAM

2. If the price change is scheduled within 3 hours, calculate the lowest temperature that can be reached by full AC-on. If the lowest temperature is out of the occupant's comfort allowance, the lowest temperature is adjusted.

3. Every one degree between the current setpoint and the (adjusted) lowest temperature are precooling setpoint candidates.

4. By using both backward prediction ($t1$+ $t2$) and forward prediction ($t3$+ $t4$), calculate AC on/off time (or duty cycle) in different price periods.

5. Find the setpoint that satisfies Eq.2.4.

2.4 Simulation Results and Field Test

Even though the DREAM still has many things to be improved and modified, the current version suggests a new direction of future residential energy management systems and shows possibility in developing a demand responsive thermostat. The following two subsections deal with simulation results of optimization and control by the DREAM and discuss the field tests that have been completed in the summer 2007. A numerical house model, called MZEST (Multi-Zone Energy Simulation Tool), was used for functionality check and control algorithm validation of the DREAM controller. Since MZEST is based on the

physical properties of a real house, HVAC devices, and occupants, it can provide realistic thermal behavior of the residential building. After sufficient tests and validation processes, a preliminary field test (deployment test) was performed for approximately one year, and fully functioning elements of the DREAM were deployed in two real houses in California.

The results related to optimization and control strategies are discussed in this section, but the topics related to house identification (or dynamic signature learning) are fully covered in Chapter 3. Further details of both simulation and real houses also can be found in the next chapter.

2.4.1 Simulation Results

Since the core of the DREAM is self-learning or self-tuning that can name the DREAM as an autonomous system, both the quantity and quality of collected data are crucial for learning. Based on the learned information, the controller can make an optimal decision. However, the DREAM should have considered the worst case in which there does not exist enough data for learning, for example when an occupant moves to a new location.

To provide reasonable[7] outputs without enough data, good defaults are used initially. The defaults for an occupant's thermal preference are set by the ASHRAE (American Society of Heating, Refrigerating and Air-Conditioning Engineers) standard and both thermal characteristics of the house and energy consumption of the air-conditioner are calculated by averaging the results from four extreme houses in California.

One aspect of the optimal setpoint determination process for the normal state in the Goal Seeking layer is seen in Fig.2.16. The plot is based on the actual outputs that the DREAM controller generates with defaults. The right plot shows the opposite tendency in the energy consumption by the AC and comfort level that an occupant might recognize with respect to the cooling setpoint. It is certain that the discomfort level will increases

[7]"reasonable" means that the prediction by the controller, for example, the indoor temperature after 30 minutes from the current time, is within an acceptable region. If the difference between real measurement and predicted value are 10F in the previous example, the prediction may not be reasonable at all.

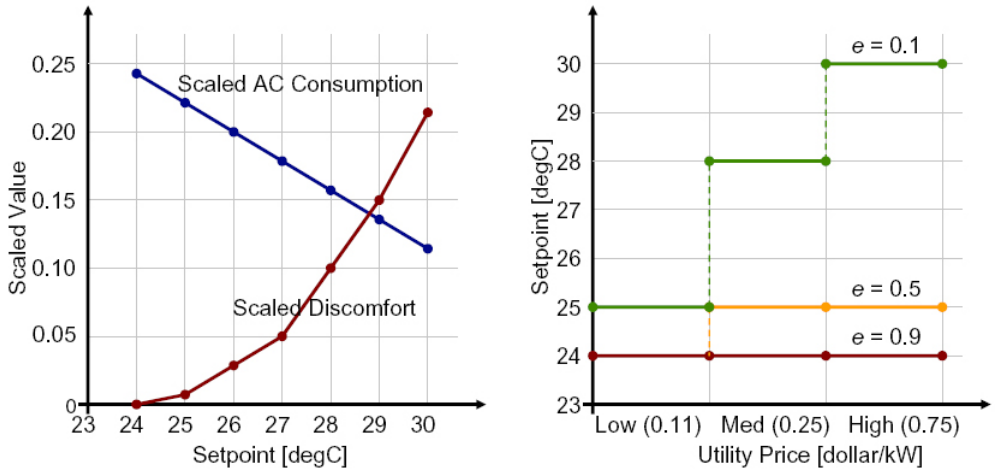


Figure 2.16: Setpoint Optimization (Right) from Energy and Comfort Defaults (Left)

as the air-conditioner decreases its operation time. The right plots imply that the cooling setpoint tends to be increased as the utility price increases, but the setpoint change by the price is less affected with a large economic index. Since these results came from the defaults, all houses would have the same outputs in prediction.

After enough data is collected, the DREAM controller has a different (more specified to the house and occupant) energy consumption curve and a different discomfort curve. The left plot in Fig.2.17 was not the result from actual learning but one example of the specific house and occupant. From the left curves, the controller performs the exact same optimization process and provides a different cooling setpoint. The most recognizable difference between Fig.2.16and Fig.2.17 (see the left plots only) is the scaled AC consumption. The AC consumption in Fig.2.17 is smaller than that in Fig. 2.16. After balancing cost and comfort, the controller dropped the cooling setpoint in 1 degree for the medium price. The increased setpoint for the low price is not explained clearly from the plots but the different discomfort profiles may affect this result.

Evaluating the performance of optimization by the DREAM controller is not trivial since a comfort level is one of two main criteria in evaluation. In the simulation, a user comfort model with defaults was used. In Fig.2.18, the results from four different setpoint

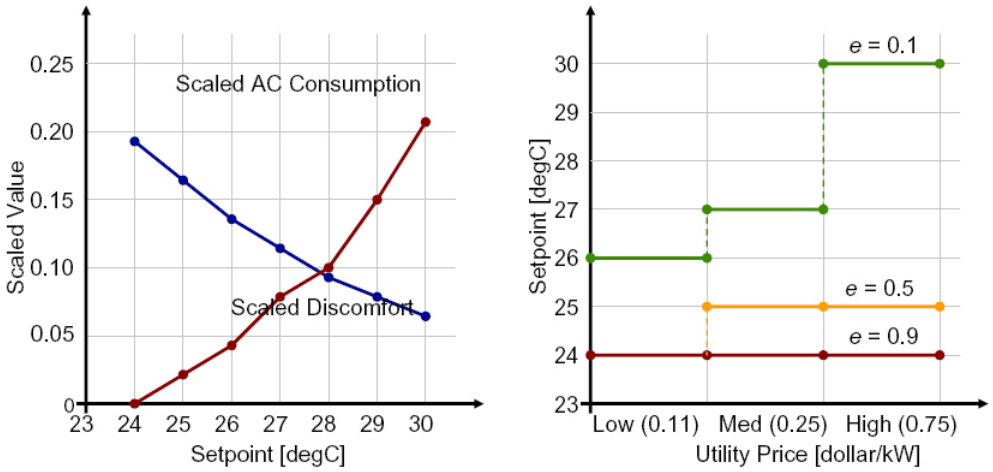


Figure 2.17: Setpoint Optimization (Right) from Learned Energy and Comfort Information (Left)

setups are compared. The simulation house is pre-1978 (constructed before 1978, usually poorly insulated) with crawl space and modeled in Sacramento, California from MZEST. A detailed setpoint setup is as follows:

- ◇ Typical programmable nighttime setback : 24.5C (76.1F) in daytime and 28C (82.4F) in nighttime.
- ◇ DREAM with 0.5 and 0.2 economic indices: Adaptive setpoint is applied based on the price and comfort level. Precooling is also performed whenever necessary. The smaller an economic index is, the more sensitive to cost an occupant is.
- ◇ Price-based setback: 25.5C (77.9F) for low price, 26.5C (79.7F) for medium price, and 28.5C (83.3F) for high price in daytime. 28C (82.4F) for all price in nighttime.

Total air-conditioning time in different price was calculated by accumulating 15 consecutive days. The plot indicates that the setpoint setup in a typical programmable thermostat is not responsive to the price at all even though the user may feel mostly comfortable. In contrast, the price-based setpoint dramatically decreases the energy use during both medium and high price periods, but the user may feel frequently uncomfortable. The setpoint setup in DREAM shows possibilities in how to balance these two extremes. Under a

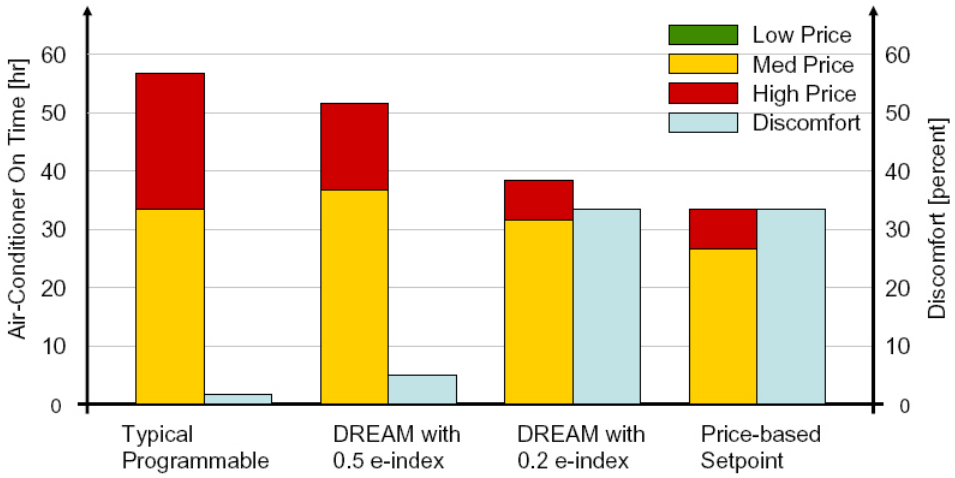
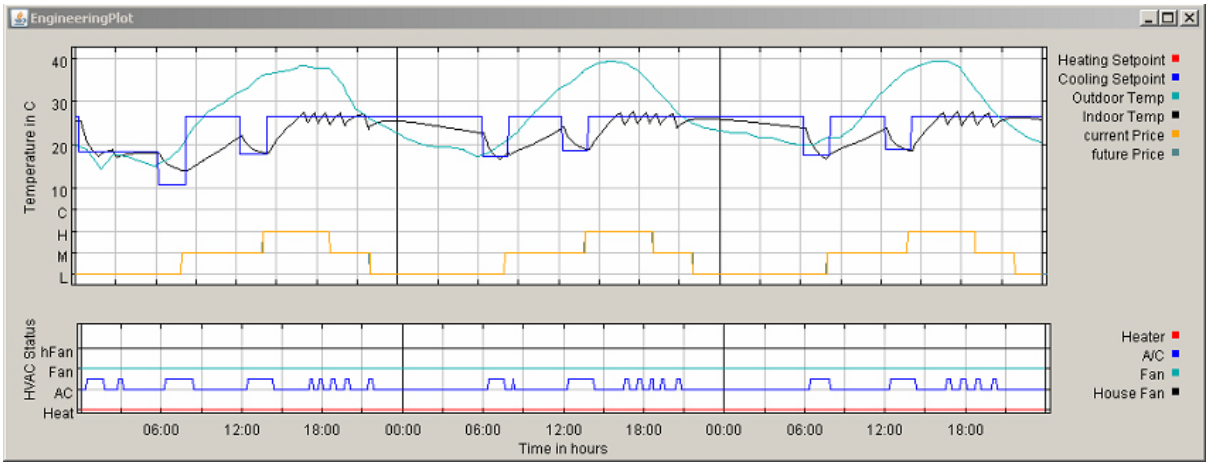


Figure 2.18: Performance Comparison : Energy Consumption vs. Discomfort

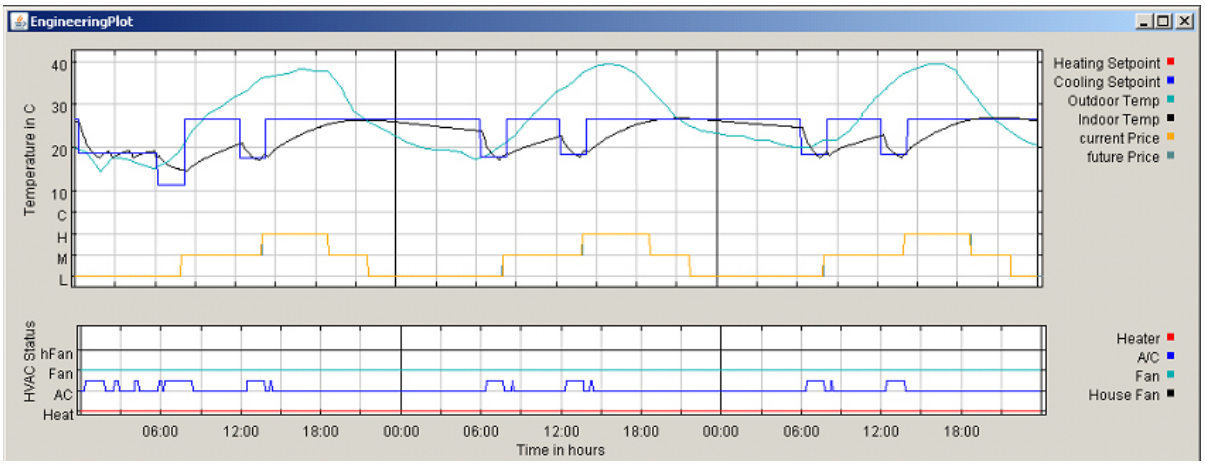
default economic index (0.5), the DREAM successfully shifted approximately 30% of the load in the high price period with a 4% increase in discomfort. With a relatively small economic index (0.2), the performance from a cost perspective is worse than that of price-based setpoint thermostat. It is true that the optimization of the DREAM may not be the best option in a specific environment. However, it may provide more diverse-situations-reflected setpoint that is impossible with fixed setpoint setup.

Fig.2.19 and Fig.2.20 are are direct screen shots of the DREAM engineering interface. They show how well the default house parameters fit different houses and how much indoor temperature prediction can be improved by learning. During a three day simulation, six total precooling events occurred at 6am and 12pm. In this simulation test, while the precooling start time is fixed, the precooling setpoint was determined by the DREAM controller. Therefore, the closer the indoor temperature (black curve) at the end of precooling is to the precooling setpoint (blue line), the more accurate the prediction is.

In Fig.2.19, the setpoint profiles in both plots are identical since the setpoint is determined based on the same default house parameters. However, the responses of the two houses are different. Since the crawl space model constructed before 1978 has relatively poor insulation, load-shifting in high price periods is not sufficient even with precooling.



(a) Pre-1978 Crawl Space Model (poorly-insulated house)



(b) Post-1992 Slab on Grade Model (well-insulated house)

Figure 2.19: 3-Day Simulation Based on the Prediction with Default Parameters

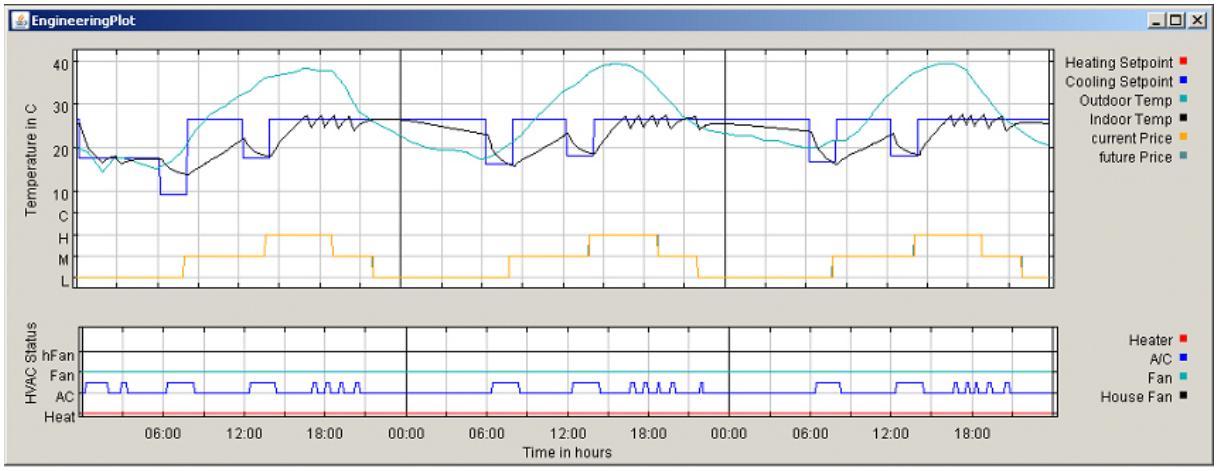


Figure 2.20: 3-Day Simulation Based on the Prediction with Learned House Parameters

	House 1	House 2
Location	Antioch, CA	Bay Point, CA
Structure	1700 square foot two-story stucco house	1500 square foot one-story house
Construction Year	1991	1984
Thermostat	White Rodgers programmable thermostat.	manual setback Honeywell Chronotherm.
HVAC System	Carrier split system air conditioner/furnace, with supply grilles in the floor throughout the house	General Electric split system air conditioner/furnace, with supply grilles in the ceiling
Occupants	Occupants are working at home. Normal setpoint in both daytime and nighttime is 74F and 79F when unoccupied. Participant opens up the windows (upstairs) at night and closes during the day.	Occupants are normally out of the house during the day. Normal setpoint in daytime is 70F and lowered or off in nighttime. The participant opens up the house at night; two windows are opened during the day as well.
Mote Deployment	Total 15 motes. A mote outside is under the southeast eave of the roof.	Total 14 motes. A mote outside is under the southeast eave of the roof.

Table 2.4: Description of Two Field Test Houses

For the slab on grade model, precooling is more beneficial due to good insulation.

All situations in Fig.2.19(a) and Fig.2.20 are equal, except what the precooling setpoint is based on. By analyzing past data and then updating parameters of the prediction model, the controller in Fig.2.20 could provide a more accurate precooling setpoint. Improvement in setpoint prediction can be seen by comparing the final indoor temperature in precooling and the setpoint.

2.4.2 Field Test Results

Test Houses and System Deployment

Two occupied houses were selected to test the function of the DREAM and verify simulation results. The feedback from participants also was important. These two single family detached houses are located in Antioch and Bay Point, California respectively and use their air-conditioning during the summer period. While the weather conditions for both houses are very similar, the house structures, HVAC systems, and residents' patterns were totally different. This diversity provided the opportunity to test the system under different conditions. Detailed house information is summarized in Tab.3.1.

Test Plan and Schedule

The whole DREAM system was scheduled to test for 37 days. The first week was spent in monitoring the existing thermostat setups, occupants' patterns, and AC power consumption without direct control by the DREAM controller. This period is defined as *continuous system check-out and calibration period*. During this period, any programming bugs and hardware failure were checked. The next two days were scheduled to test if the DREAM controller could control the existing HVAC equipment in the exact same manner. The DREAM controller followed the same setpoint that the existing thermostat used for the previous week. This period is called *mimicking and training period*. This period also provides an occupant with an opportunity in interacting with the DREAM interface. During the *test period*, control methods, optimization, and learning algorithms were validated. The field test was completed with a final interview with the occupant and hardware withdrawal. Tab.2.5 summarizes the field test plan.

Test Period		Objective
Continuous System Check Out and Calibration Period	7 days	Test system communication reliability. Analyze room occupancy and house occupancy pattern. Analyze temperature patterns. Record current energy consumption patterns, especially air conditioning. Evaluate the existing internal model . Analyze AC energy use and efficiency.
Mimicking & Training Period	2 days	Test actuation performance and ability of system to mimic existing system. Train users to interact with DREAM interface.
Test Period	2 days	Analyze AC duty cycle.
	8 days	Test optimization performance without precooling.
	9 days	Test optimization performance with precooling.
	4 days	Run new learned internal model with tuned parameters.
	4 days	Test the effect of knowing houseoccupancy on energy use and comfort.
Final Interview	1 day	Interview the participants and remove hardware.

Table 2.5: Time Schedule and Detailed Objectives for 2007 Summer Field Test

Results

The DREAM system had full control of the HVAC systems of two houses for several weeks. However, there were a few problems, some of which were solved immediately and others that warrant further analysis. The test for house 1 lasted approximately seven weeks due to fixes and working around the participant's schedule. The test for house 2 lasted about six weeks.

We captured the behavior of the house plus HVAC for warm days and very hot days; over the course of the tests, the outdoor daily high temperature ranged from 27 to 41C (80.6 to 105.8F). House 1 performed reasonably well under hot conditions, and responded well to precooling scenarios. Figure 27 shows the inside (T_i), outside (T_o) and AC supply temperature (T_{ac}) for house 1 for three days. The air conditioner was cycling all three days to cool the house, but struggled on the hottest day.

House 2, however, appeared to have an undersized HVAC unit, which could barely keep

up on hot days, and was completely underpowered for very hot days. Precooling was not an option for this house. Figure 28 shows the same three days in August. The air conditioner was cycling the first two days (albeit on for much of the time). On the hottest day the air conditioner is on constantly, but still allows the temperature to drift up.

Goal seeker and occupancy schedules One test was to evaluate the goal seeker in optimizing the temperature setpoint for occupancy and price. Although simulations show a reasonable setpoint generated by goal seeking layer, the actual results are hard to evaluate. In house 1, there was not much opportunity for optimization for the system because the acceptable temperature range by the resident was very narrow, especially for precooling. The participant complained that it was too cold during precooling mode and too hot when the temperature setpoint was raised during a simulated price increase. The goal seeker was not able to work for house 2 either. The air conditioning system for house 2 was apparently too small for the house and climate. On very hot days, the occupants turned on the air conditioner in the morning, otherwise the house would get too hot in the afternoon. On one such day, at 11 am the indoor temperature kept increasing although the AC was on. In this case, the goal seeking strategies do not work.

An occupancy switch mote was used to collect the occupancy information by requiring the participants to push a button indicating arrival or departure. The two houses showed different occupancy schedule patterns. In one house, the occupant worked from home and was home intermittently during the day. In the other house, the occupant worked away from home, and the house was unoccupied regularly for a certain interval during weekdays. The first type of schedule pattern is random and thus hard to identify. It is difficult for the system to predict the changes of status. In the second type, it is easier for the system to learn the occupancy pattern. However, the participants forgot to manipulate the occupancy switch most of the time, rendering the data unusable.

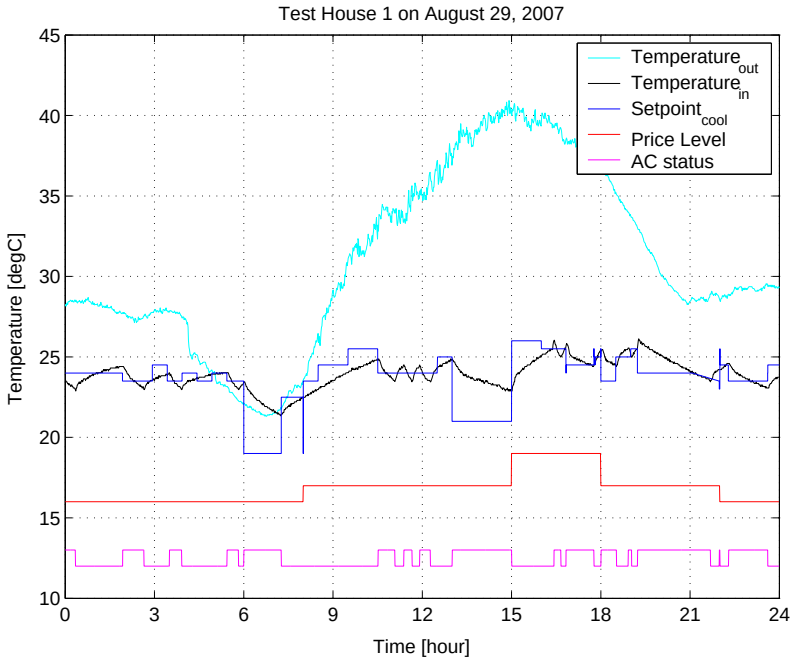


Figure 2.21: Temperature and Setpoint Profile for House 1 with Default Parameters

House Identification Results Prediction through the internal house model seemed satisfactory in simulation, but met problems in field testing. Changes in outdoor temperature and infiltration due to open windows seriously affected the prediction. Thus, the internal house model did not predict the indoor temperature as accurately as in simulation. However, from the field test, we learned about the more realistic thermal behavior of a house and characteristics of HVAC equipment.

During most of the test, the internal house model with good defaults was evaluated. The prediction results were used to determine the optimal temperature setpoint in the Goal Seeking layer. One way to check the accuracy of the prediction is to compare the predicted precooling setpoint with the real indoor temperature profile. Fig.2.21 shows outdoor temperature, indoor temperature, cooling setpoint, utility price and AC status in house 1 on August 29, 2007. There were two simulated utility price increases at 8 am and 3 pm. We can see that the precooling setpoints were set at 19C and 21C at 6 am and 1 pm respectively. The precooling setpoints chosen in both cases were somewhat lower than required. This means that the internal house model may have a higher capacity AC unit,

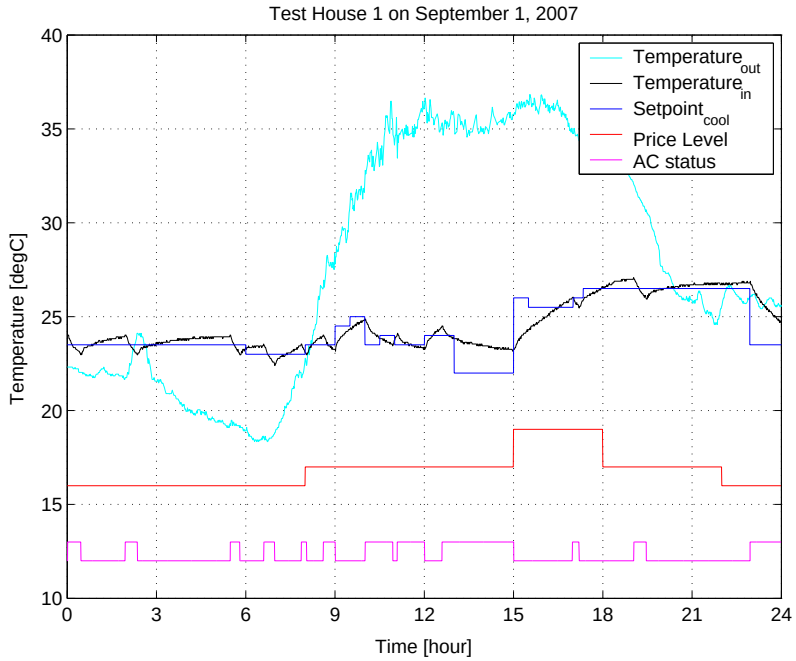


Figure 2.22: Temperature and Setpoint Profile for House 1 with Learned Parameters

or lower internal gain, or greater insulation than house 1.

In the last few days of the test, the internal house model was updated with learned parameters. The parameters were calculated from the recorded data in the remote database. According to the learned parameters, house 1 had less insulation, much higher internal gain, and slightly higher AC capacity than the default internal house model. However, due to a calculation error at that time, a lower internal gain was used for the test. Fig.2.22 shows the result. The main difference is the second precooling setpoint: a one degree higher setpoint (22C) was chosen when the outdoor temperature was 2C cooler. This implies that the learned parameters may provide a better prediction than the model with defaults. However, the first precooling setpoint was far from a reasonable prediction, thus requiring more analysis.

Chapter 3

House Dynamic Signature Learning through Data Analysis Techniques

In Chapter 2, optimization and control strategies in the DREAM controller were discussed, and it was mentioned that their success primarily depends on how accurately the indoor temperature can be predicted. A whole process in which thermal and physical characteristics of the house and HVAC equipment are identified and then indoor temperature is predicted is defined as *house identification*.

In order to identify the thermal characteristics of a specific house, the controller needs accurate and detailed information regarding the house. If an occupant already knows the house location (longitude and latitude), size, structure, orientation, air-conditioner size, and so on, the controller may predict the house dynamic signature from the given information. In commercial buildings, building engineers usually build thermal models and control their thermal behaviors based on the information. However, in residential buildings, access to that kind of information is relatively limited and, even with the information, it is impossible to make the thermal model of an individual building.

Learning is one approache to define characteristics of a system. While the characteristics of a commercial building can be derived directly from the physical properties by an engineer,

the controller for a residential building should learn the characteristics from the limited measurement by itself. Although the system identification by learning tends to be time-consuming and possibly less accurate, it is more flexible (adaptive) to various systems and environments. The house identification techniques that this chapter deals with should be understood as one algorithm that can be adopted by any autonomous energy management system for residential buildings and thus should not require high computational load.

The identification process amounts to repeatedly selecting a model structure, computing the best model in the structure, and evaluating this model's properties to see if they are satisfactory. The cycle can be itemized as follows:

1. Design an experiment and collect input-output data from the process to be identified.

2. Examine the data. Polish it so as to remove trends and outliers, and select useful portions of the original data. Possibly apply filtering to enhance important frequency ranges.

3. Select and define a model structure (a set of candidate system descriptions) within which a model is to be found.

4. Compute the best model in the model structure according to the input-output data and a given criterion of fit.

5. Examine the obtained model's properties

6. If the model is good enough, then stop; otherwise go back to Step 3 to try another model set. Possibly also try other estimation methods (Step 4) or work further on the input-output data (Steps 1 and 2).

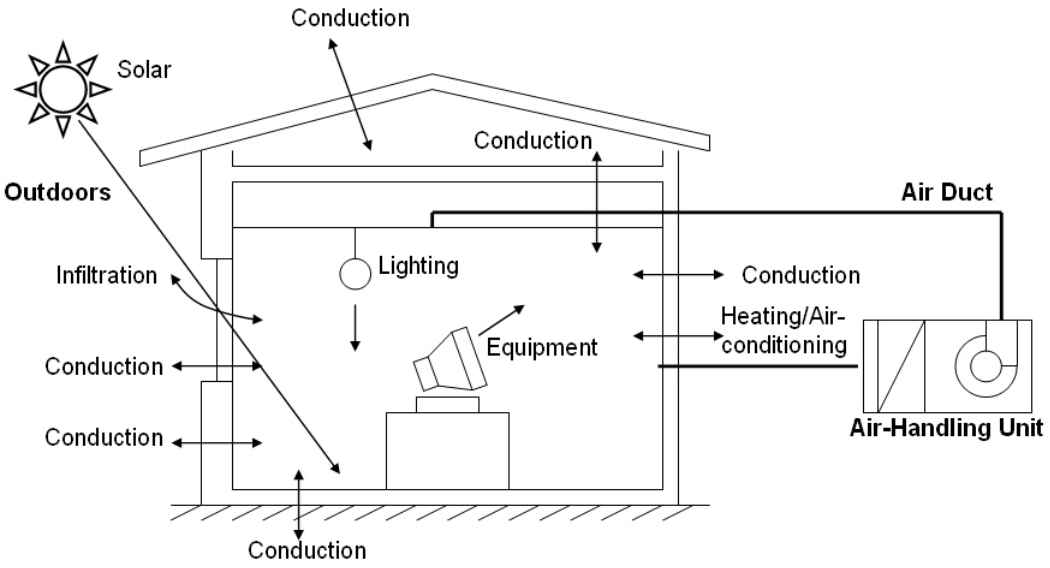
3.1 Characteristics of a Residential Building

3.1.1 Thermal properties

The envelope of a building acts as a passive climate modifier to help maintain an indoor environment that is more suitable for habitation that the outdoors. However, the building envelope alone responds to the outdoor climate in any case, and it can hardly ensure that the indoor environmental conditions will always be comfortable to the occupants, or be suitable for the intended purposes of the indoor spaces, especially in harsh outdoor conditions. Nowadays, an unfavorable indoor condition in any type of building is controlled by active means of environmental control such as central heating, ventilating and air-conditioning (HVAC) systems. It is the reason why the HVAC system should be considered as one element that defines the thermal characteristics of buildings.

Thermal characteristics of a residential building differ from those of a commercial building due to structure, size, materials, and HVAC systems. One study showed that building type determined the effectiveness of the envelope's thermal insulation on the thermal performance of buildings. It concluded that the impact of thermal insulation is more significant in the performance of a Skin-Load Dominated (SLD) building such as most residential buildings and small commercial building while it is not obvious that the internally generated heat in an Internal-load Dominated (ILD) building can be trapped more effectively with better insulation [20].

Disregarding the other differences between a residential building and a commercial building, the heat and mass transfer processes inside are very similar. Heat and mass transfer that would take place in typical residential buildings is illustrated in Fig.3.1. Although residential houses have various structural characteristics, a one-story house with a single room is sufficient to explain the heat transfer process. The room is separated from the outdoors by an external wall and a window, and above and below by a ceiling and a floor slab. The room is equipped with a HVAC system that would supply heating or cooling to



Source : Underwood, Chris P. and Yik Francis W.H. (2004) p.3.

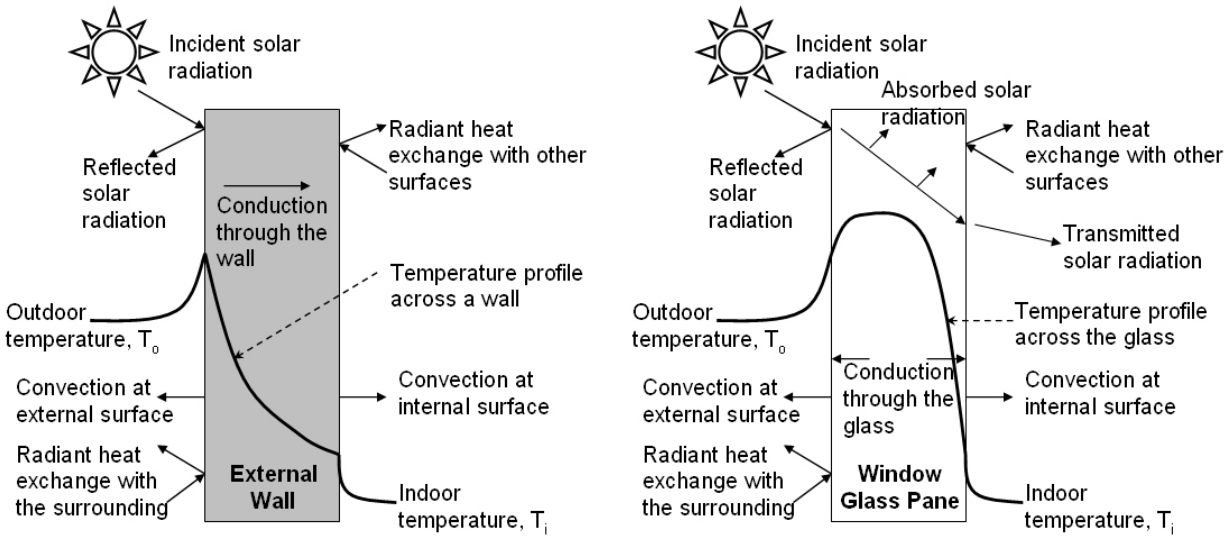
Figure 3.1: Heat and mass transfer processes involved in building energy simulation

the room by circulating air between the room and the air-handling unit via the supply and return air ducts.

The heat and mass transfer process include the following five elements.

- ◇ Conduction heat transfer through the building fabric elements, including the external walls, roof, ceiling, and floor slabs;
- ◇ Solar radiation transmission and conduction through window glazing;
- ◇ Infiltration of outdoor air;
- ◇ Heat and moisture dissipation from the lighting, equipment, occupants, and other materials inside the room; and
- ◇ Heating and cooling provided by the HVAC system.

The conductive heat transfer through an external wall or a roof is caused by the convective heat that the surface of the wall is exchanging with the surrounding air, and the radiant heat exchanges with other surfaces. The radiant heat exchange at the external side includes



(a) External Wall (b) Window Glass Pane

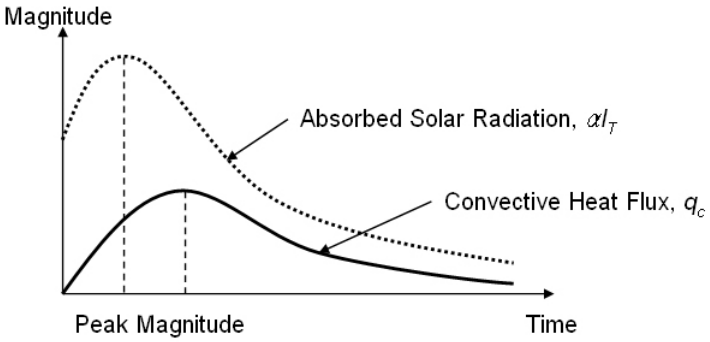
Source : Underwood, Chris P. and Yik Francis W.H. (2004) pp. 3-4.

Figure 3.2: Heat Transfer at an External Wall and a Window Glass Pane

the absorbed direct and diffuse solar radiation. However, the window glass allows part of the incident solar radiation to transmit into the indoor space. At the same time, some of absorbed solar radiation increases the temperature of the window itself and leads heat flow from the window to indoors and outdoors. The heat transfer at the external wall and a window is illustrated in Fig.3.2.

In some cases, the effect of absorbed solar radiation and outdoor to indoor temperature difference is treated together by using an equivalent outdoor air temperature, called *sol-air temperature*. The sol-air temperature causes the same amount of conduction and convection heat flow. In a similar way, *environmental temperature* is used to account for the combined effects of the convective heat transfer from the internal surface to the room air and the radiant energy gain at the surface.

The transmitted solar radiation is a component of cooling (or heating) loads in an indirect way. The temperature of the internal surfaces increases only after the transmitted solar radiation is fully absorbed. The increased surface temperature causes the conductive heat flow from the surfaces to the room air. Therefore, this cooling load differs in magnitude



Source : Underwood, Chris P. and Yik Francis W.H. (2004) p.5.

Figure 3.3: Radiant Heat Gain and the Resultant Cooling Load

and in the time of occurrence of its peak value from those of the radiant heat gain. The inconsistency between the two is shown in Fig.3.3.

The thermodynamic state of the air in the room varies with the air movement caused by pressure differences between the room, the adjoining rooms, the outdoors, and the moisture experienced by the room air, air transport into or out of the room, and heat and moisture gains by the HVAC system. Therefore, these heat transfer processes would need to be modeled for the accurate prediction of the indoor air condition or humidification or dehumidification required for maintaining the indoor air temperature at the desired setpoint [21].

3.1.2 Overview of Modeling for Energy in Buildings

While many energy modeling methods for buildings have been suggested and validated based on the knowledge obtained from experiences and fundamental laws, there have been dramatic improvements in the modeling methodology primarily due to the computer in the past 30 years.

Throughout most of the twentieth century, the thermal response of buildings has been calculated with many assumptions to simplify the calculation. For example, the outdoor temperature, wind speed and direction, and indoor temperature were considered static, and the building was assumed to be always unoccupied. While these assumptions enabled

	House 1	House 2
Location	Antioch, CA	Bay Point, CA
Structure	1700 square foot two-story stucco house	1500 square foot one-story house
Construction Year	1991	1984
Thermostat	White Rodgers programmable thermostat.	manual setback Honeywell Chronotherm.
HVAC System	Carrier split system air conditioner/furnace, with supply grilles in the floor throughout the house	General Electric split system air conditioner/furnace, with supply grilles in the ceiling
Occupants	Occupants are working at home. Normal setpoint in both daytime and nighttime is 74F and 79F when unoccupied. Participant opens up the windows (upstairs) at night and closes during the day.	Occupants are normally out of the house during the day. Normal setpoint in daytime is 70F and lowered or off in nighttime. The participant opens up the house at night; two windows are opened during the day as well.
Mote Deployment	Total 15 motes. A mote outside is under the southeast eave of the roof.	Total 14 motes. A mote outside is under the southeast eave of the roof.

Table 3.1: Description of Two Field Test Houses

the engineer to predict the thermal behavior very roughly (it might be useful to calculate heat load in a worst case), it could not explain the impact of the thermal capacity of the building envelope material or the dynamics of the indoor temperature with respect to a varying outside climate. In the later part of the twentieth century, more factors such as lagging effect of solar heat transfer through walls and windows began to be considered.

Since the 1970s, digital computers have been used as a main computational tool, which means that the building energy modeling is free from static boundary conditions. In this period, two main approaches were commonly used. One is the numerical method based on the governing Fourier equation using finite difference methods, and the other is an analytical method in which response factor is derived from the time response to a unit pulse. Common practice in North America is to use the analytical method due to the

computational inexpensiveness.

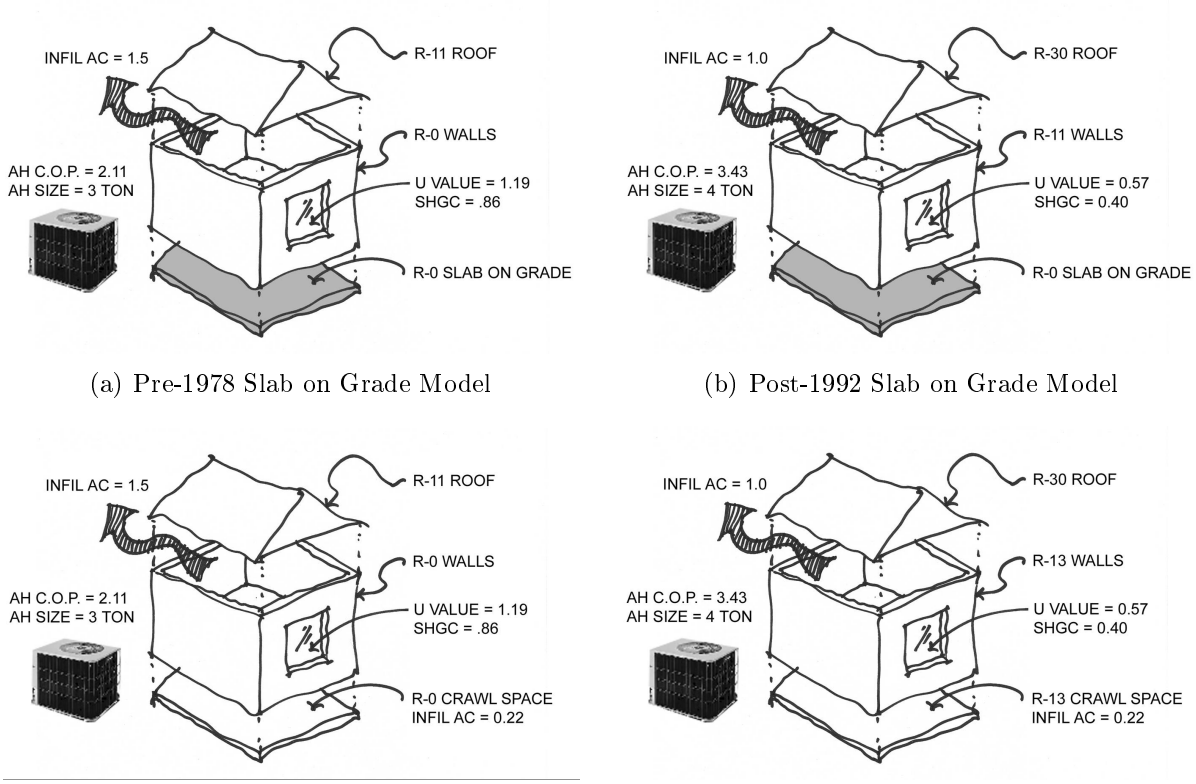
In the 1980s, the first commercial codes for design calculation using steady-state methods of energy in buildings became available. Two main methods, Response Factor method and Finite-difference method, were still dominant during this period. At the latter part of this decade, HVAC system modeling was introduced, and the moisture sorption in building fabric elements was treated in some codes.

In the 1990s, much research was focused on the neglected problem of fluid flow in and around buildings and the first computational fluid dynamics (CFD) codes for the case dealing with the low Reynolds number flow associated with building ventilation. However, the high computational demand of most CFD problems provoked people to doubt the embracing and comprehensive description in the building modeling method. Accordingly, the current modeling became more specified and simplified in accordance with the purpose and application.

Throughout this chapter, several methods are introduced. It should be clarified that the goal of those methods is not to include all features in a building but to predict indoor temperature within an acceptable error range with limited computational effort.

3.2 House Description for Algorithm Test and Validation

The house identification algorithms in this chapter were suggested and developed with different research purposes and testing environments. The following simulation houses and real houses were used in both algorithm development and validation. Different aspects in these houses allow algorithm developers to easily find problems or limitations of the suggested algorithm as well as consider various phenomena in a development stage.



(a) Pre-1978 Slab on Grade Model

(b) Post-1992 Slab on Grade Model

(c) Pre-1978 Crawl Space Model

(d) Post-1992 Crawl Space Model

Source : Hand-drawing by Kyle Konis, the Department of Architecture, University of California at Berkeley.

Figure 3.4: Four MZEST Models Based on Typical California Houses

3.2.1 Multi-Zone Energy Simulation Tool (MZEST) Houses

As an infrastructure of the DREAM system, the DREAM controller needed simulation houses to validate its control strategies prior to a field test. The Multi-Zone Energy Simulation Tool (MZEST) was developed to simulate the energy use and thermal response of houses. MZEST originated from the simulation code, CNE (California Non-residential Engine), which the energy simulation software distributed by the California Energy Commission used for demonstrating compliance with state residential Title 24 energy standards.

In order for MZEST to simulate thermal response and energy use, certain input parameters such as the specification of a weather file, designated house construction information, and HVAC on/off status are required. The weather file is hourly climate data in the form of TMY2 (Typical Meteorological Year) and the house construction parameters are speci-

fied in the input files. House size, construction material properties, window specification, infiltration rates, and HVAC equipment size are determined in this file. The HVAC on/off status is transferred from the DREAM controller in the form of XML (Extensible Markup Language). Two main outputs are temperature of each zone and energy use of the house. MZEST reports these outputs every sampling time in the form of XML. Although the default sampling time is five minutes, it can be decreased to one minute without losing accuracy.

The initial MZEST model is based on a real single-story house in Moraga, California. The simulation results were validated against measured indoor zone temperature using on-site outdoor weather condition[22]. However, to cover most typical houses in California[1], four additional houses were designed. In Fig.3.4, the characteristics of the four houses are summarized.

These houses are categorized with respect to construction year and foundation type. Roughly two-thirds (67%) of the occupied existing California housing stock was constructed before the first Title-24 energy standards took effect in 1978[23]. These old houses have relatively poor insulation with single-pane windows and air-conditioning and heating efficiency typical of the 1970s. In contrast, a post-1992 model represents a generic house meeting the minimum for Title-24 compliance (insulated envelope and double-paned windows). In addition, since the thermal mass plays in the attenuation of heating load, both a crawl-space and slab-on-grade foundations were applied as well.

3.2.2 Test Houses for the DREAM System

Two house owners in Northern California voluntarily provided their houses as a test bed for the DREAM system (Fig.3.5). Data acquisition frequency was determined by the sensor type and location. The mote that was located outside measured weather information every three minutes, and indoor temperature and relative humidity were measured every one

[1]While MZEST can be used for simulation of any house under diverse climate zones, the initial development was designated to simulate residential buildings in California.

(a) House 1 (House ID: C1)

(b) House 2 (House ID: C2)

Figure 3.5: Test Houses in Northern California

minute. The power information was updated every second. Occupancy information was updated whenever the status changed. Among the data from more than 50 sensors in each house, outdoor temperature, solar radiation, indoor temperature in the living room, and HVAC status were mainly used for house identification.

House 1 (House ID: C1)

House C1 is a $1700 ft^2$ two-story stucco house built in 1991. The living room has a cathedral ceiling that is open to the stairs and upstairs hallway and four bedrooms are located upstairs side by side. Most of the windows face east or west. The HVAC system is a Carrier split system air conditioner/furnace, with supply grilles in the floor throughout the house. Three ceiling fans are controlled manually in the living room, kitchen and master bedroom.

There were two occupants who work at home. One dog is inside most of the time. The owner normally keeps the thermostat set to 74F all the time, but he sets the thermostat to 79F when he leaves. He opens the windows at night and closes them during the day.

House 2 (House ID: C2)

House C2 is a $1500 ft^2$ one-story house built in 1984. One ceiling fan continuously runs in the family room. The HVAC system is a General Electric split system air conditioner/furnace,

with supply grilles in the ceiling. The house has an attic fan and two skylights in the roof.

There were two occupants who are normally out of the house during the day, but during a portion of the test were at home. The participant decides whether or not to use the air conditioning from the weather forecast. The setpoint during the day is 70F, and lowered in the evening, and turned off at night. If the weather is hot, the setpoint is 68F and 70F at night. The participant opens up the house at night and two windows are opened during the day as well.

3.2.3 Test Houses for the EcoFactor Thermostat System

EcoFactor, a start-up company in SunnyVale, CA, had co-worked with the DR thermostat research group at University of California, Berkeley for six months. This company deployed their first thermostat system in July, 2007 and started collecting information that is necessary for the advanced thermostat control. The following houses, located in Minnesota and Adelaide, are their test houses, and the data used in this chapter came from their own database.

Houses in Minnesota (House ID: M1 - M12)

Twelve houses in various locations in Minnesota were involved in the study. Data from these houses were uploaded to the database in November, 2007, and winter data were primarily used for analysis. The house size ranged from 946 to $2204 ft^2$ with an average of $1330 ft^2$. It should be noted that the following house information is based on the homeowner's reports. Therefore, some houses do not have complete information. The construction year ranged from 1937 to 1987, and the average age was 42 years. All houses are reported to have insulation in both the attic and the walls. The primary house structure in 10 houses is wood, and in one is masonry. The last house was not identified. Six houses were built on slab, two have basements, and two have raised perimeter. Ten of the houses are traditional detached houses, one is half of a duplex, and one is a town home. Due to the severe winter

House ID	Size (ft^2)	Construction	House ID	Construction	Comments
M1	1501 - 2500	1985	A1	1890	
M2	1501 - 2500	2003	A2	1999	structurally identical
M3	2501 - 3500	1987	A3	1999	
M4	1501 - 2500	1937	A4	1999	
M5	1501 - 2500	1968	A5	1999	structurally identical
M6	1000 - 1500	1951	A6	1999	
M7	1501 - 2500	1984	A7	N/A	
M8	less than 1000	1962	A8	N/A	
M9	1501 - 2500	1956	A9	N/A	
M10	1501 - 2500	1958	A10	N/A	
M11	1000 - 1500	1957			
M12	1501 - 2500	1957			
C3*	2860	1974			

*C3 is an additional test house in California

Table 3.2: Summary of the Test Houses in Minnesota and Australia

climate, nine of the houses have double-glazes or better windows. The location, size, and construction year of the houses are summarized in Tab.3.2.

Each home has one single-stage natural gas furnace. Approximately half of the houses are equipped with less than five-year-old furnaces and the rest of them are equipped with 5-10 year-old ones. Two of them have more than 10 year-old furnaces. For the study, the existing thermostats were replaced with off-the-shelf devices manufactured by Proliphix, Inc.

Among nine homeowners that had programmable thermostats, three reported changing the programming often, five reported adjusting the programming very rarely (once or twice per year), and the rest never changed it. Five houses are occupied most of the day, and four houses are mostly unoccupied during the day.

The sampling time (data acquisition and uploading time) of the thermostat system is one minute for all Minnesota houses. Indoor temperature is measured by the thermistor of the thermostat, and other control properties such as setpoints, HVAC status, and HVAC mode are uploaded to the database through wireless networking. Weather information (outdoor temperature, humidity, cloud condition, and wind speed / direction) is obtained

Figure 3.6: Five Houses on the Same Street in Adelaide, Australia

in the server side directly from the website of the local weather stations. Sampling time in the weather stations ranged from 10 minutes to one hour.

Australia Houses (House ID: A1 - A10)

In February 2008, EcoFactor deployed their thermostats in ten houses located in Adelaide, Australia. These houses provided thermal behavior of the houses mostly in a summer period. Five houses (house ID: A2, A3, A4, A5, and A6) were constructed in the same recent development. Among them, three houses (A2, A3, and A4) are structurally identical. (They have the same floor plan, construction methods, etc.) The other two (A5 and A6) are identical to one another. All five houses are located on the same street shown in Fig.3.6. The house information is summarized in Tab.3.2.

Each house has one single-stage AC-heat pump unit. Three of the houses have less than 5-year-old AC units and five have less than 10-year old units. One interesting thing is that the ducting in every house has in-line motorized dampers controlled by switches, which makes zone control possible.

All homeowners previously had programmable thermostats and had an experience programming them at least once. Six of them adjusted the programming once or twice in the previous year and the rest never changed it. Seven houses are usually occupied most of the day and three houses are unoccupied during the day.

The data acquisition and updating procedure are exactly the same as that in the Min-

nesota houses. The only difference is that the weather information for the 10 houses comes from one weather station, and the sampling time in the weather station is one hour. Therefore, the outside temperature (one major piece of information for dynamic signature learning) may not be reliable, and additional filtering will be required.

3.3 1st Order Physical Model

3.3.1 Model Development

As a first candidate, identification through a complex physical model was not recommended. A complicated house model with many unknown parameters would certainly provide better prediction in that the model might be able to represent even nonlinear characteristics of the house thermal behavior. MZEST that is used as an algorithm validation tool in this study is a good example of complex models. However, this type of model requires a large amount of data and high computational load to identify the unknown parameters. A simplified house model, thus, can streamline the learning process by minimizing the number of unknown parameters. A limit was imposed on the properties of the proposed model: It should be low order, have less than five unknown parameters, and require a small amount of data for prediction. The initial internal model has the following form:

$$T_{in}(t+\Delta t) = function(T_{in}(t),\, T_{out}(t),\, RH_{in}(t),\, I_{global}(t),\, Occupancy(t),\, etc.) \qquad (3.1)$$

where $T_{in}(t+\Delta t)$: indoor temperature at time $t+\Delta t$

$T_{in}(t)$: indoor temperature at time t

$T_{out}(t)$: outdoor temperature at time t

$RH_{in}(t)$: indoor relative humidity at time t

$I_{global}(t)$: global radiation at time t

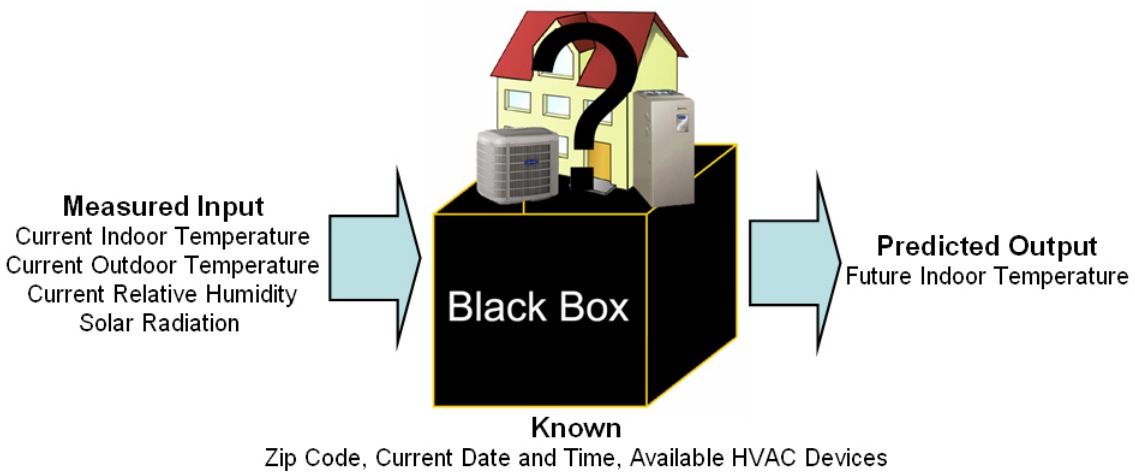
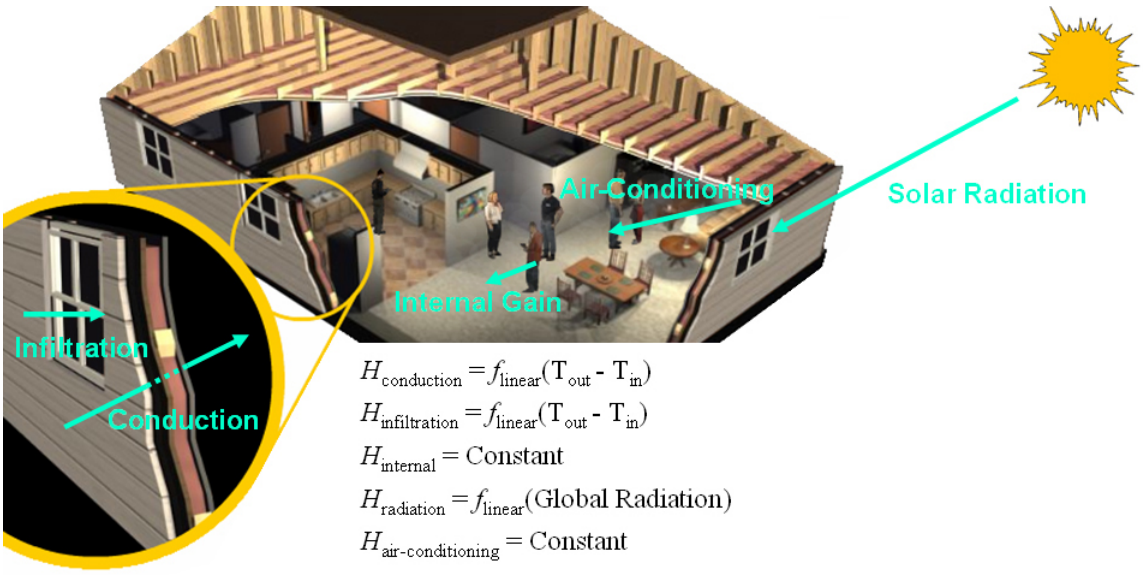


Figure 3.7: Inputs and Outputs of the Internal House Model

$Occupancy(t)$: house occupancy status at time t

An example of how the internal house model works is shown in Fig.3.7. Upon installation of the DREAM system, the house characteristics including the characteristics of HVAC equipment are unknown—the proverbial black box. The indoor temperature, outdoor temperature, relative humidity and solar radiation are sensory data inputs. Other inputs such as zip code and available HVAC devices will be entered by the occupant. The zip code will provide general location of the house, which is used to predict global radiation. Current time and date will be automatically configured. The output of the model is the predicted indoor temperature.

To simplify the model, all residential buildings were assumed to be a single-room controlled by a single HVAC unit. The model considers five heat sources: conduction, infiltration, internal gains, solar radiation, and heating or air-conditioning (Fig.3.8). Heat balance caused by convective heat transfer was ignored in this house model. Heat gain through the external walls, windows, doors, ceilings, and floors by conduction and infiltration can be expressed by the following equations.



Original graphics: www.ce.utexas.edu/bmeb/scenarios/heatingCooling.cfm

Figure 3.8: Five Heat Sources in the 1st Order Model

$$H_{conduction} = AU(T_o - T_i) \tag{3.2}$$

$$H_{infiltration} = sdnV(T_o - T_i) \tag{3.3}$$

where H: heat gained (W)

A: area of exposed surface (m^2)

U: overall coefficient of heat transmission (W/m^2K)

s: specific heat capacity of air (kJ/kgK)

d: density of air (kg/m^3)

n: number of air changes (s^{-1})

V: volume of room (m^3)

T_i: inside air temperature $(°C)$

T_o: outside air temperature $(°C)$

In Eq.3.2 and 3.3, other parameters except $T_i$and T_o are uniquely determined by the house or assumed to be constant. This implies that heat transfer due to conduction and

Still Air												
Air Temperature ($°C$)	10	12	14	16	18	20	22	24	26	28	30	32
Sensible Heat (W)	136	126	115	106	98	92	85	77	69	58	47	33
Latent Heat (W)	21	21	21	21	23	27	33	41	49	60	69	81
Total (W)	157	147	136	127	121	119	118	118	118	118	116	114
Moisture (g/hr)	31	31	31	31	34	40	48	60	73	88	102	120
Air Velocity 1m/s												
Air Temperature ($°C$)	10	12	14	16	18	20	22	24	26	28	30	32
Sensible Heat (W)	152	142	131	122	112	104	97	88	81	69	55	38
Latent Heat (W)	19	19	19	19	19	20	25	32	38	49	61	77
Total (W)	171	161	150	143	131	124	122	120	119	118	116	115
Moisture (g/hr)	28	28	28	28	28	29	36	47	57	73	89	114

Source: Porges, F (1995) p.116[24].

Table 3.3: Heat Emitted by Human Body (Light Office or Domestic Work)

infiltration can be expressed as a linear function of temperature difference. The magnitude of an internal heat gain inside the house is affected by the number of people, their activity, and household equipment such as lights. Even from the same number of people, total body heat varies because a human body emits a different amount of heat depending on its surrounding air temperature and velocity (see Tab.3.3). The effect of solar radiation on inside air temperature is more subtle. Although heat induced by the solar radiation is calculated from a very simple equation (Eq.3.4), three parameters F, α, and I are dependent on many factors including surface material, type of shading, and orientation.

$$H_{radiation} = AF\alpha I \tag{3.4}$$

where H: heat gained (W)

A: area of exposed surface (m^2)

U: overall coefficient of heat transmission (W/m^2K)

F: radiation factor, proportion of absorbed radiation transmitted to interior (W/m^2)

α: absorption coefficient, proportion of incident radiation absorbed

I: intensity of solar radiation striking the surface (W/m^2)

To reduce unknown parameters in the internal model, the following four assumptions were applied. Heat gained by conduction and infiltration is a linear function of outdoor to indoor temperature difference; internal gain is constant; heat gained by solar radiation does not have a time lag and can be written in a linear function of the global radiation value; and efficiency of a heating or an air-conditioning unit is not dependent on the indoor and outdoor temperature, which implies that the unit supplies constant heat to the room. Based on these assumptions, the future indoor temperature is calculated by

$$\frac{T_{in}(t+\Delta t) - T_{in}(t)}{\Delta t} = \alpha \times (T_{out}(t) - T_{in}(t)) + \beta + \gamma \times I_{global}(t) + \delta \times (AC\ status) \quad (3.5)$$

where α: coefficient corresponding to conduction and infiltration heat transfer (hr^{-1})

β: coefficient corresponding to internal heat gain (K/hr)

γ: coefficient corresponding to the heat by solar radiation (Km^2/Whr)

δ: coefficient corresponding to the heat gained by air-conditioning (K/hr)

$I_{global}(t)$: global radiation at time t

Four parameters in Eq.3.5 determine the thermal behavior of a given house. However, these parameters would be possibly affected by time of day, season, or other factors because the effect of those factors was excluded by the assumptions. γ, for example, may be relatively small during a certain time period if the house is partially shaded by woods.. For this reason, the parameter learning needs to be recursive with recent data. To minimize seasonal effects as well as the variation of occupant's behavior pattern, recent three-day-long data were used for parameter learning.

The learning process starts from data categorizing. Instead of using the same data set for all parameters, different portions of the data were applied to different parameters. For α and β, the data corresponded to the period in which there was no air-conditioning event

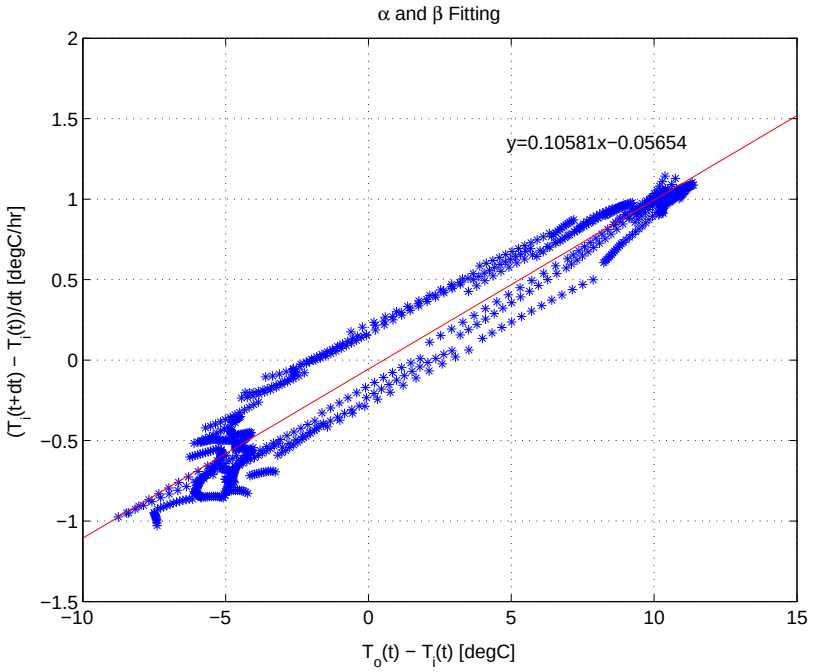


Figure 3.9: α and β Calculation through Least Squares (Pre-1978 Slab on Grade Model)

and ignorable global radiation. Most of the night-time data belongs to this category. Since the third and fourth terms in the right side of Eq.3.5 are ignored, only α and β remain, and then they can be easily calculated by linear least squares[2]. With the α and β obtained through the previous process, the next two parameters, γ and δ, are calculated from the rest of data corresponding to no air-conditioning and air-conditioning periods respectively. Fig.3.9 shows how the parameters are calculated from least squares. In this simulation, the outdoor temperature came from the TMY2 weather file for Sacramento, California [25].

By following the above process, four parameter sets with respect to four different MZEST houses were calculated. The sets are compared in Tab.3.4. It is interesting that each group in the same construction year has similar values while the parameters between two groups are considerably different. Improved insulation in the post-1992 houses seems to prevent the outside weather condition from affecting the inside air temperature. The

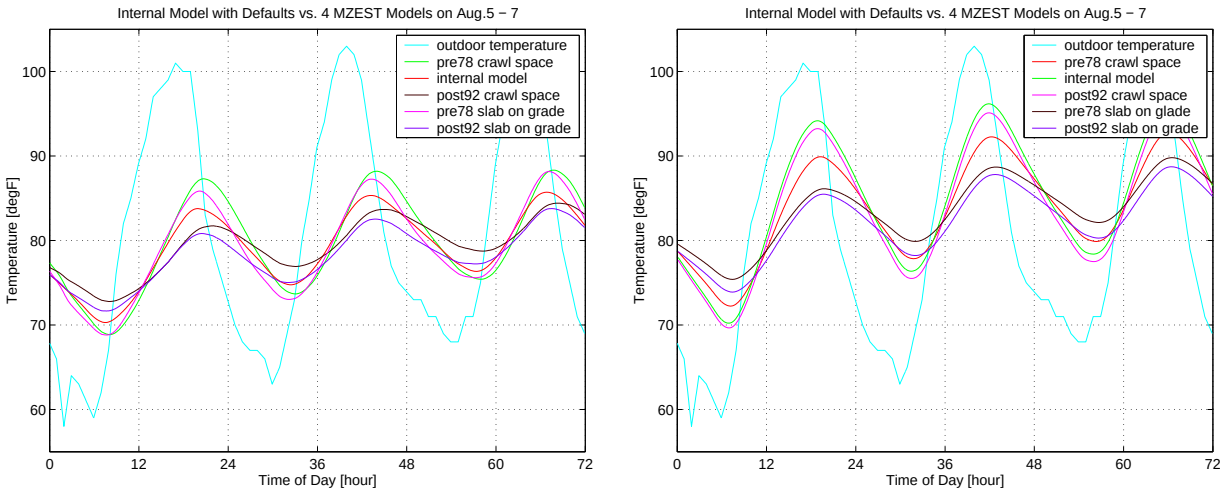
[2]The method of least squares, also known as regression analysis, is used to model numerical data obtained from observations by adjusting the parameters of a model so as to get an optimal fit of the data.(from Wikipedia)

MZEST Model	$\alpha \times 10^3$ (hr^{-1})	$\beta \times 10^3 (K/hr)$	$\gamma \times 10^3 (Km^2/Whr)$	$\delta \times 10^{-3}$ (K/hr)
Pre-1978 Slab on Grade	105.81	-56.54	0.94952	-16.591
Pre-1978 Crawl Space	104.82	-17.29	0.99593	-18.253
Post-1992 Slab on Grade	48.29	-31.49	0.37017	-27.003
Post-1992 Crawl Space	43.05	10.99	0.33565	-26.828
Default	73.37	-26.79	0.66296	-21.307

Table 3.4: Parameter Comparison

post-1992 houses are approximately two times less affected by the outside weather than the pre-1978 houses in that α value represents the sensitivity of the indoor temperature to the outside temperature. In addition, heat from the solar radiation tends to be relatively small in the post-1992 houses. The sizes of air-conditioning units in two groups are initially different. Old houses have 3 ton units and new houses have 4 ton units. (The size ratio is 1 to 1.3) Despite considering the size difference, heat loss by air-conditioning in the post-1992 houses is more significant. (The δratio is 1 to 1.6.) In contrast to our expectation, the data provided a negative β value, with the exception of one house. A negative internal gain in the house does not make any physical sense, but other factors that were not considered in this very simple model might affect the sign of β. Although this model does not fully explain the physics inside the house, it successfully identified the thermal characteristics of the house to some degree.

To guarantee reasonable out-of-box performance of the DREAM controller that can be deployed in any California house, preset parameters (defaults) of the internal model were also obtained. Since the defaults are based on the four MZEST models, they should be updated later. Nevertheless, they would help the controller to predict indoor temperature within an acceptable error range. In Fig.3.10, indoor temperature predicted with defaults is compared with the simulation results from the learned parameters associated with four different MZEST houses.



(a) No Air-Conditioning and No Solar Radiation (b) No Air-Conditioning but Solar Radiation

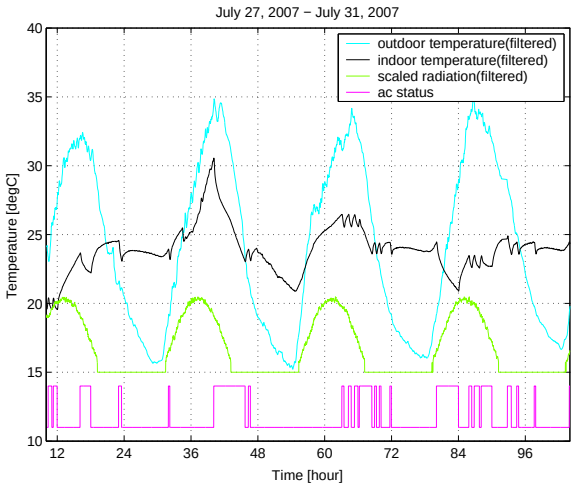
Figure 3.10: Default Parameters

3.3.2 Performance Evaluation

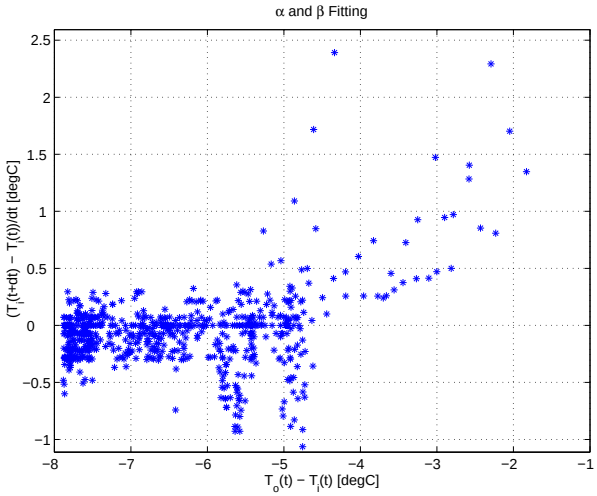
While the parameters from the MZEST houses tend to be constant regardless of time change, the parameters from the real house were seriously affected by disturbance. The consecutive five days from July 27 to July 31, 2007 were analyzed in Fig.3.11. In order to see consistency of the learned parameters, four α-β sets were calculated from the data obtained during each night time period under no air-conditioning. α and β are corresponding to the slope and y-intercept of the regression line respectively in Fig.3.11(c)(d)(e) and (f).

Except the second parameter set ($\alpha = 0.4539$ and $\beta = 2.0643$), the other three ($\alpha =$ 0.1533, 0.1946, 0.1471 and $\beta = 0.9841$, 1.3612, 0.9833) are close. Also, the signs of the β values, which were negative in the three out of four MZEST houses, are all positive here. However, even if the second case is excluded for parameter calculation, the scattered data in Fig.3.11(b) does not show the linear relationship of current outdoor/indoor temperature difference, $T_o(t) - T_i(t)$, and indoor temperature increment, $T_i(t + \Delta t) - T_i(t)$, clearly unlike the MZEST houses in Fig3.9. The shape of scattering is rather more exponential than straight.

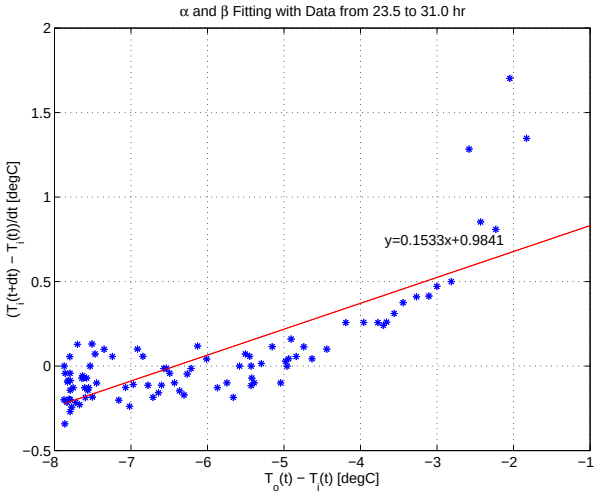
The anomaly in the second parameter set can be explained by comparing the temper-



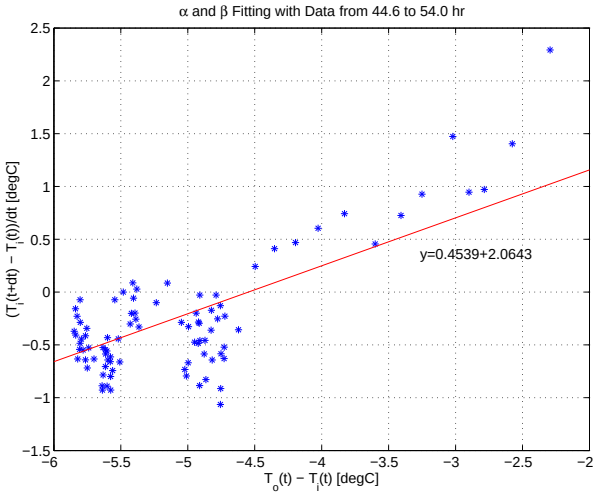
(a) Original Data Profile



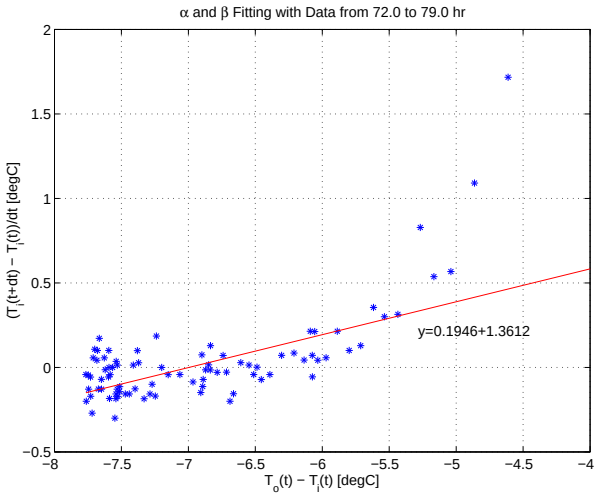
(b) Parameters Learned from 4 Days Data



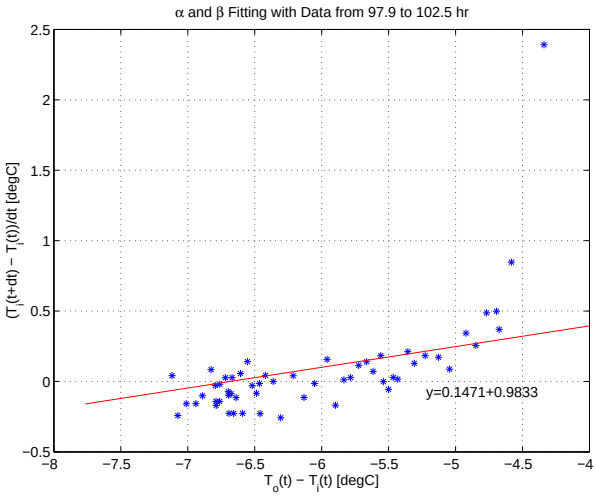
(c) Parameters Learned from the First Day Data



(d) Parameters Learned from the Second Day Data



(e) Parameters Learned from the Third Day Data



(f) Parameters Learned from the Forth Day Data

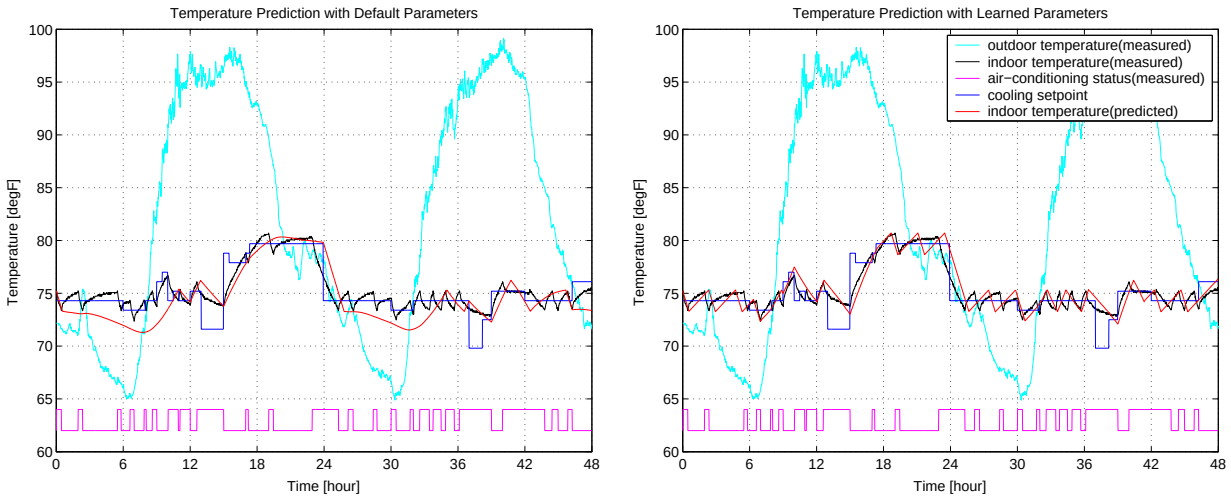
Figure 3.11: Parameter Consistency (Antioch, California from Jul.27-Jul.31)

ature behavior and air-conditioning status in 3.11(a). The weather conditions during the selected night periods are very similar. However, the indoor temperature change rate in the second night (44.6 - 54.0 hr) is considerably different from those in the other three nights. Therefore, it can be inferred that there was an unusual heat loss during the night, and it might be caused by the heat exchange between indoor and relatively cool outdoor through open windows. If the controller can compare parameters by itself, automatic fault detection or continuous commissioning would be possible.

Fig.3.12 shows how much improved the prediction quality is after parameter learning. Predicted indoor temperature in Fig.3.12(a) (red line) follows actual indoor temperature (black line) within an acceptable error range. However, especially in a morning period (from 0 to 10am), the AC on/off status of real indoor temperature and predicted temperature is totally different in that no AC on event has been predicted. Compared with predicted indoor temperature in Fig.3.12(a), predicted temperature in Fig.3.12(b) is much closer to the real temperature. Although the temperature slope during AC on period is not well matched due to the simplicity of the internal model, it can be seen that AC on/off cycle follows the actual on/off cycle pretty well. Imperfection of the suggested internal model is certainly the main reason for error. Even with a more complicated internal model, the prediction error still exists. The error would come from unexpected occupant behavior including opening windows and cooking. This would change infiltration and internal heat gain of a house. Therefore, if this detailed information could be provided, the prediction will be more accurate. Even without them, the current house learning algorithm is still promising.

3.4 Tabular Method

Now it is questionable whether simplified numerical models can describe complicated thermal dynamics inside a house. As noted, the non-linearity of the house behavior was assumed



(a) Prediction with Defaults

(b) Prediction with Learned Parameters

Figure 3.12: Indoor Temperature Prediction Before/After Parameter Learning (Antioch, California from Sep.1, 2007 – Sep.2, 2007)

to be linear due to complexity of calculation in the previous method. It is always a dilemma that both accuracy and simplicity of the prediction mode is required for the DREAM controller. At this time, a totally different approach is suggested, which starts from skepticism on numerical methods.

A tabular method does not fit the measured data to a predefined model. Instead, it reorganizes the data in a predefined structure. Fig.3.13 demonstrates how a tabular method is applied in house dynamic signature learning. Among many factors that affect indoor temperature fluctuation, outdoor temperature, solar radiation, and the on/off condition of the heating or cooling devices are most dominant. Indoor temperature should also be saved as a reference value in a prediction procedure. In Fig.3.13, three tables are used for an off mode, a cooling mode, and a heating mode respectively. This mode is automatically determined based on the real-time on/off status from the HVAC switch. Each table is filled with temperature change rates (degree/hour) in given outdoor and indoor temperature. The table size is determined by the reachable indoor and outdoor temperature ranges and their discretization measure. If solar radiation is also considered, each table will require one more axis for it.

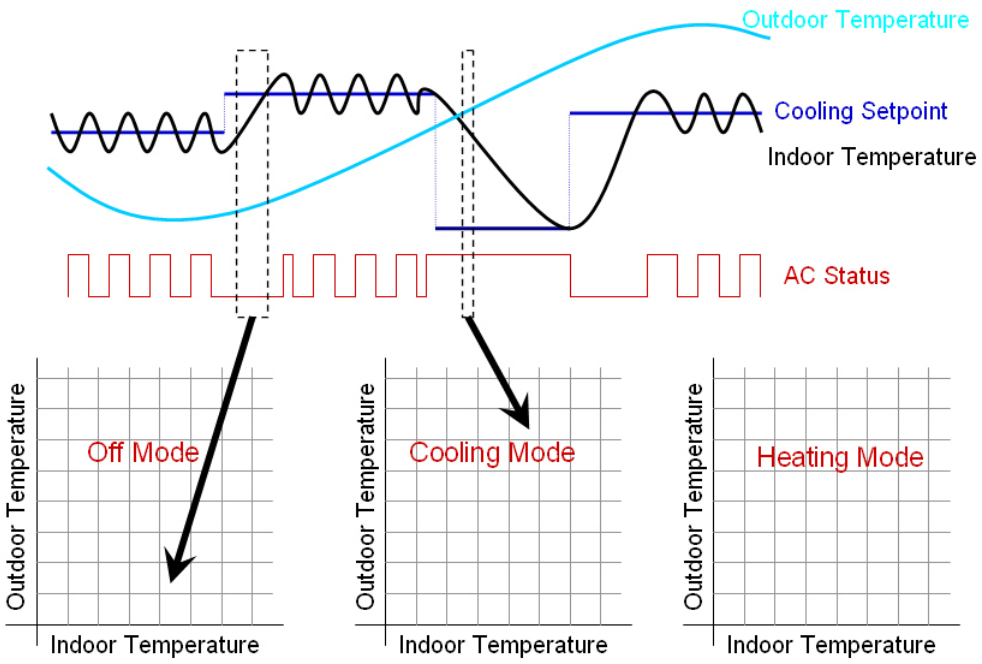
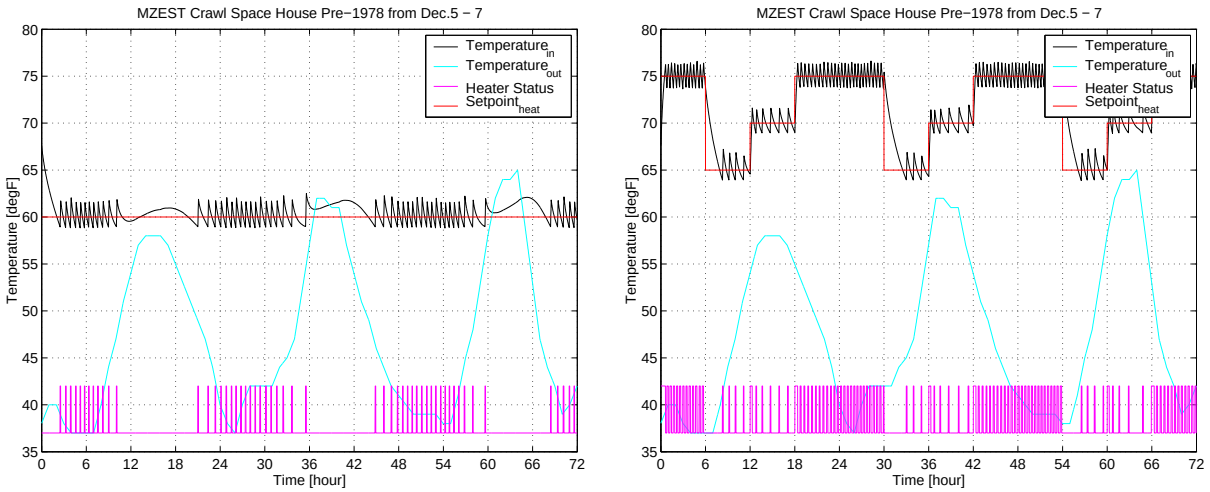


Figure 3.13: Schematics of the Tabular Method for Dynamic Signature Learning

If five-minute sampling time is assumed, temperature change rate is calculated every five minutes and this value is stored in the corresponding (off, heating, or cooling) table. The outdoor and indoor temperature at that moment determines the exact location where the value should be put in the table. For example, if indoor and outdoor temperatures are 70.0F and 82.0F at 2:30pm, and 69.5F and 82.1F at 2:35pm under air-conditioning, the temperature change rate is -6.0$^{\circ}F$/hr (-0.1 $^{\circ}F$/min) and the location is (82, 70) of the second table in Fig.3.13.

3.4.1 Instant Slope Method

In this method, an instant slope at every sampling time is calculated by linearizing three consecutive indoor temperature points. However, it is necessary to decrease the total number of reference conditions (outdoor temperature, indoor temperature, and solar radiation). While the diverse references may provide an opportunity to consider more delicate situations, it requires, at the same time, more CPU memory for the controller and increases the



(a) Constant Setpoint

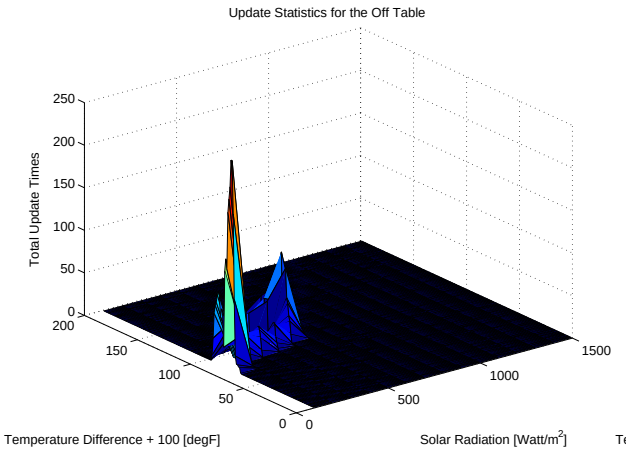
(b) Changing Setpoint

Figure 3.14: Two MZEST Simulations with Different Setpoint Setups

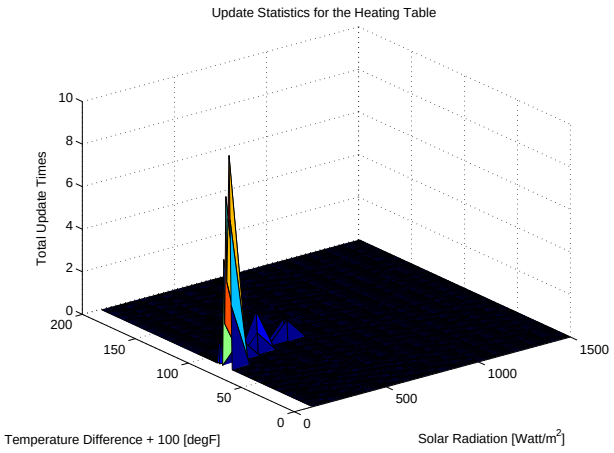
learning time[3]. Finally, the temperature difference between outside and inside is substituted for outdoor temperature and indoor temperature in that heat transfer is basically caused by the temperature difference. It is certain that some properties are definitely related to the respective value rather than the difference. Nevertheless, it would be a reasonable approach.

This method was applied to the one of the MZEST houses (pre-1978 crawl space house) from Dec. 5 to Dec. 7. For comparison, a fixed value and a changing value in a wide range were used as a heating setpoint. In Fig.3.14, two simulated indoor temperature profiles are shown. Fig.3.15 and Fig.3.16 show the correlation between a temperature change rate (slope) and either temperature difference or solar radiation. From the results, it could not be clearly shown how the solar radiation affects the indoor temperature change rate. However, the proportional relationship between the temperature difference and the temperature change rate was found both in the off mode and in the heating mode. The slope distribution with respect to the temperature difference in the off mode seems to be parabolic and that in the heating mode is approximately linear. One notable finding is that the distributed spots in Fig.3.16(f) are apparently divided into two separate lines while the

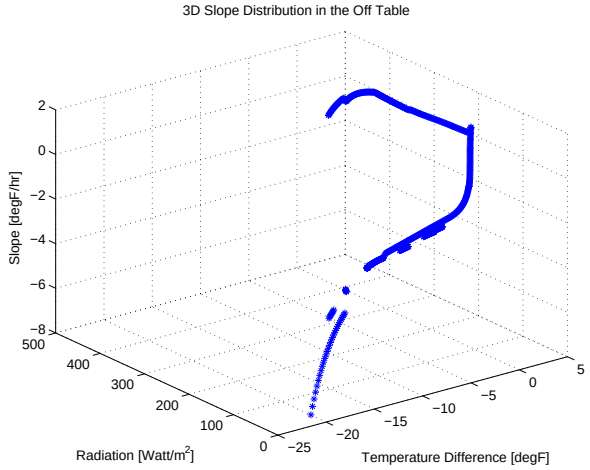
[3]It can be defined as a total time within which all tables are filled with a value.



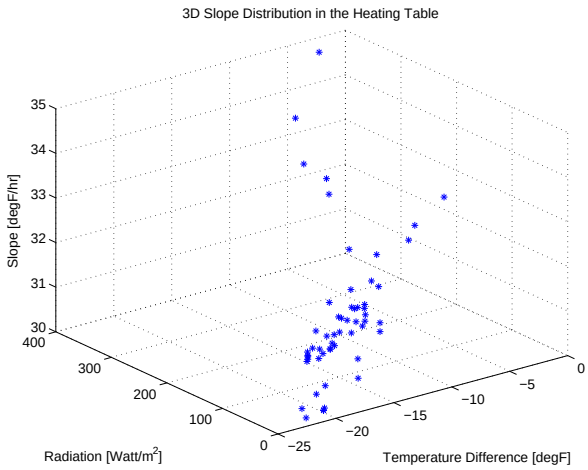
(a) Updata Statistics of the Off Table



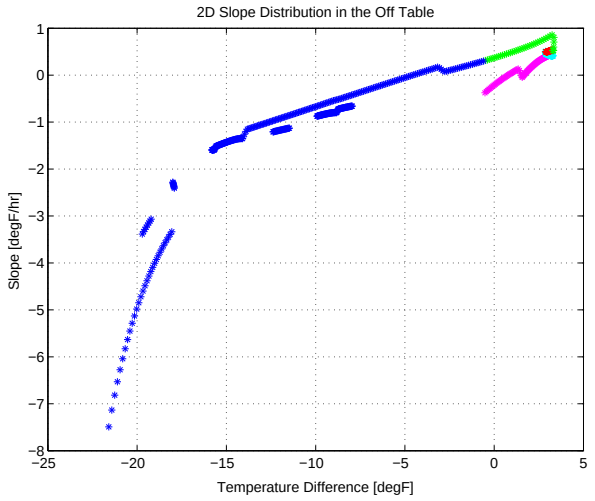
(b) Update Statistics of the Heating Table



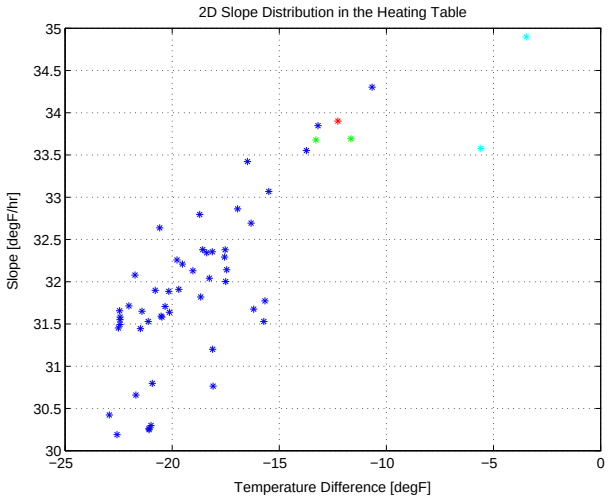
(c) 3D Slope Distribution in the Off Table



(d) 3D Slope Distribution in the Heating Table

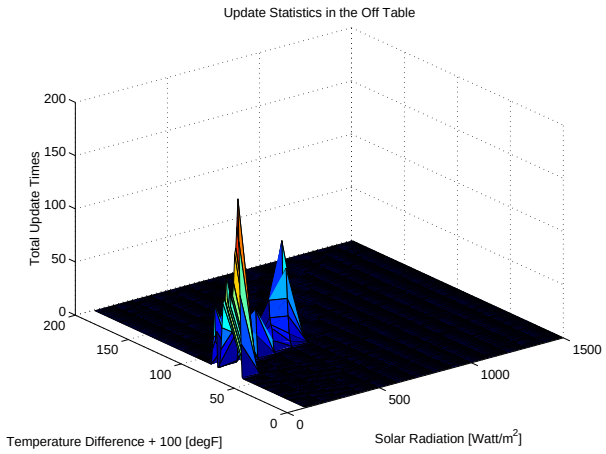


(e) 2D Slope Distribution in the Off Table

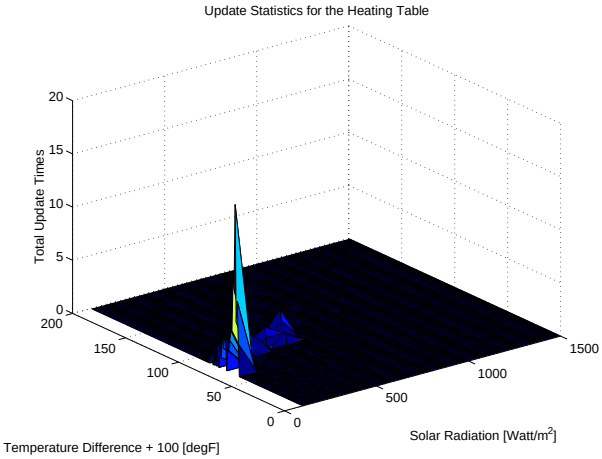


(f) 2D Slope Distribution in the Heating Table

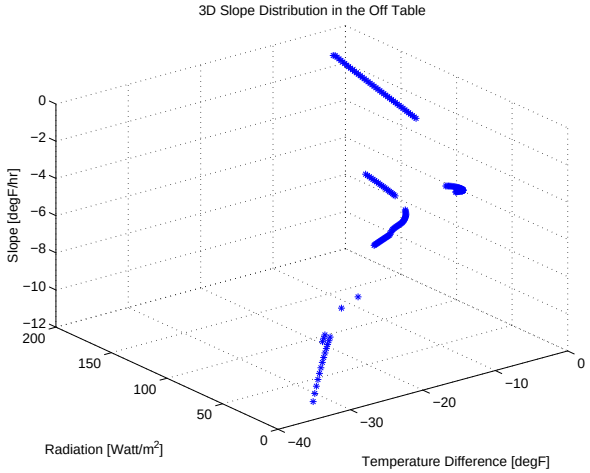
Figure 3.15: Correlation between Slope, Temperature Difference, and Solar Radiation under Constant Setpoint Setup (MZEST Pre-1978 Crawl Space, Dec. 5 - Dec. 7)



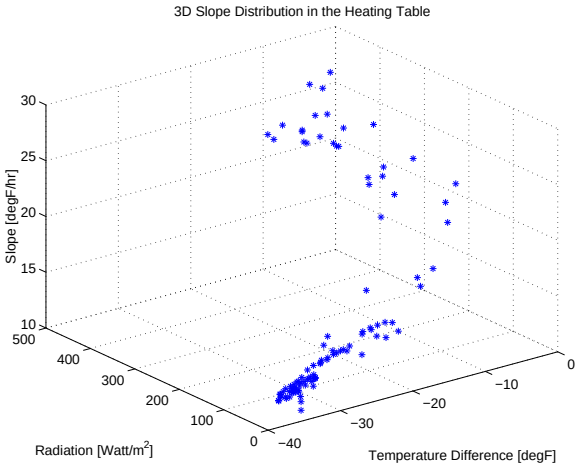
(a) Update Statistics of the Off Table



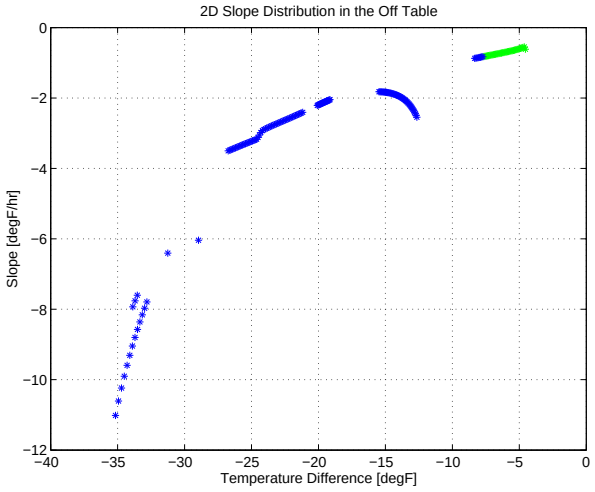
(b) Update Statistics of the Heating Table



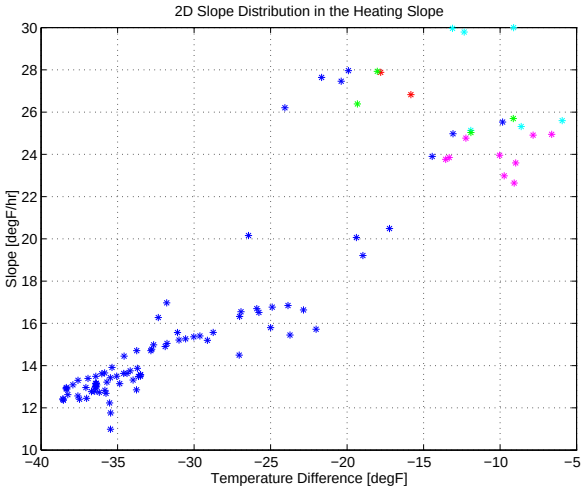
(c) 3D Slope Distribution in the Off Table



(d) 3D Slope Distribution in the Heating Table

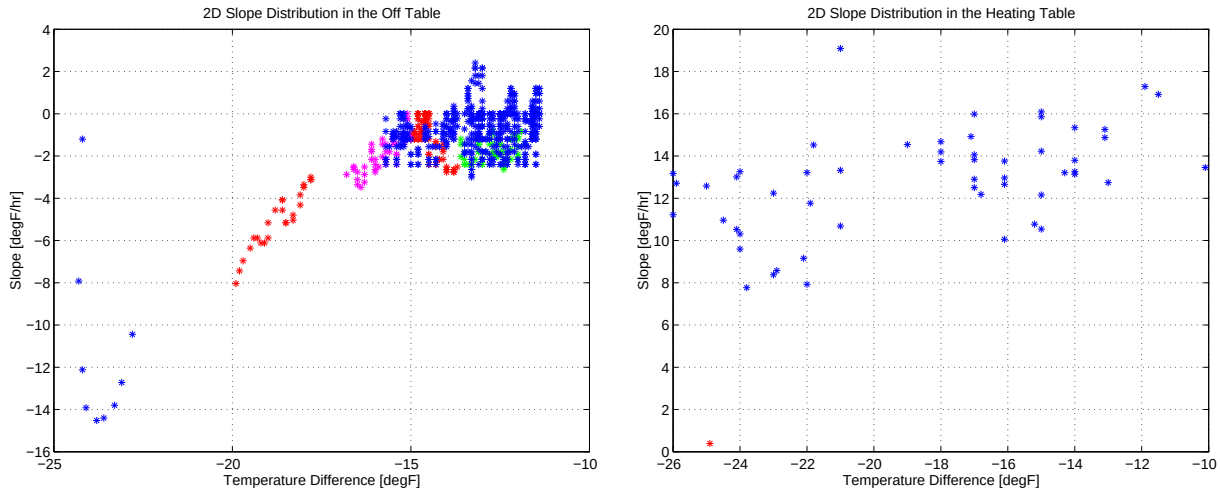


(e) 2D Slope Distribution in the Off Table



(f) 2D Slope Distribution in the Heating Table

Figure 3.16: Correlation between Slope, Temperature Difference, and Solar Radiation under Changing Setpoint Setup (MZEST Pre-1978 Crawl Space, Dec. 5 - Dec. 7)



(a) Slope in the Off Table

(b) Slope in the Heating Table

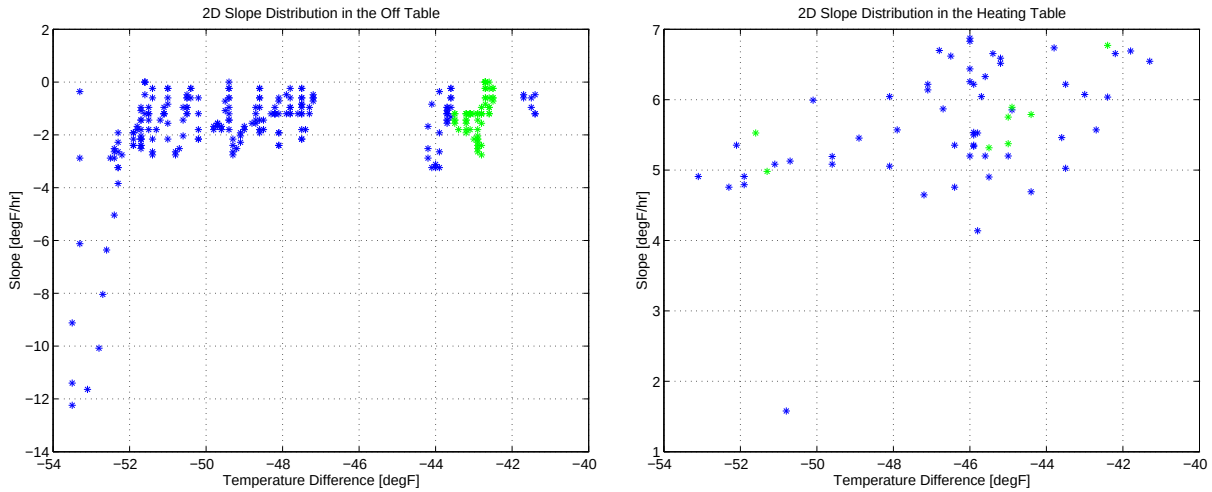
Figure 3.17: Slope Distribution of the House C3 (Dec. 28 - Dec. 31, 2007)

separation is less clear in Fig.3.15(f).

Fig.3.17 and Fig.3.18 are based on 4-day-long real data from the House C3 and M4 respectively. Different colors in the plots display different magnitudes of solar radiation. The slope distribution in the off table of the C3 house (Fig.3.17(a)) resembles closely the shape of MZEST simulation if the degree of concentration is not the matter. In the same standpoint, the heating slope approximates a straight line.

However, the proportional relationship is not always clear in the data from real houses like the case in Fig. 3.18. Only 4 out of 12 test houses in Minnesota showed the similar distribution pattern and 3 houses did not have any pattern in either the off or heating tables. The rest of the houses could not show any relationship between two variables in both tables.

While the indoor temperature that MZEST outputs is numerically calculated, the measurement from the real houses is easily disturbed by uncontrolled factors. These factors include inaccuracy of the indoor temperature sensor resolution due to rough resolution, information from a weather station that cannot reflect the actual outdoor temperature around the house, and unpredictable human behavior that might affect the thermal behav-



(a) Slope in the Off Table

(b) Slope in the Heating Table

Figure 3.18: Slope Distribution of the House M4 (Dec. 28 - Dec. 31, 2007)

ior of the house. Due to the continuity in temperature changing, the slopes within a very short period should be close if there is no disturbance. For example, two slopes at 6:21pm and 6:22pm may not be very different. However, even in the constant outside temperature condition within a relatively short period, actual indoor temperature measurements are noisy. As a result, two totally different slopes can be calculated for the same slot in the table.

3.4.2 Two-Slope Method

Even though the previous results from the MZEST and real houses show possibility in using a slope method as a dynamic signature learning method, a more consistent method that is less sensitive to disturbances is required for better prediction. When the indoor temperature of the Minnesota houses was examined closely, it was found that the temperature change rate (or slope[4]) was also affected by uneven heat balance. When the heater was repeatedly turned on and off around the heating setpoint, the slopes in both off and heating modes were affected by the outdoor temperature. With roughly constant outdoor temperature,

[4]A temperature change rate is the slope in a time-temperature plot

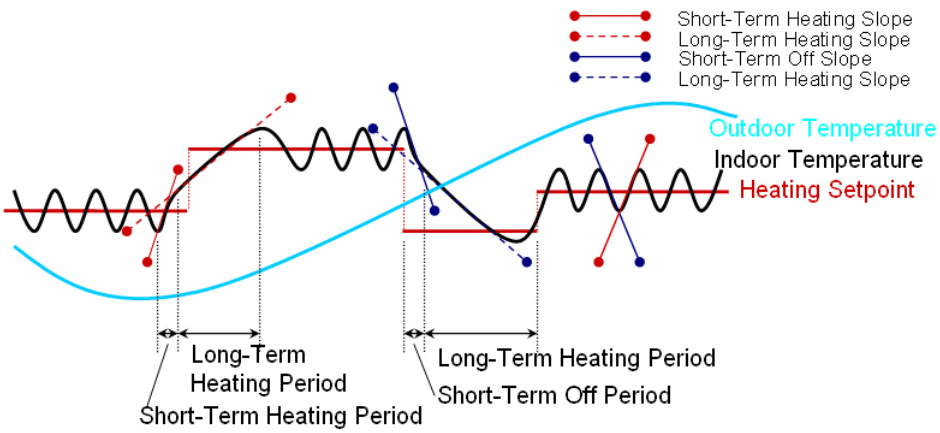


Figure 3.19: Short-Term Slope vs. Long-Term Slope in a Heating Period

these slopes were expected to be similar. However, if a relatively long-term event occurs, the slope tends to change dramatically at a certain point. In Fig.3.19, this dramatic changes occurs in the middle of the transient period between two different heating setpoints. Usually the slope magnitudes during long-term heating and long-term off periods are smaller than those during short-term periods. The distinctive slope changes possibly cause the two lines in Fig.3.16(f).

This phenomenon can be explained from a heat transfer point of view. Heat capacity[5] is the measure of the heat energy required to increase the temperature of an object by a certain temperature interval. Since the inside of a residential building consists of various materials (mostly air, wooden walls or concrete walls, furniture, and carpets) that have different heat capacity, temperature change rate in the individual material is different. In addition, at the beginning of a heating or cooling event, inside air gets heated or cooled first via heat convection and, after some time period, the changed temperature will be balanced with other materials via heat conduction. This explanation, however, comes from the completely theoretical hypothesis. The actual thermal behavior inside the house may not be as simple as the hypothesis.

[5]Heat capacity, C_p is an extensive property because its value is proportional to the amount of material in the object. As distinct from heat capacity, specific heat capacity, also known simply as specific heat, is the measure of the heat energy required to increase the temperature of a unit quantity of a substance by a certain temperature interval.

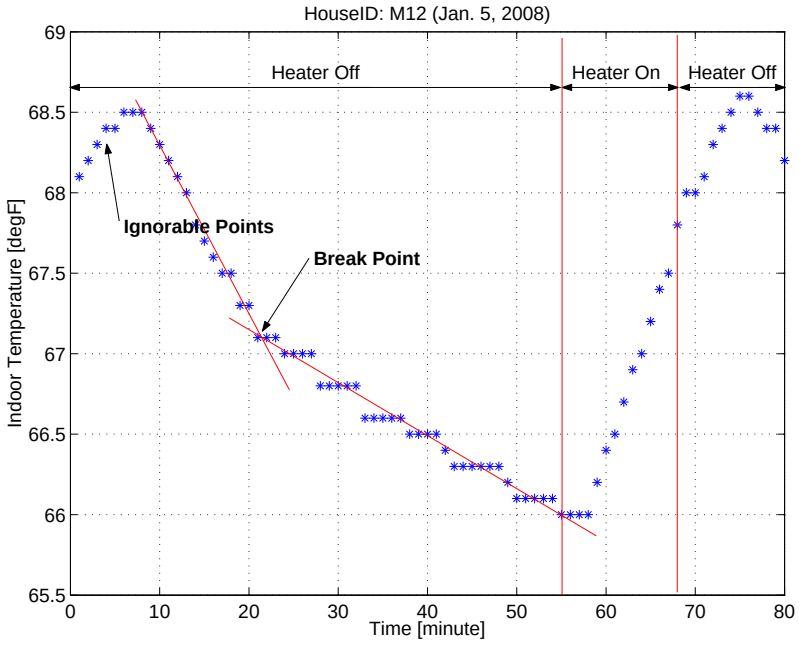


Figure 3.20: Ignorable Period and Break Point (House M12, Jan. 5, 2008)

In the two-slope method, a series of indoor temperature values (off, heating, and cooling) in an individual event is approximated to at most two lines with different slopes, a *short-term slope* and a *long-term slope*. Depending on the length of the event, some events would be approximated to a single line. Compared to the instant slope method, the two-slope method is less sensitive to the measurement noise. While the former method uses only three consecutive points to calculate the temperature change rate that might change every minute, the latter method uses many more points to get two lines per event.

Fig.3.20 shows a good example of how to fit indoor temperature to two lines with different slopes. During the period shown in Fig.3.20, there is one heater-off event (0-55 min) and one heater-on event (55-68 min). The partial heater-off event (68-80 min) is not considered here. Two events were divided based on the actual heater on/off time. The most challenging process in the slope method is to find an exact break point. The break point is defined as a point from which a long-term period starts. In most cases, it is difficult to tell where two lines meet because the indoor temperature usually follows a curve around the break point.

Besides determining the break point, it is important to determine how long the indoor temperature would be affected by the previous event. Even after the heater was turned off in Fig.3.20, the indoor temperature kept increasing for six minutes and began to drop after two more minutes. This is due to residual heat around the heater. The same thing happens during the heater-on period as well. The length of this delay depends on the characteristics of a house and an HVAC unit, temperature difference, and type of events (heater on/off or AC on/off). Therefore, this property also needs to be learned for accurate temperature prediction. Because the temperature values during this period are excluded in calculating the short-term slope, the corresponding points are called *ignorable points*.

Moving Window with a Fixed Sample Size

A simple way to find the break point is to compare the coefficients of determination[6] of linear regression with fixed number of samples. If all samples in the window are located closely to the regression line, R^2would be close to 1. Otherwise, it would be much smaller than 1. Until the window reaches the break point, calculated R^2 values would be close each other as well as close to 1. However, R^2 decreases due to the lack of linearity when the sample window includes the break point. Therefore, the point at which R^2 is minimized would be the break point and two slopes corresponding to maximum R^2 before and after the break point would be the short-term and long-term slopes. These two slopes are stored to the corresponding stable.

The sample size is crucial to determine the sensitivity to measurement noise in this method. In the case that the sample size is too big, the R^2 values around the break point are not clear enough to be detectable. The big sample size is problematic in the house that has relatively short heating and air-conditioning on/off cycles. If an event lasts for less than the time corresponding to the sample size, the algorithm automatically excludes

[6]In statistics, the coefficient of determination, R^2, is the proportion of variability in a data set that is accounted for by a statistical model. The better the regression is, the closer the value is to 1. In the case of linear regression, R^2 is simply the square of a correlation coefficient.

the event to increase the regression accuracy. Therefore, the tables for this type of houses require longer time to be filled with valid data as long as a long event such as precooling and preheating happens. A small sample size, in contrast, increases sensitivity to the noise. If the temperature did not change for three minutes due to the failure of sensing and dramatically dropped afterwards, the algorithm using a small sample size may conclude that the break point locates around these three constant temperatures.

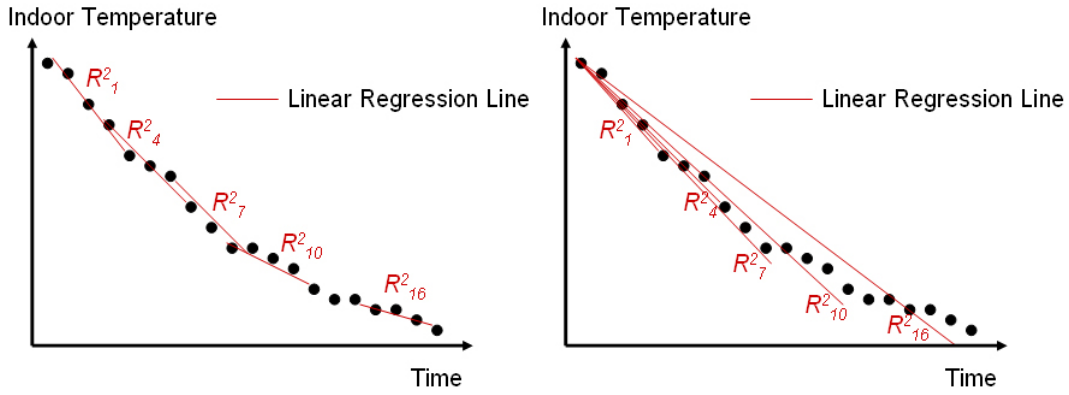
When a default sample size (n=5) was applied to different house data, the method frequently failed to pick a correct break point, which degraded the quality of the slope data in tables.

Fixed Window with an Incremental Sample Size

A flexible sample size that is adjusted by the characteristics of events can overcome the limitation of the fixed sample size. Instead of using a moving window to detect a slope change, a new method fixes the initial sample point but increases the window size (or sample size). The differences in the two methods are compared in Fig.3.21. The initial sample size is 5. The algorithm to find the break point is exactly the same as the previous method. Fig.3.22 explains how the coefficient of determination (or correlation coefficient) changes as the window size increases in both heater-off and heater-on periods. The resultant slope lines are also drawn in the plots. The difference of short-term and long-term slopes is usually bigger in heater-off events than in heater heater-on events.

As shown in Fig.3.22, the lengths of different short-term periods are also different. Even though different lengths do not matter to fill up the table with slope values, the short-term length should be fixed to determine which tables (short-term and long-term) should be referred for prediction. In this method, the short-term length is determined by averaging the values that are distributed within a 99% confidence interval[7]. After the break point is determined, the short-term slope and, if available, the long-term slope are recalculated

[7]When it is assumed that the lengths follow the Gaussian distribution, a 99% confidence interval corresponds to $\bar{X} \pm 2.58\frac{s}{\sqrt{n}}$, where $\bar{X}$: sample mean, s: sample standard deviation, and n: sample size



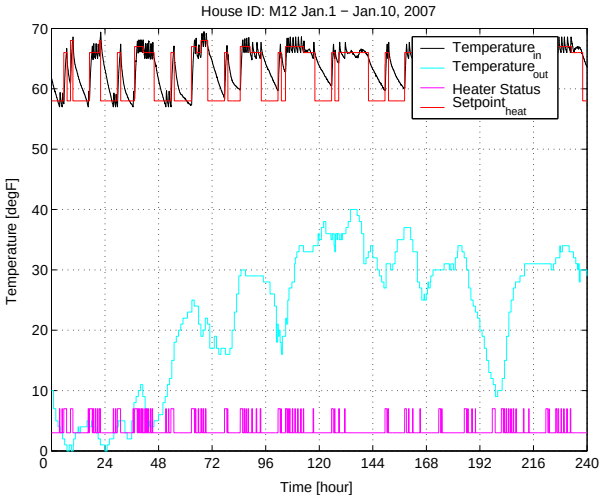
(a) Moving Window w/ a Fixed Sample Size (b) Fixed Window w/ an Incremental Sample Size

Figure 3.21: Moving Window vs. Fixed Window in Linear Regression

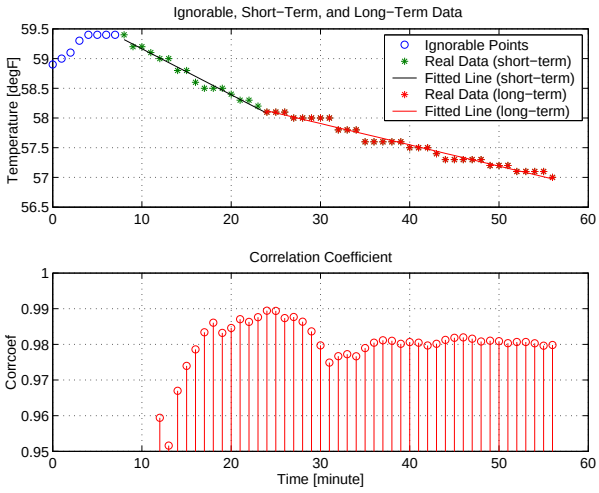
based on the points before/after the break point. Two flow charts in Fig.3.23 summarize how the method works.

Fig.3.24 and Fig.3.25 show the relationship between temperature difference and slope (indoor temperature change rate). The results were obtained from 10 days data (Jan. 1 - Jan. 10, 2008). Except the short-term heater-on period in both houses, the indoor temperature change rate tends to decrease as outdoor to indoor temperature difference decreases in heating mode. . This result is consistent with what we expected in that indoor temperature in that it drops quickly in off mode and increases slowly in heating mode when the outdoor temperature is much lower than the indoor temperature. In addition, the magnitude of the short-term slope is clearly smaller than that of the long-term slope at the same outdoor to indoor temperature difference. For the house M4 (Fig.3.24), two slopes at -40F difference are approximately -10 and -1.7F/hr in off mode and 9 and 3.5 F/hr in heating mode. For the house M12 (Fig.3.25), these values are -6 and -1.2F/hr in off mode, and 10 and 7.7F/hr in heating mode.

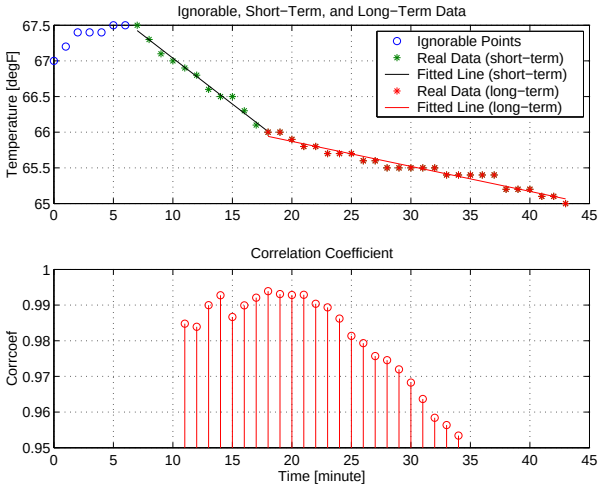
One interesting issue here is that the thermal characteristics of the two houses can be accounted for only by comparing slope values in their tables. The temperature change rates in the long-term off table ranges from -3F/hr at -70F to -0.5F/hr at -20F in house



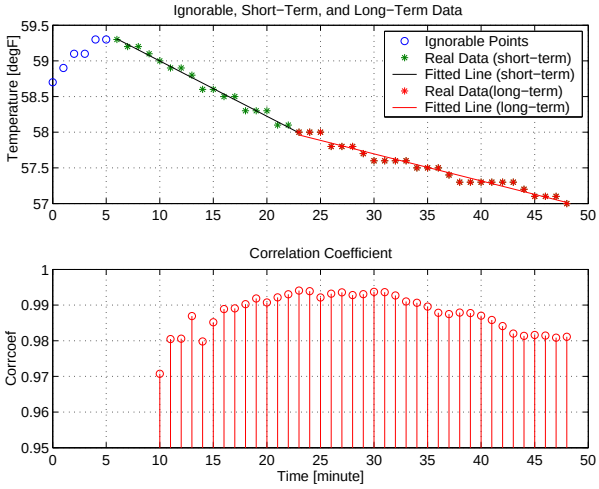
(a) Original Data Profile



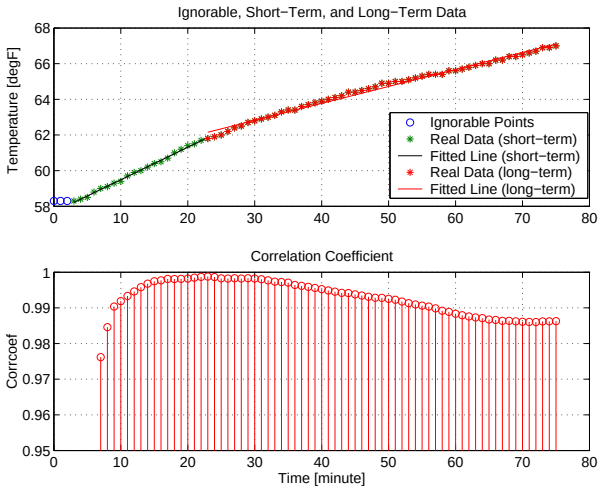
(b) 1st Off Event



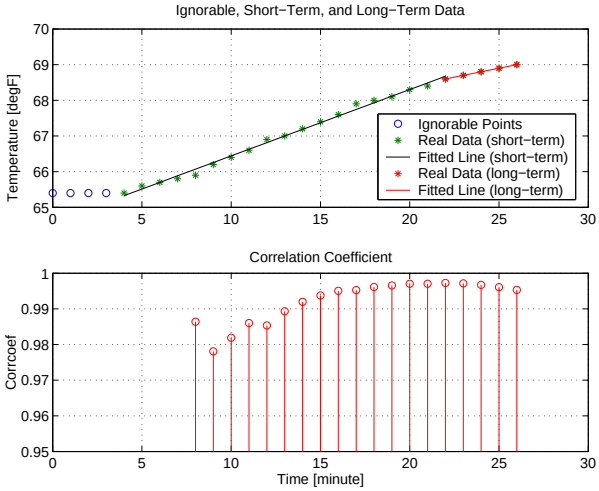
(c) 8th Off Event



(d) 11th Off Event

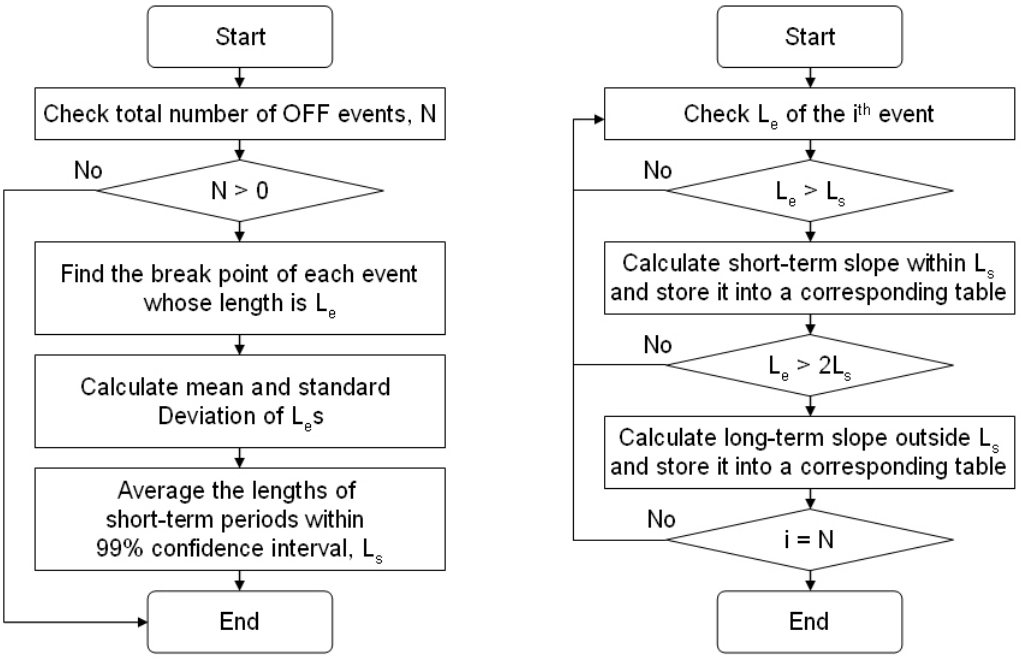


(e) 6th Heater Event



(f) 31st Heater Event

Figure 3.22: Break Point Detecting and Resultant Slope Fitting in Two-Slope Method (House M12, Jan.1 - Jan.10, 2008)



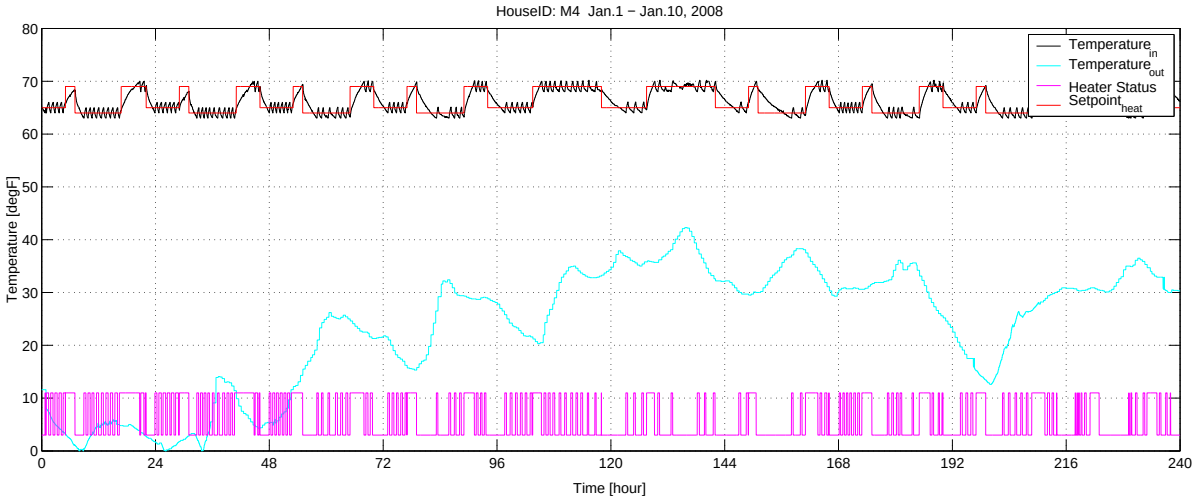
(a) Break Point Determining (b) Table Filling with Slopes

Figure 3.23: Flow Charts for a Break Point Determining Process and a Table Filling Process

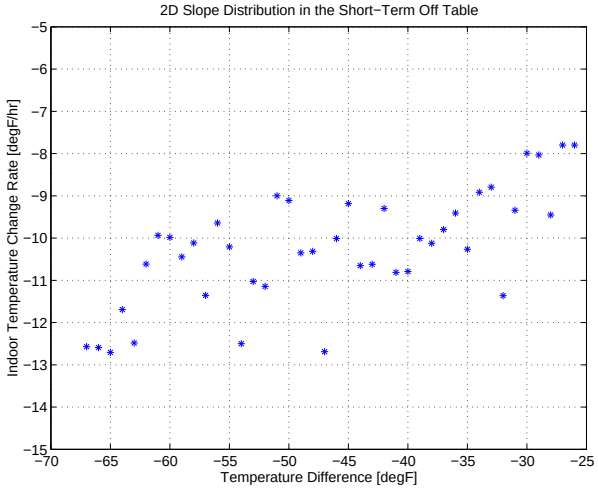
M4. However, house M13 has a smaller range (from -2.1F/hr to -0.6F/hr). The fact that temperature change rate is less affected by the outside condition means that the latter house has better insulation. From Tab.3.2, it can be inferred that the construction year is strongly related to the house insulation. (House M4 was built 20 years earlier than house M12.) In addition, compared with the result from the instant slope method (see Fig.3.18), the proportional relationship is more explicit in the two-slope method.

Although the values in Fig.3.24 and Fig.3.25 are distributed approximately linearly, these values are not adequate for prediction in that these are neither continuous nor perfectly linear. If adjoining three values in the same table are out of line or one of slots is empty, the raw tables do not provide reliable prediction at all. As a method to fill up the empty slots as well as smooth the raw data in the tables, linear regression is performed again. Since the data distribution in the table is not necessarily linear, the linear regression should be considered one option for filtering.

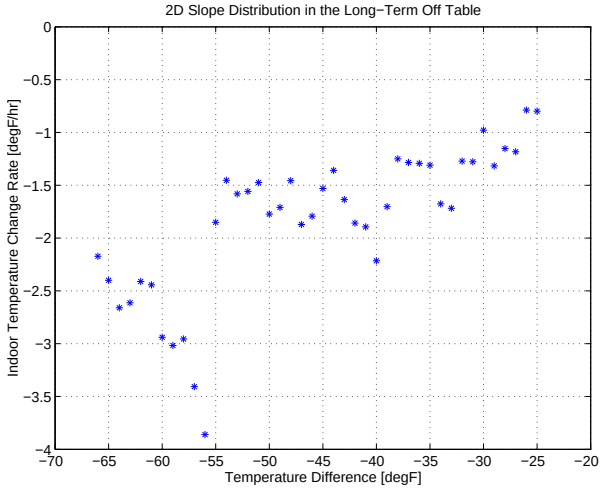
To evaluate the method, predicted indoor temperature and measured indoor temper-



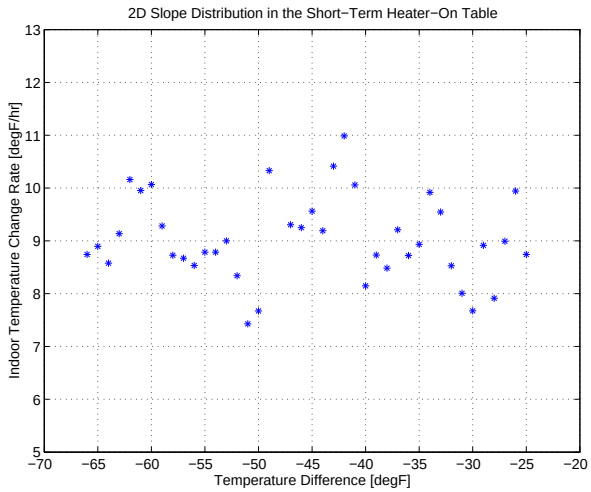
(a) Original Data Profile



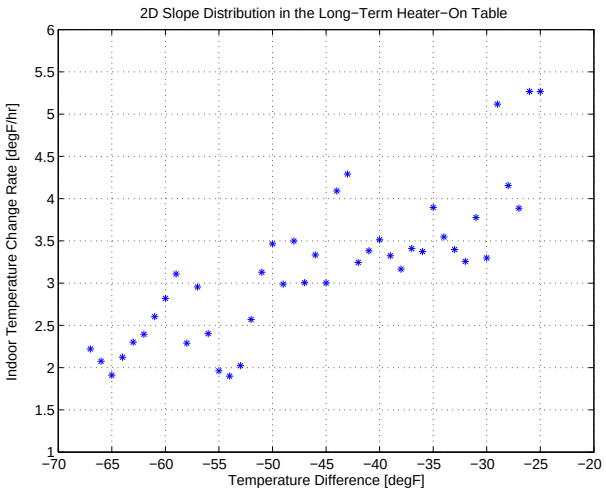
(b) Short-Term Off Table



(c) Long-Term Off Table

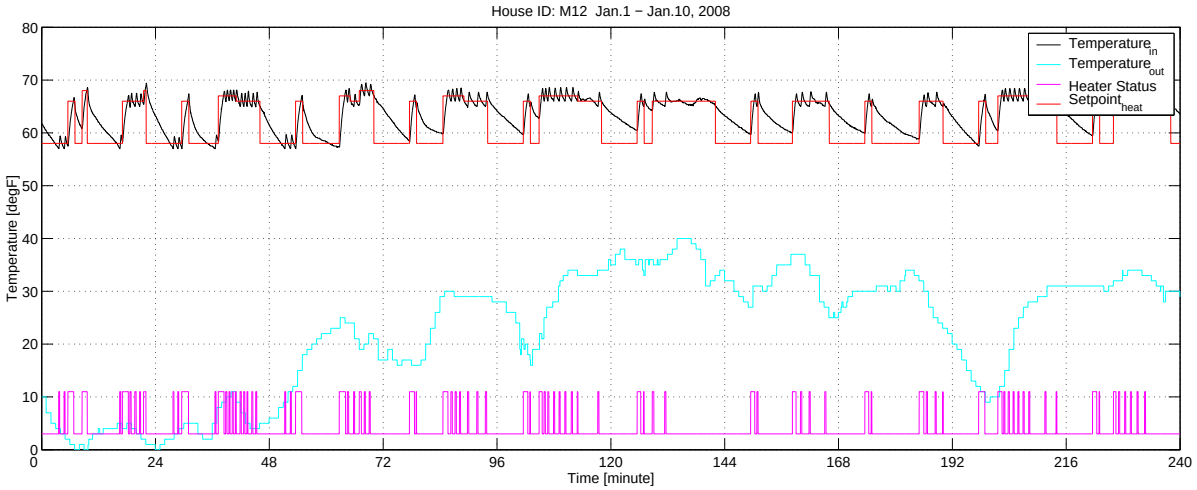


(d) Short-Term Heater-On Table

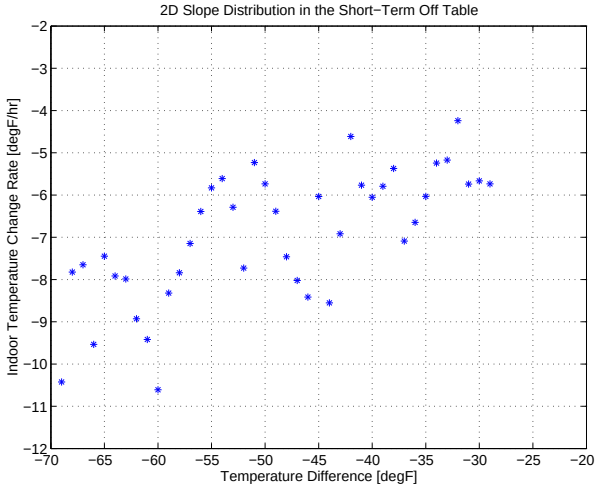


(e) Long-Term Heater-On Table

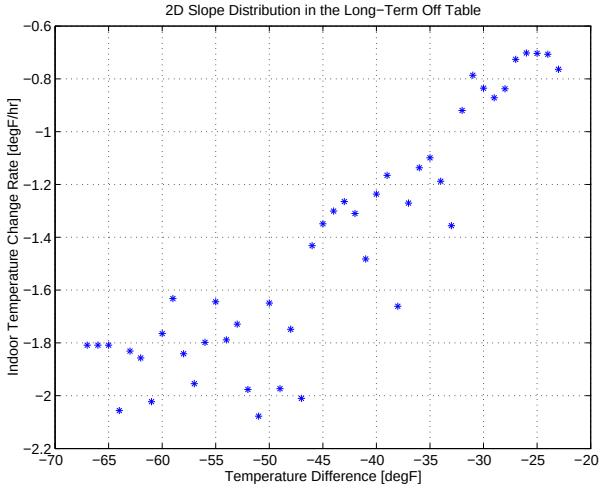
Figure 3.24: Slope Distribution of the House M4 (Jan. 1 - Jan. 10, 2008)



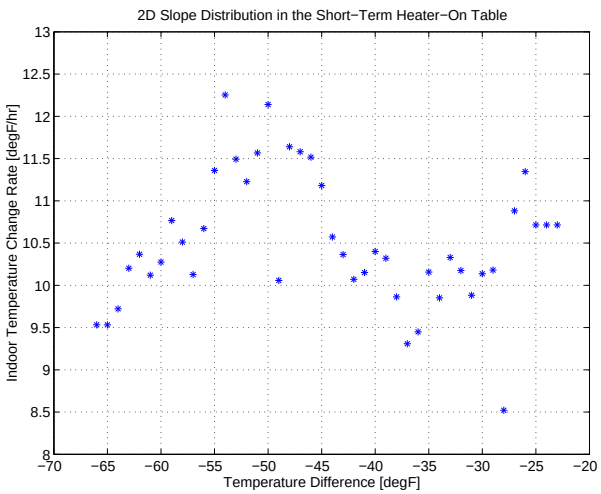
(a) Original Data Profile



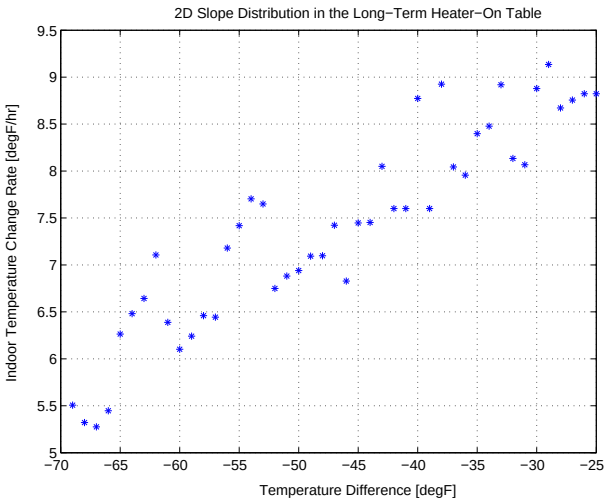
(b) Short-Term Off Table



(c) Long-Term Off Table



(d) Short-Term Heater-On Table



(e) Long-Term Heater-On Table

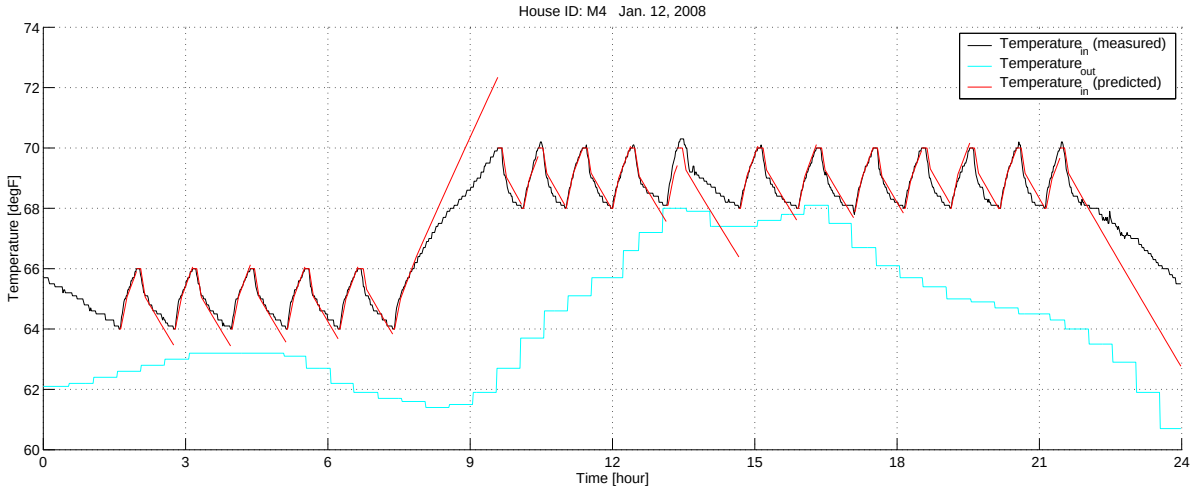
Figure 3.25: Slope Distribution of the House M12 (Jan. 1 - Jan. 10, 2008)

ature for three consecutive days were compared. The prediction was based on the tables learned during past 10 days. Two temperature values were synchronized at the beginning of each event. The results are shown in Fig.3.26. While the predicted value for short-term events follows the actual measurement pretty well, the long-term prediction is not reliable especially for the heater-on event.

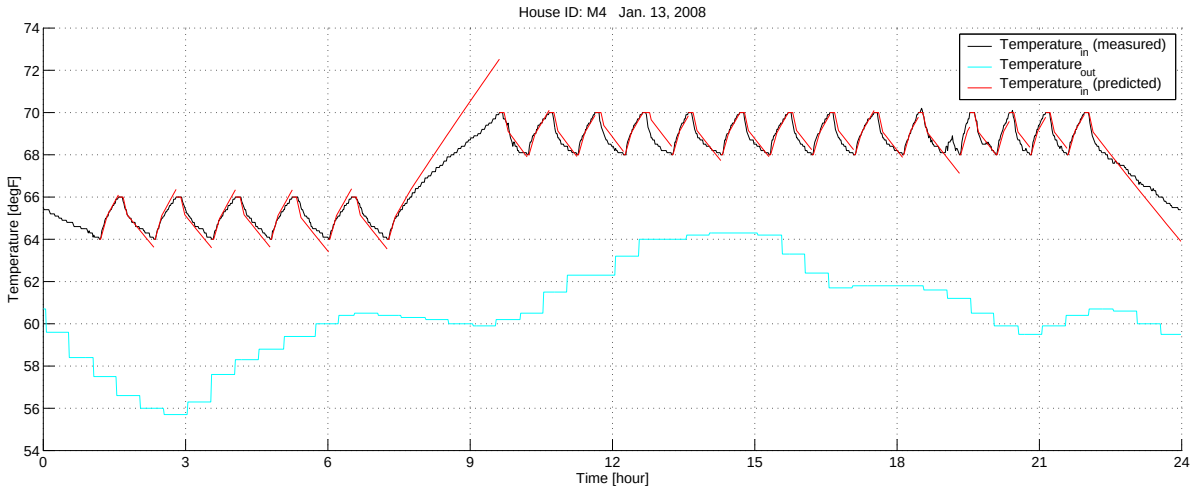
3.4.3 Three-Slope Method

When the time of an event is long enough for the indoor temperature to reach approximately its balance point, two lines may not be sufficient to approximate the whole temperature profile of a single event. Three-slope method is not much different from the two-slope method in that the algorithms in it are the same as those in the two-slope method. The only difference is one more break point must be determined in the long-term period, and two lines are used for the long-term data. The period after the second break point is called an *extra-long-term* period for convenience. In Fig.3.27, the 112-minute-long off event is approximated with both two and three lines. It is certain that the regression error drops down with an additional line.

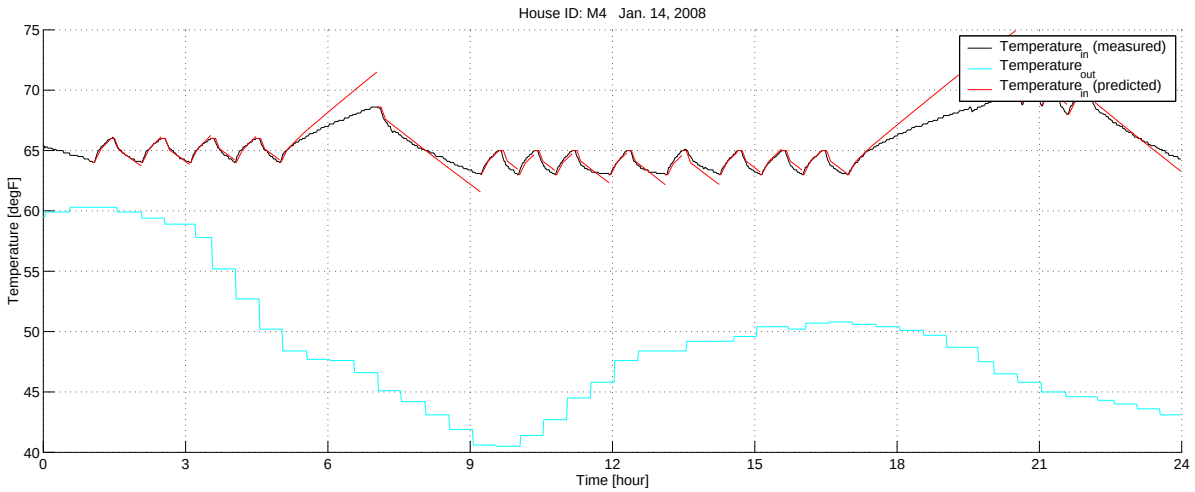
Besides using three lines in data fitting, solar radiation effect on the slope change was also considered in this method. If the temperature change rate is significantly affected by the solar radiation, performance of the tabular method can be enhanced by adding one more dimension to the existing table structure. In analysis on solar radiation effect from the following results, two facts should be considered. First, the solar radiation data for the houses in Minnesota and Australia are not the real measurement but the simulated data. The global solar radiation was predicted based on weather information from the local weather station and the Bras method [26].A turbidity factor n_{fac} in the Bras method could be determined by mapping weather condition to a specific value that ranges from 2 to 5. Using predicted data that do not reflect actual solar radiation around the house possibly leads to inaccurate results. Second, the solar radiation effect on the house temperature is



(a) Prediction on Jan.12, 2008 (Tables Learned from Jan.1 - Jan.10, 2008)

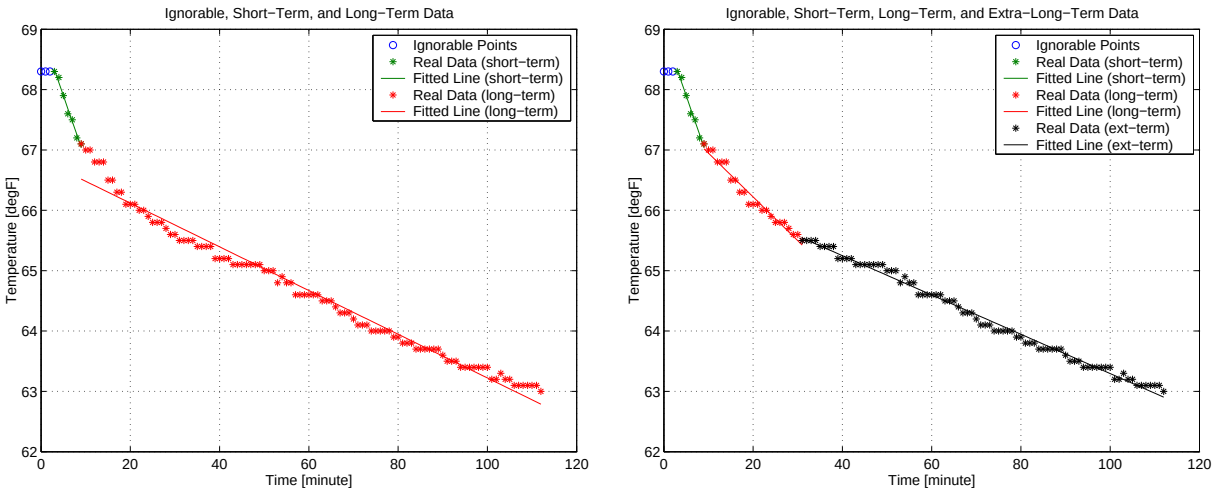


(b) Prediction on Jan.13, 2008 (Tables Learned from Jan.1 - Jan.10, 2008)



(c) Prediction on Jan.14, 2008 (Tables Learned from Jan.1 - Jan.10, 2008)

Figure 3.26: Prediction with Two-Slope Method (House M4, Jan.12 - Jan.14, 2008)



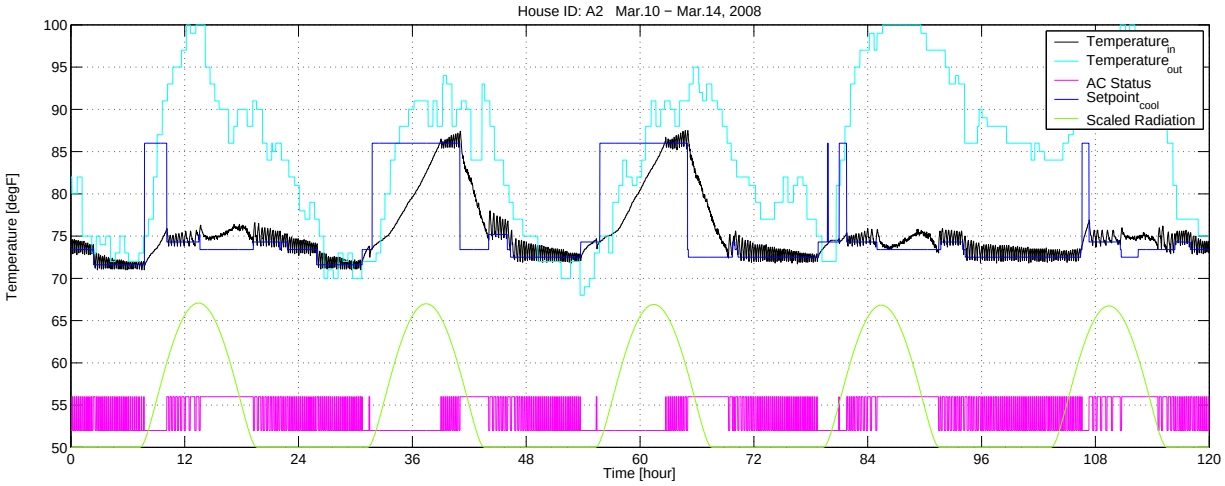
(a) Two Slopes (b) Three Slopes

Figure 3.27: Two-Slope Approximation vs. Three-Slope Approximation (House M4, Jan. 1, 2008)

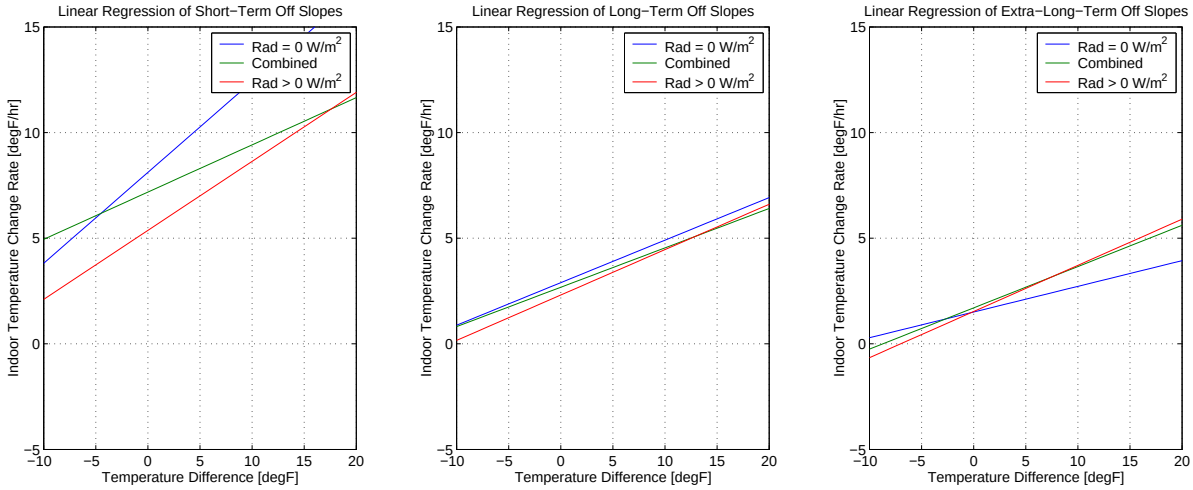
indirect. As demonstrated in Fig.3.3, there exists a delay. House structures such as window type, direction, trees, and surrounding buildings also cause the indirect consequence.

In spite of possible corruption induced by the above uncertainty, the following results still exhibit the meaningful correlation between the strength of solar radiation and temperature change rate. Each plot in Fig.3.28(b) and (c) contains three linear regression lines corresponding to the temperature change rate under no solar radiation (blue) and solar radiation (red) as well as whole data (green). Five consecutive days retaining clear sky (constant n_{fac} value) were used for this analysis. While the solar radiation increases the temperature rate over all ranges in the air-conditioning mode (Fig.3.28(c)), this correlation is not clearly shown in the off mode (Fig.3.28(b)). The temperature change rate is much higher under no solar radiation particularly in the short-term periods of the off mode. Fig.3.29 and 3.30 display the slope distribution and linear regression line of each case in detail. It was discovered that the solar radiation is related to the temperature change rate but how they are related was not evident in this analysis.

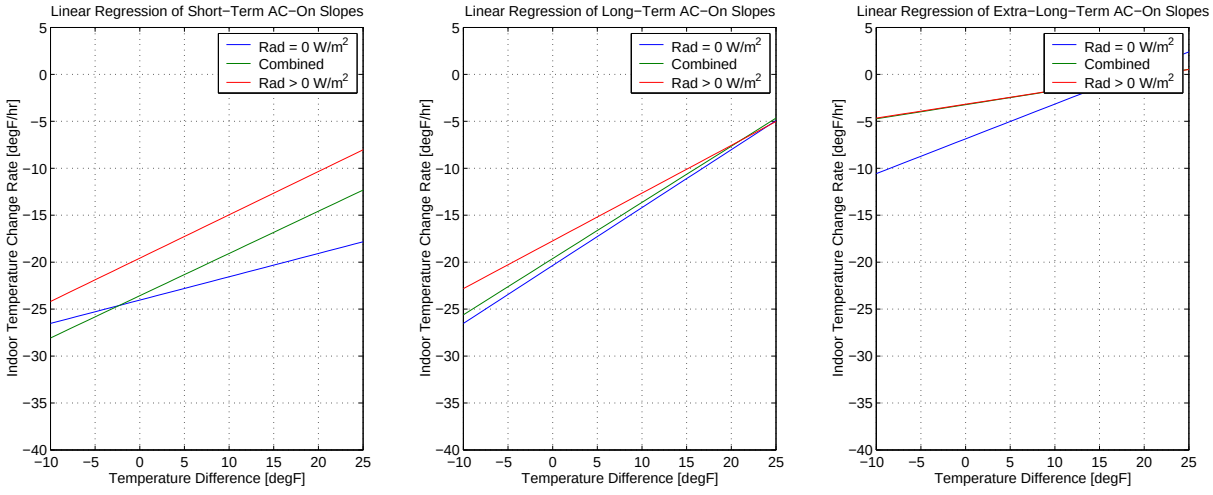
Fig.3.31 includes the indoor temperature prediction based on the three-slope tables learned for the past 10 days. When it is compared with Fig.3.26, it turns out that the



(a) Original Data Profile

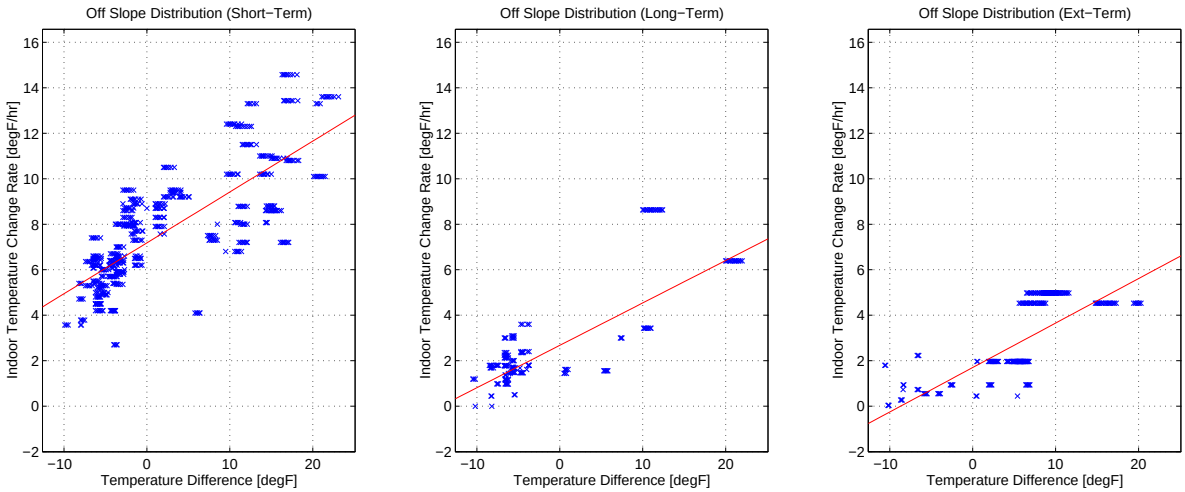


(b) 3 Linear Regression Lines for the Off Table

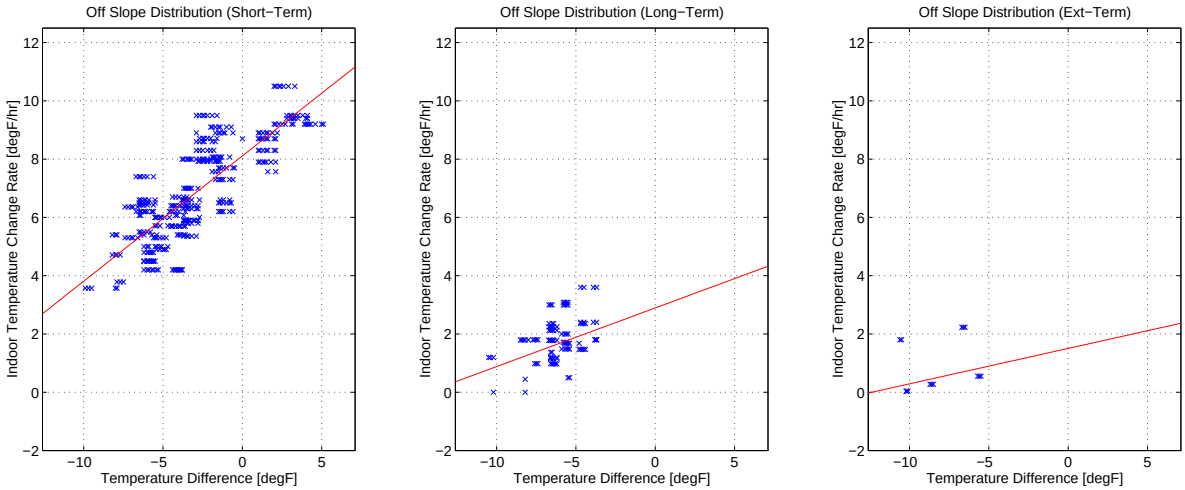


(c) 3 Linear Regression Lines for the AC-On Table

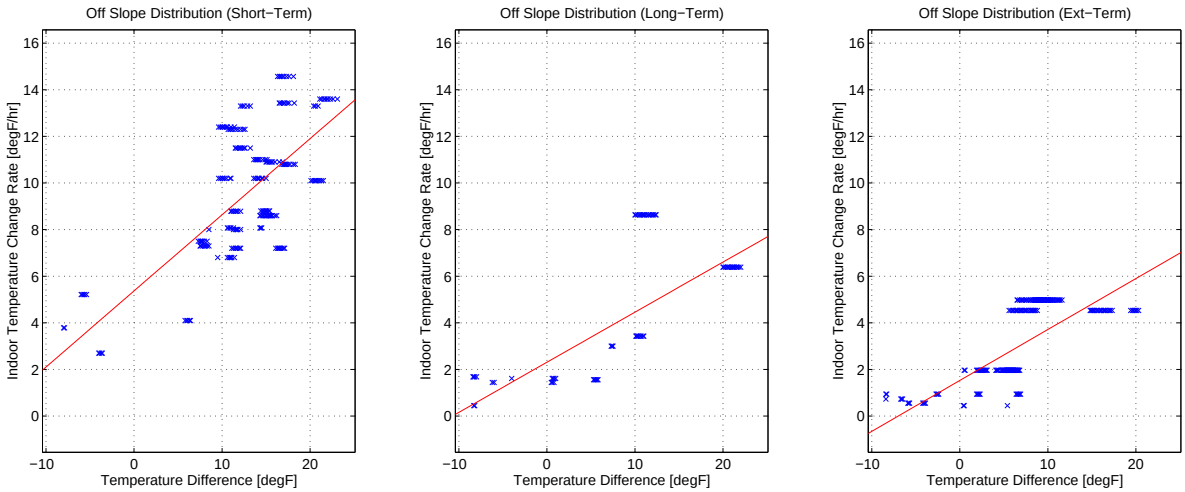
Figure 3.28: Solar Radiation Effect on the Slope in the House A2 (Mar. 10 - Mar. 14, 2008)



(a) Off Slope Distribution and Regression Line for All Data

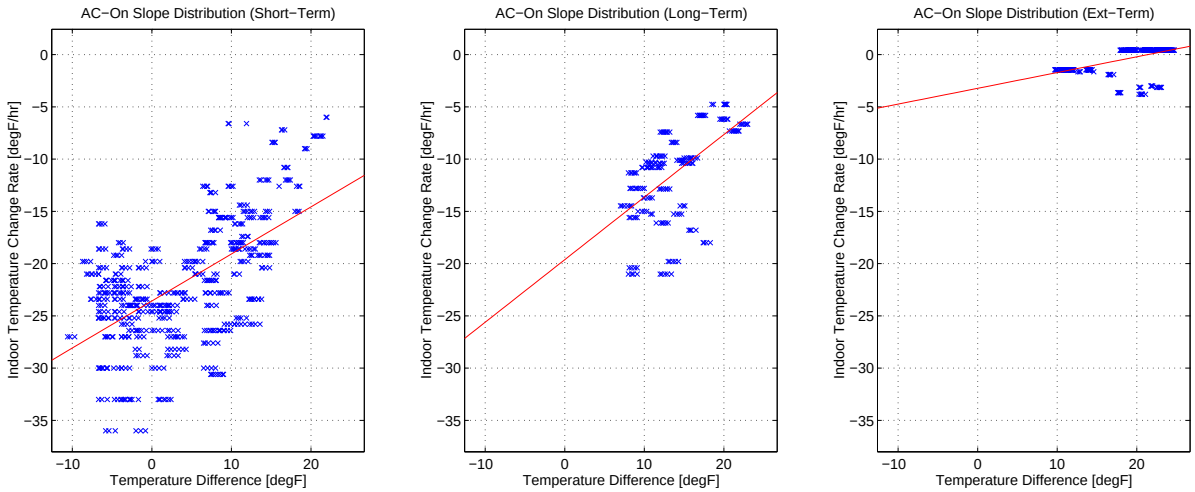


(b) Off Slope Distribution and Regression Line for Data below 250W/m2

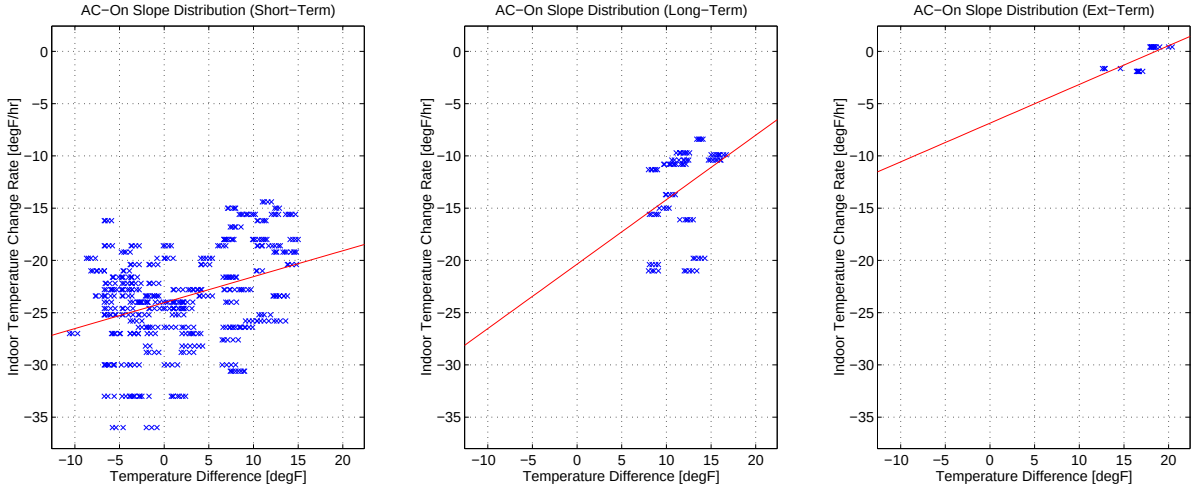


(c) Off Slope Distribution and Regression Line for Data above 250W/m2

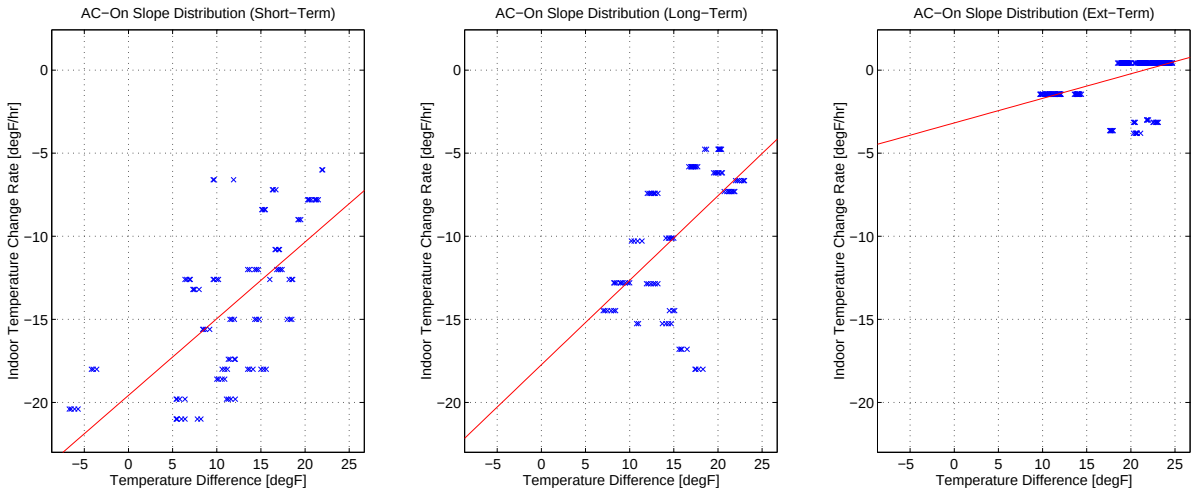
Figure 3.29: Solar Radiation Effect on the Off Slope in the House A2 (detailed)



(a) AC-On Slope Distribution and Regression Line for All Data



(b) AC-On Slope Distribution and Regression Line for Data below 250W/m2



(c) AC-On Slope Distribution and Regression Line for Data above 250W/m2

Figure 3.30: Solar Radiation Effect on the AC -On Slope in the House A2 (detailed)

three-slope method characterizes the temperature behavior more accurately especially for long events. Prediction for the last off event on Jan. 13 (10pm - 12am) and the last long heater-on event on Jan. 14 (from 5pm to 8pm) shows significant improvement. However, the slope method does not always provide reliable results. When it was applied to summer data from the houses in Australia, the prediction error was much bigger than that of the winter data from the houses in Minnesota. While occupants usually keep the windows closed during a winter season, they frequently open the windows during a summer season. This behavior changes the thermal characteristics of the whole house and allows tables to be mixed with factitious data[8].

3.5 ARX (Auto-Regressive with Exogenous Input) Model

Parametric models describe systems in terms of differential equations and transfer functions. A system can generally be described using 3.6, which is known as the general-linear polynomial model or the general-linear model. The general-linear model structure, shown in Fig.3.32, provides flexibility for both the system dynamics and stochastic dynamics.

$$y(n) = q^{-k}G(q^{-1}, \theta)u(n) + H(q^{-1}, \theta)e(n) \tag{3.6}$$

where $u(n)$ and $y(n)$: input and output of the system respectively

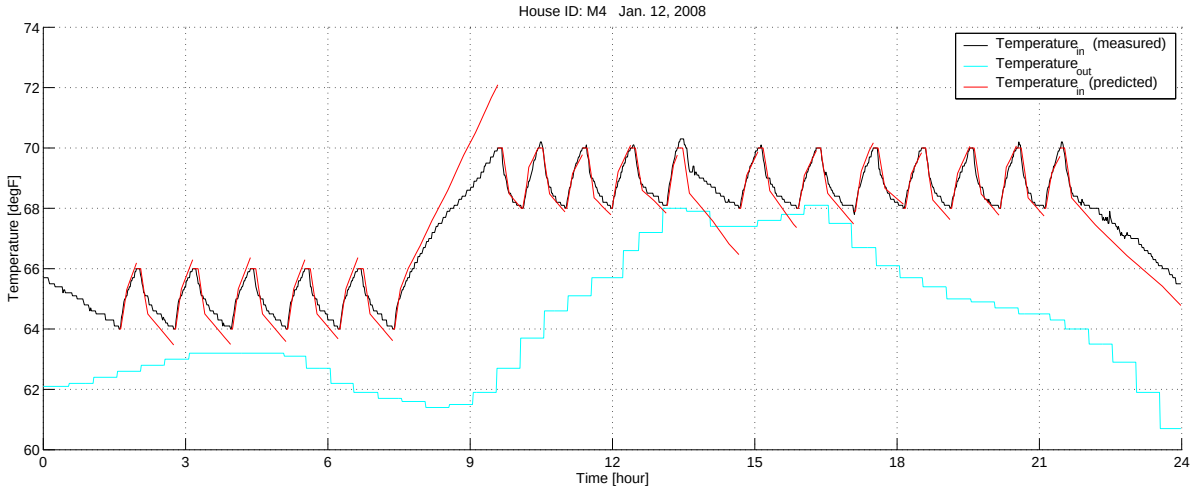
$e(n)$: zero-mean white noise, or disturbance of the system

$G(q^{-1}, \theta)$: transfer function of the deterministic part of the system

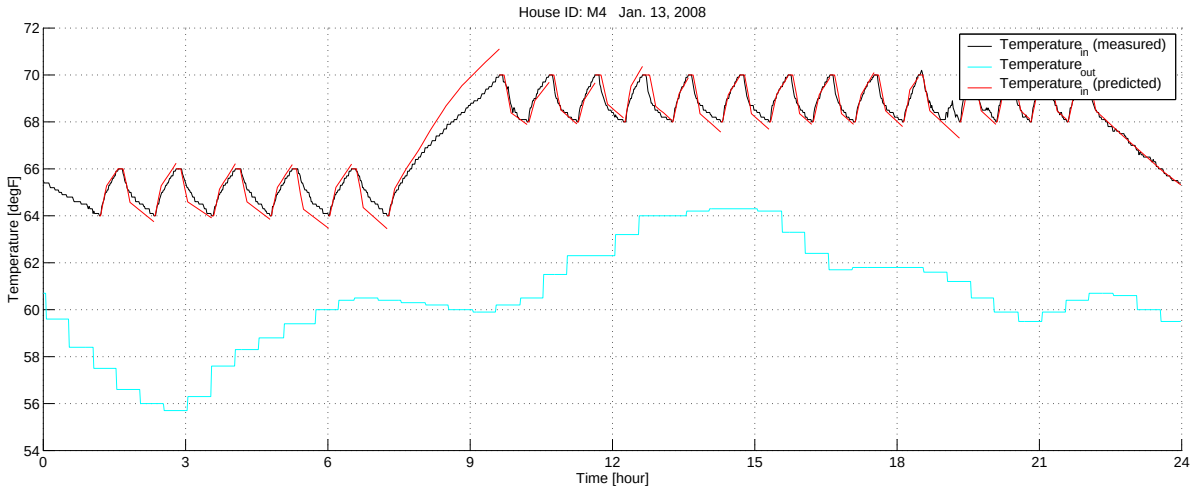
$H(q^{-1}, \theta)$: transfer function of the stochastic part of the system

Among parametric models, the ARX model is the most efficient of the polynomial estimation methods because it is the result of solving linear regression equations in analytic form and its solution is unique. However, the coupling between the deterministic and stochastic dynamics can bias the estimation of the ARX model when the disturbance e of

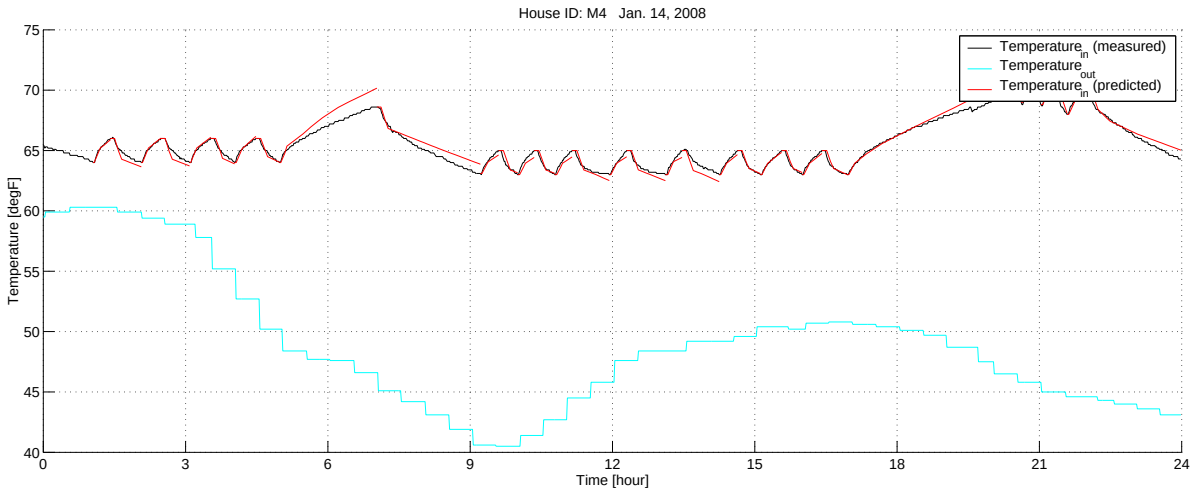
[8]This data can be useful to characterize the house thermal properties under a specific condition



(a) Prediction on Jan.12, 2008 (Tables Learned from Jan.1 - Jan.10, 2008)

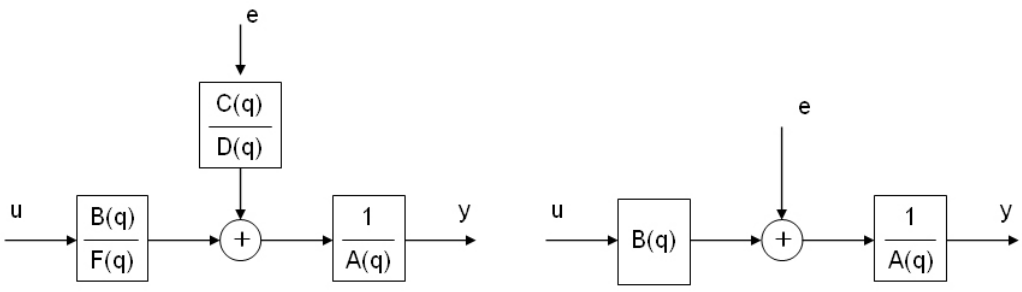


(b) Prediction on Jan.13, 2008 (Tables Learned from Jan.1 - Jan.10, 2008)



(c) Prediction on Jan.14, 2008 (Tables Learned from Jan.1 - Jan.10, 2008)

Figure 3.31: Prediction with Three-Slope Method (House M4, Jan.12 - Jan.14, 2008)



(a) Genera-Linear Model (b) ARX Model

Figure 3.32: General-Linear and ARX Model Structure

the system is not white noise. In order to minimize the equation error, the model can have a higher order than the actual system under low signal-to-noise ratio.

A model with more freedom or parameters does not necessarily provide better estimation as it can result in the modeling of nonexistent dynamics and noise characteristics. Therefore, a certain level of physical insight into a system is required to build a proper model. The following section briefly describes a modeling procedure for an HVAC unit and building structures, which was initially proposed by William Burke and David Auslander in 2007[27]. Instead of finding physical parameters that are associated with thermal properties of the actual system, coefficients of the polynomials, which do not have any physical meaning, were calculated using a prediction error/Maximum Likelihood method.

3.5.1 Modeling

Despite the diversity in a building structure, the house equipped with a single zone HVAC system was considered for modeling. This model includes the four temperatures of the indoor air (T_{air}), internal walls (T_{iw}), external walls (T_{xw}), and cooler (or heater) mass (T_{mass}) as states. In modeling, several handbooks regarding HVAC and heat-mass transfer were referenced [28, 29, 30, 31, 32, 33].

HVAC Unit

When the length between the inlet and outlet of the cooler is L and the temperatures are T_{mass} and T_{sup} respectively, it can be assumed that the temperature in the middle ($T(x)$) is distributed exponentially as in Eq.3.7.

$$T(x) = T_{air} + (T_{mass} - T_{air})(1 - e^{-\frac{x}{k_3 L}}) \tag{3.7}$$
$$T_{sup} = T(L) = T_{air} + (T_{mass} - T_{air})(1 - e^{-\frac{1}{k_3}})$$

Conductive heat transfer from the cooler mass to outside can be considered heat loss associated with the HVAC unit and it is expressed by Eq. 3.8. Eq.3.9 explains the heat transfer to the inside air between the outlet and the inlet of the cooler. Then, the temperature variation of the cooler mass follows Eq.3.10 when Q_{adj} is the adjusted heat input. Q_{adj} is determined by the thermal efficiency of the cooler, which is mainly dependent on the outside temperature. Finally, the supply air and the indoor air exchange the heat through the convective heat transfer as in Eq.3.11.

$$Q_{loss} = Q_{mass2out} = k_2(T_{mass} - T_{out}) \tag{3.8}$$

$$\begin{aligned} Q_{mass2air} &= k_1 \frac{1}{L} \int_0^L (T(x) - T_{air}) dx \\ &= k_1 (T_{mass} - T_{air})(1 + k_3[e^{-\frac{1}{k_3}} - 1]) \end{aligned} \tag{3.9}$$

$$\begin{aligned} \dot{T}_{mass} &= \frac{Q_{adj} - Q_{loss} - Q_{mass2air}}{c_{mass} m_{mass}} \\ &= \frac{Q_{adj} - k_2(T_{mass} - T_{out}) - k_1(T_{mass} - T_{air})(1 + k_3[e^{-\frac{1}{k_3}} - 1])}{c_{mass} m_{mass}} \end{aligned} \tag{3.10}$$

$$q_{sup2air} = \frac{\nu_{mass}}{m_{air}} (T_{sup} - T_{air}) \tag{3.11}$$

External Walls and Window Glass Panes

Assuming that walls are composed of uniform external material, the heat transfer between the outdoor air and the indoor air by conduction follows Eq.3.12 and 3.13. If the heat by the adsorbed solar radiation is also considered, the differential equation for the external walls is given by Eq.3.14.

$$Q_{xw2air} = k_4(T_{xw} - T_{air}) \tag{3.12}$$

$$Q_{xw2out} = k_4(T_{xw} - T_{out}) \tag{3.13}$$

$$\begin{aligned} \dot{T}_{xw} &= \frac{-Q_{xw2air} - Q_{xw2out} + Q_{rad2xw}}{c_{xw}m_{xw}} \\ &= \frac{-k_4(T_{xw} - T_{air}) - k_4(T_{xw} - T_{out}) + c_{rad2xw}I_{global}}{c_{xw}m_{xw}} \\ &= \frac{k_4(T_{air} + T_{out} - 2T_{xw}) + c_{rad2xw}I_{global}}{c_{xw}m_{xw}} \end{aligned} \tag{3.14}$$

Besides the heat transfer at the surface of the external walls, an amount of heat is transferred to the indoor air through the windows. The absorbed solar radiation and infiltrated outdoor air affect the indoor air temperature (Eq.3.15 and 3.16)

$$Q_{win2air} = \frac{1}{3600}A_{win}C_{win}(T_{out} - T_{air}) \tag{3.15}$$

$$q_{inf} = \frac{\nu_{inf}}{m_{air}}(T_{out} - T_{air}) \tag{3.16}$$

Internal Walls

Internal walls include any mass inside the house such as furniture, floors, and walls. The conductive heat transfer between internal walls and indoor air is given by Eq.3.17 and the heat due to the transmitted solar radiation is combined to calculate temperature change in the internal walls (Eq.3.18).

$$Q_{iw2air} = k_5(T_{iw} - T_{air}) \tag{3.17}$$

$$\begin{aligned}\dot{T}_{iw} &= \frac{-Q_{iw2air} + Q_{rad2iw}}{c_{iw}m_{iw}} \\ &= \frac{k_5(T_{air} - T_{iw}) + c_{rad2iw}I_{global}}{c_{iw}m_{iw}}\end{aligned} \tag{3.18}$$

Indoor Air

The dynamics of the indoor air temperature is determined by the heat transfer associated with the all elements mentioned above as well as internal heat sources such as human bodies and electrical appliances. The indoor air temperature is calculated by Eq.3.19.

$$\begin{aligned}\dot{T}_{air} &= q_{sup2air} + q_{inf} + \frac{Q_{int} + Q_{iw2air} + Q_{xw2air} + Q_{win2air}}{c_{air}m_{air}} \\ &= \frac{v_{hvac}}{m_{air}}(T_{mass} - T_{air})(1 - e^{-\frac{1}{k_3}}) + \frac{\nu_{inf}}{m_{air}}(T_{out} - T_{air}) \\ &+ \frac{Q_{int}}{c_{air}m_{air}} + \frac{k_4}{c_{air}m_{air}}(T_{iw} - T_{air}) + \frac{k_5}{c_{air}m_{air}}(T_{xw} - T_{air}) \\ &+ \frac{A_{win}c_{win}}{3600c_{air}m_{air}}(T_{out} - T_{air})\end{aligned} \tag{3.19}$$

By combining Eq. 3.10, 3.14, 3.18, and 3.19 and substituting T_{mass}, T_{xw}, T_{iw}, and $T_{air}$with x_1, x_2, x_3, and x_4 respectively, each state is given by Eq.3.20, 3.21, 3.22, and 3.23.

$$\dot{x}_1 = \lambda_{11}Q_{adj} + \lambda_{12}(x_1 - T_{out}) + \lambda_{13}(x_1 - x_4) \tag{3.20}$$

$$\dot{x}_2 = \lambda_{21}(x_4 + T_{out} - 2x_2) + \lambda_{22}I_{global} \tag{3.21}$$

$$\dot{x}_3 = \lambda_{31}(x_4 - x_3) + \lambda_{32}I_{global} \tag{3.22}$$

$$\begin{aligned}\dot{x}_4 &= \lambda_{41}(x_1 - x_4) + \lambda_{42}(T_{out} - x_4) + \lambda_{43}Q_{int} + \lambda_{44}(x_3 - x_4) \\ &+ \lambda_{45}(x_2 - x_4) + \lambda_{46}(T_{out} - x_4)\end{aligned} \tag{3.23}$$

The above equations can be written in the form of the general-linear model.

$$A(q)x_4(t) = B_1(q)T_{out}(t) + B_2(q)Q_{adj}(t) + B_3(q)I_{global}(t) + B_4(q)Q_{int}(t) + e(t) \quad (3.24)$$

where $A(q) = 1 + a_1q^{-1} + a_2q^{-2} + a_3q^{-3} + a_4q^{-4}$

$B_1(q) = b_{11}q^{-1} + b_{12}q^{-2} + b_{13}q^{-3} + b_{14}q^{-4}$

$B_2(q) = b_{22}q^{-2} + b_{23}q^{-3} + b_{24}q^{-4}$

$B_3(q) = b_{32}q^{-2} + b_{33}q^{-3} + b_{34}q^{-4}$

$B_4(q) = b_{41}q^{-1}$

When the efficiency of the cooler is assumed to be constant, and heat from the internal heat sources does not fluctuate considerably over time, a unit function $1_{hvac}(t)$ (1 for on and 0 for off) can be substituted for $Q_{adj}(t)$ and $Q_{int}(t)$ can be set 1. Then the system identification process boils down to optimization of 15 parameters in the 4th order ARX model with 4 inputs (outdoor temperature, cooler on/off status, global solar radiation, and internal heat gain) and one output (indoor temperature). From the collected input and output data, the parameters can be calculated by the least square. The solution is given by Eq.3.25.

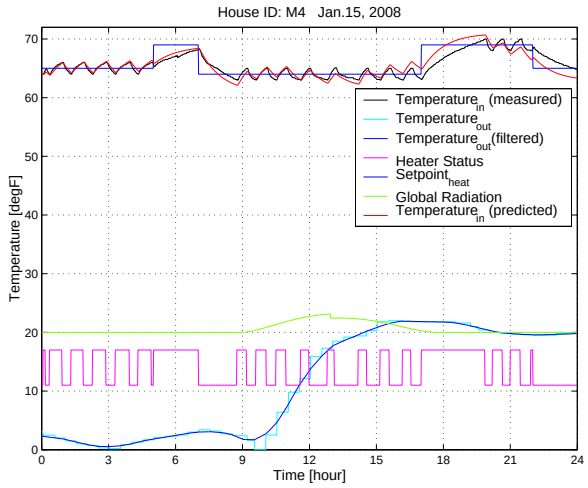
$$[a_1 \cdots a_4 \mid b_{11} \cdots b_{14} \mid b_{22} \cdots b_{24} \mid b_{32} \cdots b_{34} \mid b_{41}]^\top = \left(X^\top X\right)^{-1} X^\top Y \quad (3.25)$$

where $X = [T_{air}(k+3:k) \mid T_{out}(k+3:k) \mid 1_{hvac}(k+2:k+1) \mid I_{global}(k+2:k+1) \mid 1]$

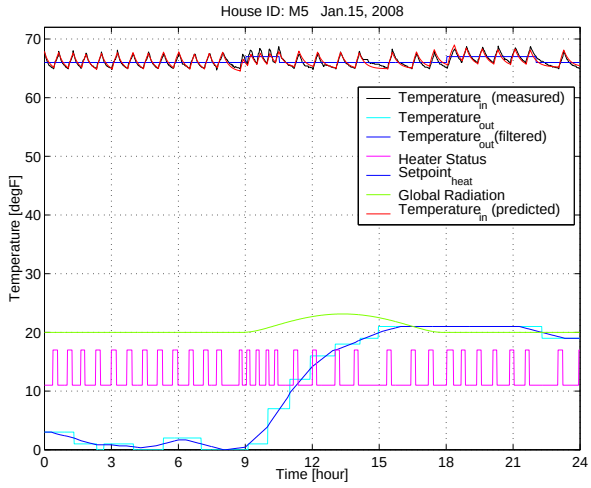
$Y = [T_{air}(5)\, T_{air}(6) \cdots T_{air}(n)]^\top$, n is a data size and $k = 0 \cdots n-1$

3.5.2 Performance Evaluation

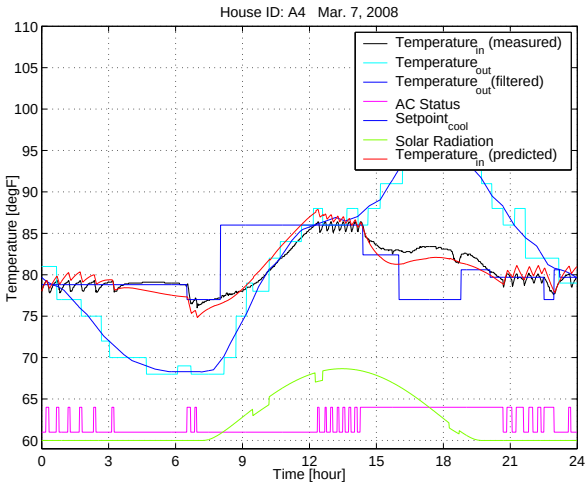
The model validation is to check how accurately the output of the ARX model follows the behavior of the actual indoor temperature. The validation data normally differ from the data that are used to determine parameters. However, by using the exact same data both in parameter learning and for temperature prediction, the 4th order is enough to realize



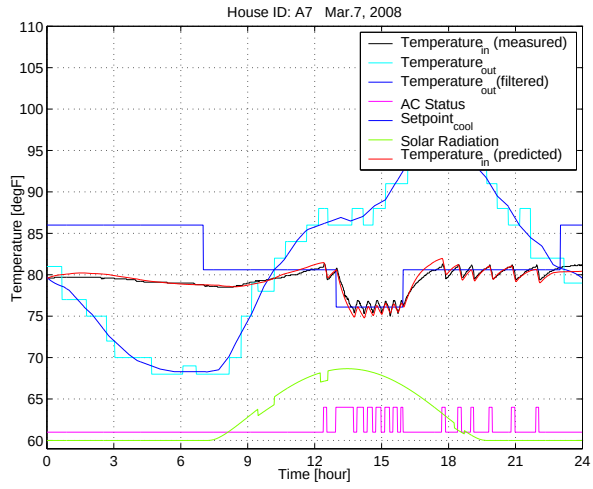
(a) House M4 (Jan.15, 2008)



(b) House M5 (Jan.15, 2008)



(c) House A4 (Mar.7, 2008)



(d) House A7 (Mar.7, 2008)

Figure 3.33: Prediction By the ARX Model Learned from the Same Day Data

a complex system. In Fig.3.33, the simulated temperature from the ARX model whose parameters were calculated from the data on Jan. 15, 2008 for the Minnesota houses and data on Mar. 7, 2008 for the Australia houses are compared with the indoor temperature of the same day.

The simulated temperature was not synchronized with the real temperature during 24 hours, with the exception of the initial start point while the model used the same HVAC on/off time as the real data. Within $2°F$ error, the ARX model could reproduce the real indoor temperature of house M4 and A4. In the house M5 and house A7 cases, the

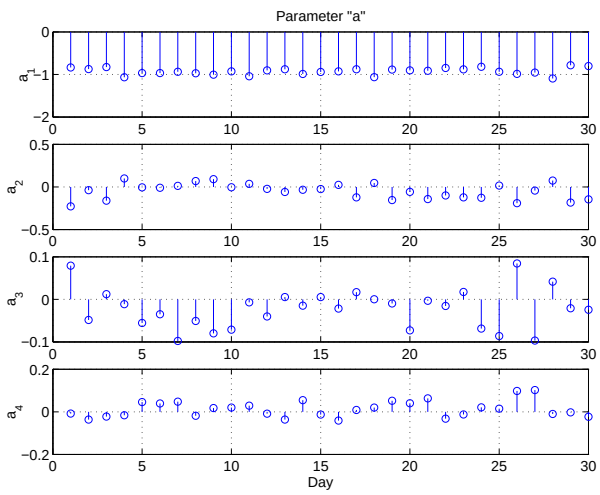
predictions were more accurate. Overall, the 4th order model provided good prediction results for both summer and winter periods when it used the same data for learning and prediction.

Although the parameters of the ARX model do not have any physical meaning by themselves, they should be constant over time in that they are combinations of thermal parameters, which are fixed for a given house. If the parameters calculated from different data sets are close to each other, the model would represent the house thermal properties very well. In order to probe consistency of 15 parameters, parameters calculated from a 24-hour-long data period were compared over 30 days.

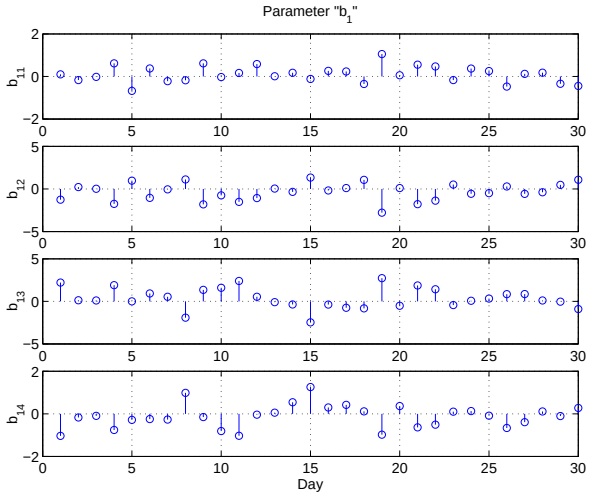
Fig.3.34 shows the parameter change over time in house A4. Unfortunately, parameters were not consistent, except for a_1. This result implies that the parameters learned from past data may not be useful for prediction. Among 30 days, 5 days (Mar.5, 6, 7, 9, and 10) on which the fitted curve follows the actual indoor temperature with relatively small error were considered again. During this period, a_1, a_2, a_3, a_4, b_{13}, b_{14}, b_{22}, b_{23}, b_{24}, b_{34}, and b_{41} out of 15 parameters are close to each other. It is interesting that $b_1$and b_3, the most inconsistent parameters, are associated with outdoor temperature and global solar radiation, which were obtained from the local weather station while a and b_3, relatively consistent parameters, are associated with indoor temperature and HVAC status that were real measurements. It is also notable that b_4s can be split into two groups. Internal gains from Mar. 5 to Mar. 17 might be much bigger than those from the rest of days.

For the house M5 case in Fig.3.35, a_1, b_{23}, and b_{41} show minor changes over the whole test period. However, there was no consistency in b_1. Unlike the house A4, b_3, the parameter associated with solar radiation, is all zero except for three days. A relatively small effect of solar radiation on the indoor temperature during winter might be reflected in the parameters. A major difference from the above result is that $b_{41}$is fairly constant in this house.

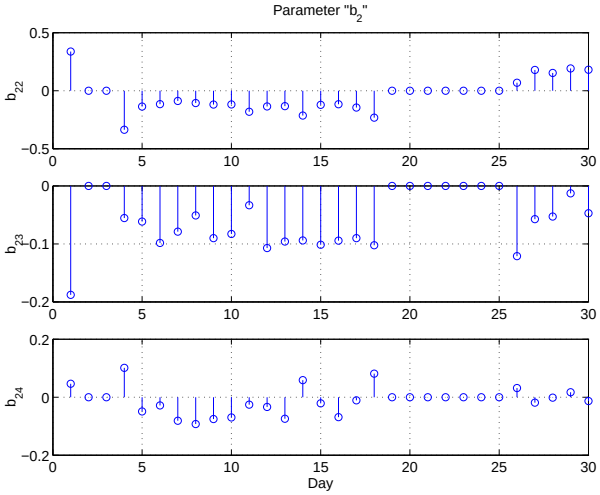
One of issues in this method is how to filter anomalies that result in learning wrong



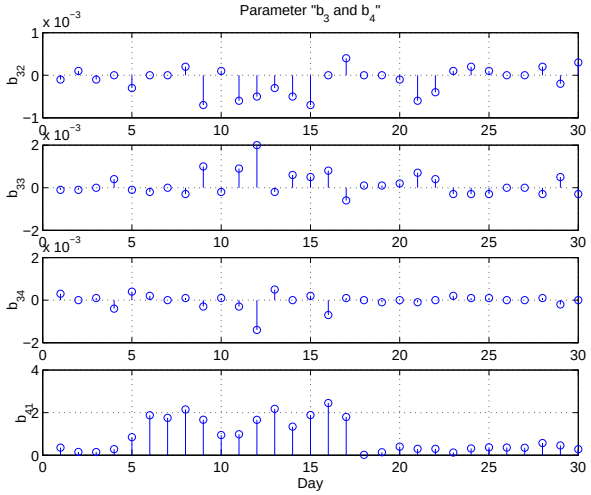
(a) Parameter a



(b) Parameter b1

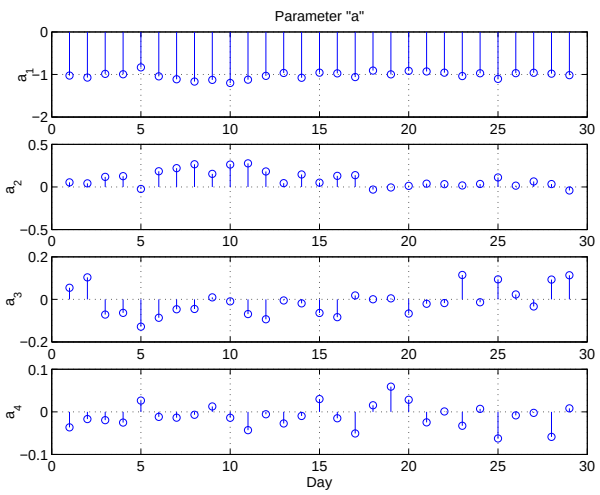


(c) Parameter b2

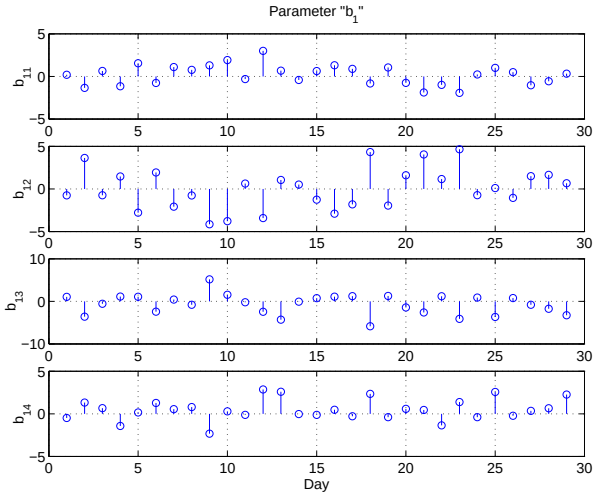


(d) Parameter b3 and b4

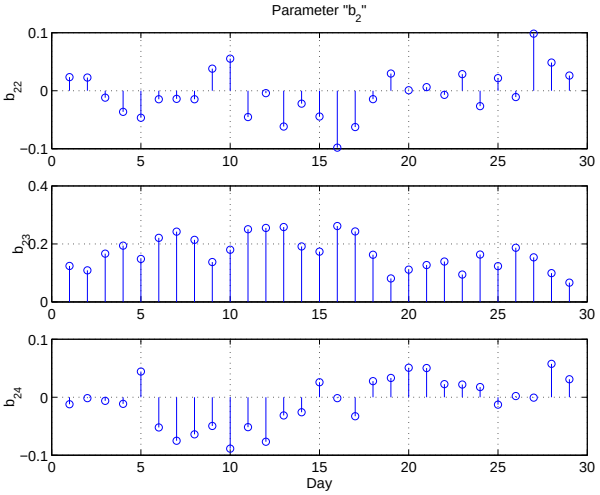
Figure 3.34: Consistency of ARX Model Parameters (House A4, Mar.1 - Mar.30, 2008)



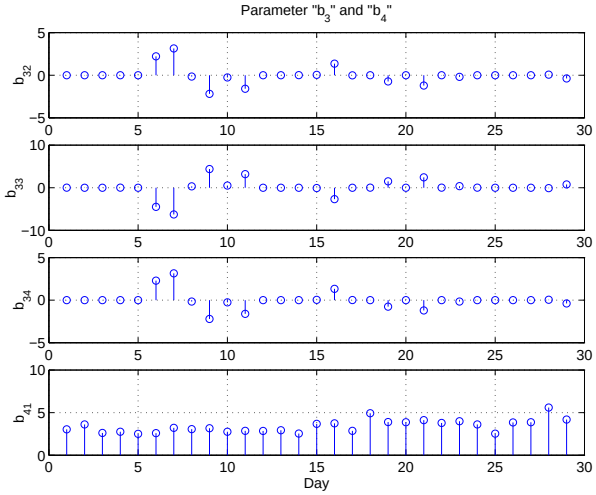
(a) Parameter a



(b) Parameter b1

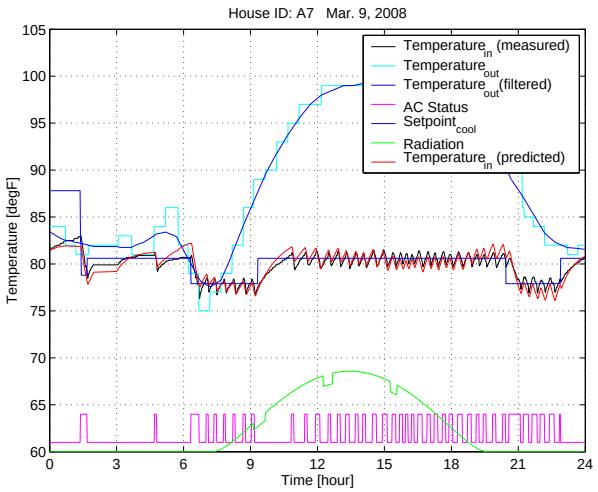


(c) Parameter b2

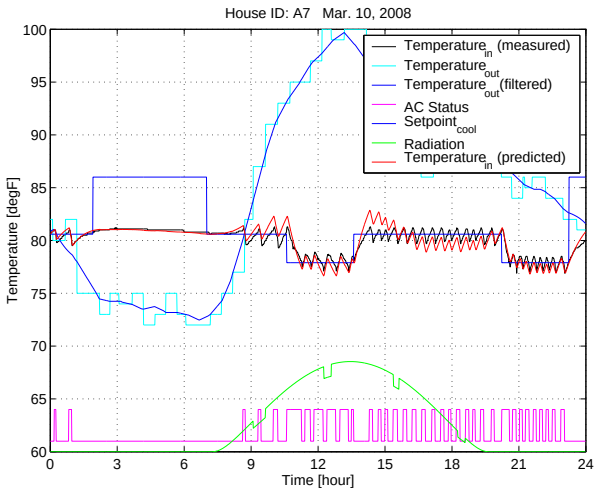


(d) Parameter b3 and b4

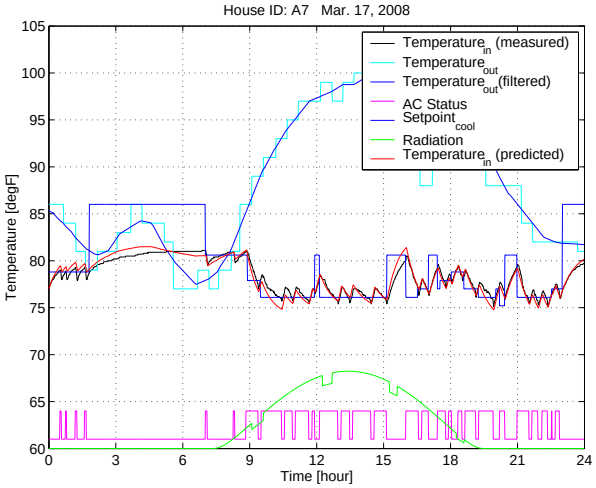
Figure 3.35: Consistency of ARX Model Parameters (House M5, Jan.1 - Jan.29, 2008)



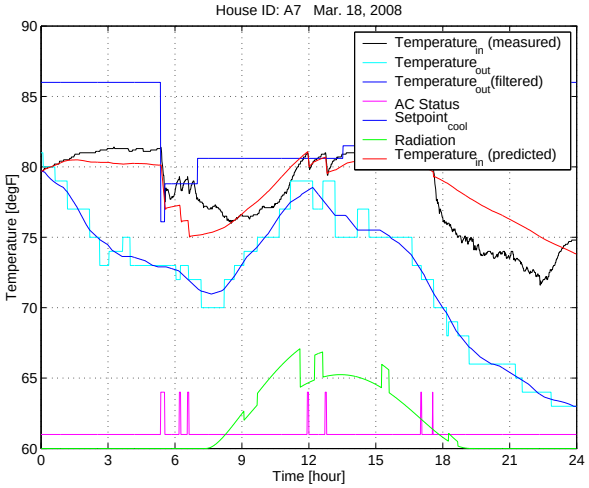
(a) March 9, 2008



(b) March 10, 2008



(c) March 17, 2008



(d) March 18, 2008

Figure 3.36: Anomaly Detection

parameters. Anomalous data can be obtained when the occupant keeps windows open, turns on an additional heater or cooler, or invites other people. A broken sensor or system malfunction may also cause anomalies. The following 4 plots in Fig.3.36 shows potentialities of the ARX model in both filtering anomalies for parameter learning and detecting system failure in real time. For prediction, the ARX model used the parameters that were calculated based on the same day as the prediction date. It should be noted that the parameter learning from the same day data is only useful to filter the occurred anomalies but useless for real-time fault detection.

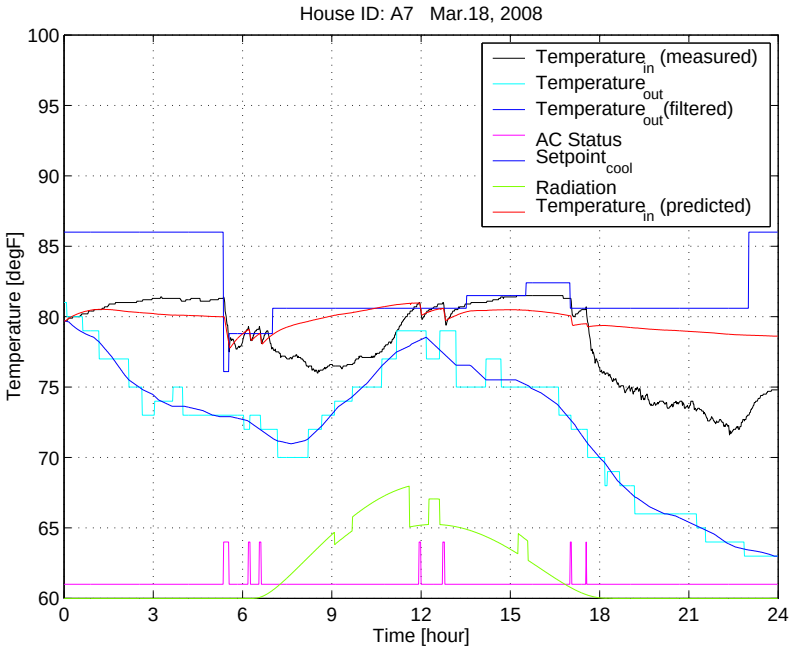


Figure 3.37: Real-Time Anomaly Detection

While the RMS error between the real indoor temperature and simulated temperature is small in the first three plots (Fig.3.36(a), (b), and (c)), the fourth plot shows a large error overall. This implies that the given ARX model could not realize the thermal behavior of the house on that day in particular. Since the model provided good prediction results for the other three days, an anomaly might exist. From the indoor temperature and air-conditioning status, it turns out that the real temperature from approximately 7am to 12pm and after 6pm was influenced by unusual factors. The noisy indoor temperature during these periods suggests that the windows were open and cool outside air was mixed with the inside air.

As the anomaly detection described above uses the same day data, the consistency of the parameters is not important. It merely determines whether the ARX model in the predefined order is matched with the real indoor temperature or not. It is useful when the calculated parameters considerably change over time. At the same time, it is possible that this method cannot detect anomalies if these anomalies are well fitted with the model. Fig.3.37 shows the anomaly detection with the parameters based on the one day ahead

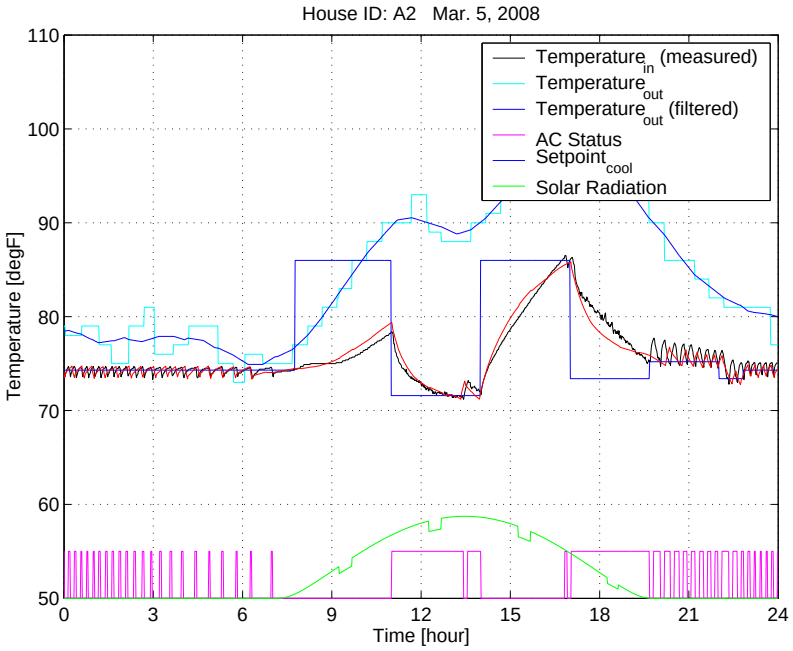


Figure 3.38: Prediction by the ARX Model Learned from the Previous Day Data (House A2, Mar.5, 2008)

data. Compared with Fig.3.36, the predicted temperature in Fig.3.37 is more realistic. Because this method uses the parameters already learned before prediction, real-time fault detection is possible. However, reliability of the parameters is crucial for it.

The final prediction result using the ARX model is shown in Fig.3.38. Instead of using the same HVAC on/off time as the real system, on/off signals based on the given setpoint were provided to the ARX model. In order to check improvement in prediction for long-term events, the data from house A2 were selected. House A2 was precooled from 11am to 2pm and there was no air-conditioning for four hours right after the precooling event on Mar.5, 2008. Initial temperature input for the ARX model was synchronized with the actual indoor temperature only at 12am, and there was no compensation for 24 hours. The parameters were calculated from the data on Mar. 4, 2008. Predicted temperature followed the real indoor temperature with only small error. Both the long-term cooling and long-term off events were also successfully predicted.

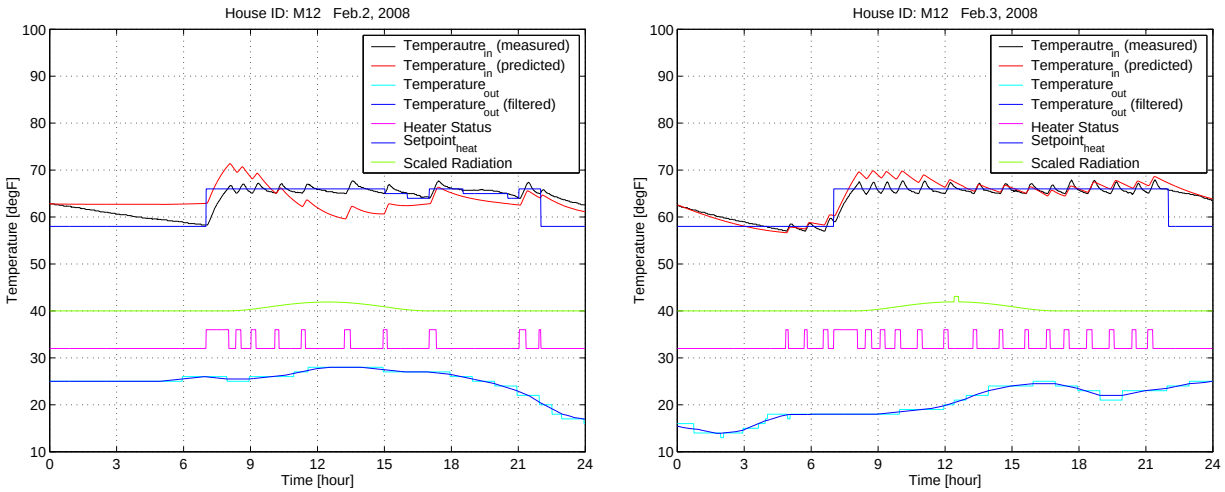
Chapter 4

Multiple-Model Switching Algorithm for House Thermal Behavior Prediction

4.1 Problem Description

In chapter 3, three different approaches for house dynamic signature learning were suggested and their performance was evaluated by comparing the actual indoor temperature measurement and the estimation from the models. The 1st-order model reflecting on basic heat and mass transfer characterized house thermal properties with four parameters. The measure of house insulation and air-conditioning could be inferred directly from the magnitude of parameters associated with the conduction and infiltration, internal heat gain, intensity of solar radiation, and supplied heat from the HVAC equipment. However, the model was not delicate enough to represent the dynamics of air-conditioning. The tabular method, despite inaccuracy in long-term prediction, showed improved performance for short-term events. While the ARX model provided the best performance among three methods, the prediction quality varies over time, which ultimately reduces reliability of this method.

Two prediction results from the same house on two consecutive days were shown in



(a) On Feb.2, 2008 (House M12) (b) On Feb.3, 2008 (House M12)

Figure 4.1: Variation of Prediction Quality in the ARX Model Method

Fig.4.1. There is definite variation of prediction quality over time. The model parameters for prediction were calculated from 1 day ahead 24-hour-long data and the heater on/off input for prediction was based on the actual heater status instead of the setpoint. The predicted indoor temperature on the first day (Fig.4.1(a)) was far from the actual measurement. While the cold outside temperature lowered the actual indoor temperature by five degrees for the first seven hours, the prediction model could not represent this. On the other hand, the prediction on the second day (Fig.4.1(b)) seems to be successful. The predicted indoor temperature during the first long-term off and heater-on events starting at 12am and 7am respectively are close to the actual behavior. Except the error that originated mainly with underestimation of the heat transfer due to temperature difference, the overall prediction quality is acceptable.

As discussed in the previous chapter, parameters of the ARX model fluctuated over time. If two consecutive days generate similar parameter sets, second day prediction based on the first day would be accurate. It is because the prediction with the same day data in the ARX model method normally does not result in large errors, except when there are anomalies present. Therefore, it tends to determine success of prediction which data were used for parameter calculation. Ten-day-long data can decrease the probability of

poor prediction caused by totally different parameters. However, more data can generate blunt parameters with which disastrous prediction may not happen but, at the same time, accurate prediction is impossible. It is also controversial that the latest data set is always beneficial for parameter calculation.

In the case that blow-by-blow modeling is impossible and, even with the perfect model, behavior of the system is frequently influenced by uncertainty, an adaptive technique would be useful to increase consistency of the prediction quality. Adaptive prediction, in the similar sense of adaptive control, involves modifying the prediction law to cope with the fact that the parameters of the system are slowly time-varying or uncertain. The following two methods adjust model parameters by selecting one of parameter sets among several existing candidates instead of calculating a new set. That is the reason why they are called model switching.

4.2 Proximity-Based Model Switching Algorithm

4.2.1 Algorithm Structure and Proximity Calculation

One of multiple-model methods that is widely used for state prediction is the Interacting Multiple Model (IMM) algorithm. The IMM algorithm, one of the most cost-effective filters for tracking maneuvering targets, was introduced in the 1980s by H.A.P. Blom as a safe adaptation or soft switching approach [34]. The basic idea of the IMM algorithm is to weigh each mode by mode probability and combine these estimates with no hard decision. Different modes are defined as different dynamics of the system. For example, the acceleration mode and constant velocity mode have different dynamics, which can be represented in different state-space models. The algorithm performs four sequential processes (interacting of estimates, model-conditional filtering, updating mode probability, and combining estimates) recursively.

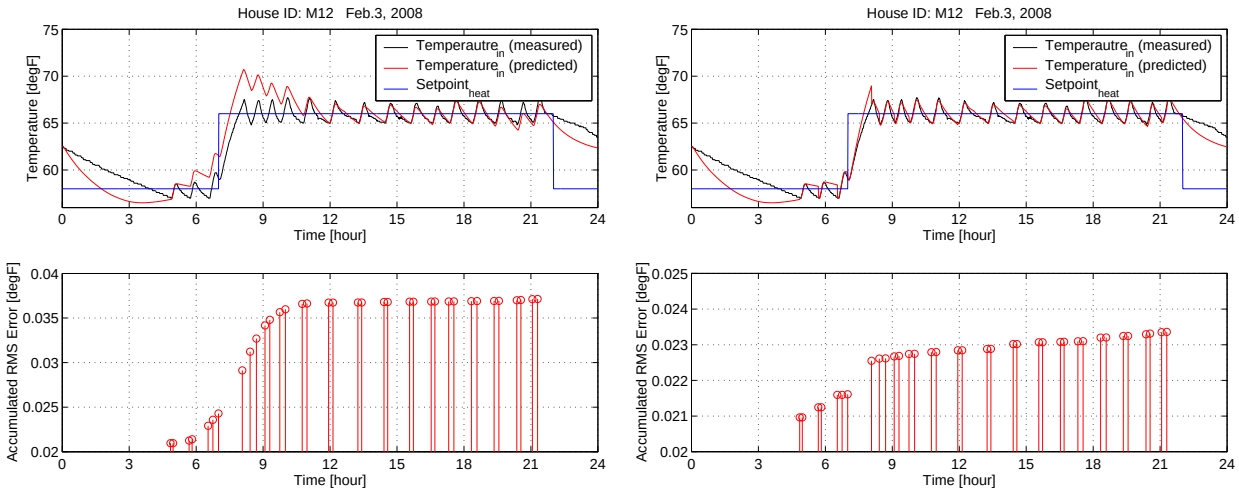
Due to the different system characteristics and different prediction constraints, the IMM

algorithm cannot be applied to the temperature prediction directly. While the IMM algorithm depends on real-time error calculation between actual sensor measurement and estimations from multiple models, estimation (or prediction) in our system should be performed at least a couple of hours ahead of measurement acquisition. This means that the first two steps in the IMM algorithm are useless for our algorithm.

In the proposed model switching algorithm, the models are distinguished only by different parameters. Except parameter difference, they are represented in the exact same ARX model that is given in chapter 3.5. The parameters of each model, which are calculated from 24-hour-long data (from 12am to 11:59pm), provide the best fit to the indoor temperature measurement. The algorithm used the recent seven models that are corresponding to the seven days prior to the current day on which indoor temperature is predicted because seven days are assumed to be long enough to cover the outside weather characteristics. Further, a seven day time frame is appropriate from a computational cost point of view. Model index implies which day the model corresponds to. For example, the parameters of model 3 are calculated three days prior to the current day.

The mode probability in the IMM algorithm is analogous to proximity in the multiple-model switching algorithm. The proximity denotes how closely the model represents the current temperature behavior with the same inputs. It can also be considered that parameters of two models are very close to each other. In order to measure proximity of the model candidates, sample data from the same day when prediction is scheduled should be selected. The sample data can be any data that are available at the prediction point.

The proximity is basically evaluated from an RMS regression error. Small RMS error implies high proximity between the system and model. Depending on how to calculate the error, seven models can be listed in different order. Two factors, the HVAC on/off input and temperature synchronization, were considered in error calculation. In order to evaluate the prediction performance of a given model, the HVAC on/off input for the model was determined by the setpoint in chapter 3. The thermostat control allows the predicted



(a) Without Temperature Synchronization (b) With Temperature Synchronization

Figure 4.2: Proximity Evaluation from Different Error Calculation

indoor temperature to follow the actual trajectory of the measurement. Large error between estimation and measurement, however, does not necessarily mean poor prediction in this case because time lag can seriously accumulate error. Using the same HVAC on/off input, the measured indoor temperature and four indoor temperature inputs[1] for the model were synchronized at the beginning of every new HVAC event. The event-based synchronization prevents undesirable error accumulation that is not caused by poor prediction performance and, at the same time, allows sufficient penalty for mismatches.

Fig.4.2 demonstrates how the event-based indoor temperature synchronization eliminates accumulated error that is irrelevant to prediction performance. Compared with Fig.4.2(b), error occurring from 6am to 11am in Fig.4.2(a) was mainly generated by the asynchronous initial condition. However, the big jump around 7:30am in Fig.4.2(b) indicates that penalty was also imposed on the faulty prediction.

Based on the error from the seven models, indoor temperature is predicted from two different methods. If a single model that has least error is used for prediction, it is called a *hard* switching method. On the other hand, if the estimates from the seven models are

[1]The ARX model uses past 4 indoor temperatures, $x(t-1)$, $\cdots$,$x(t-4)$ to calculate current indoor temperature, $x(t)$.

weighted according to their RMS error sizes and combined, it is called a *soft* switching method. In soft switching, the estimates prediction is calculated by

$$w(n) = \frac{\frac{1}{e(n)}}{\sum \frac{1}{e(i)}}, \quad i = 1, 2, \cdots, 7 \tag{4.1}$$

$$T_{in}(t + \Delta t) = \sum w(i) T_{in}^{i}(t + \Delta t), \quad i = 1, 2, \cdots, 7 \tag{4.2}$$

where $w(n)$: weight of the nth model

$e(n)$: RMS error between measurement and estimation from the$n$th model

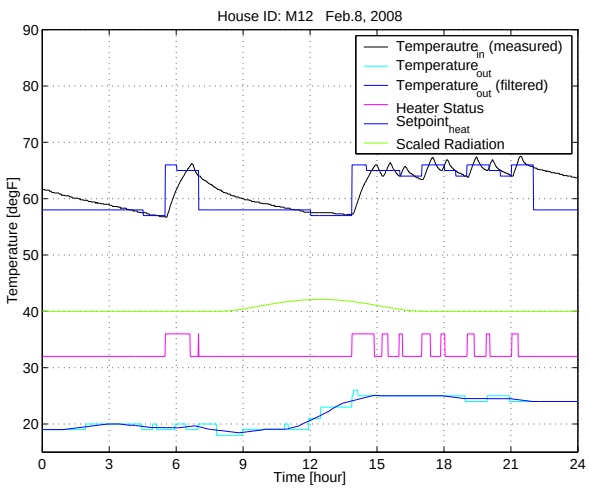
$T_{in}(t + \Delta t)$: final predicted temperature at $t + \Delta t$

$T_{in}^{i}(t + \Delta t)$: predicted temperature from the ith model at $t + \Delta t$

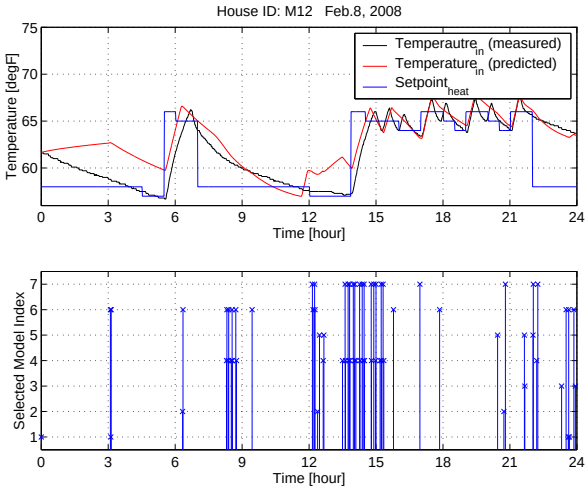
4.2.2 Results from the Multiple-Model Hard Switching (MMHS)

The MMHS algorithm was applied to one winter house (House M12) and one summer house (House A2). The results are shown in Fig.4.3 and 4.4 respectively. The algorithm predicted 3-hour ahead indoor temperature, and model proximity was calculated based on the sample data from 3 hour prior to the prediction point. The model that had the highest proximity (least regression error) was used to predict indoor temperature. The lower plots of Fig.4.3(b) and Fig.4.4(b) display which model was selected in this 24-hour simulation. HVAC operation for the predicted indoor temperature was controlled by the scheduled setpoint instead of following the original HVAC commands.

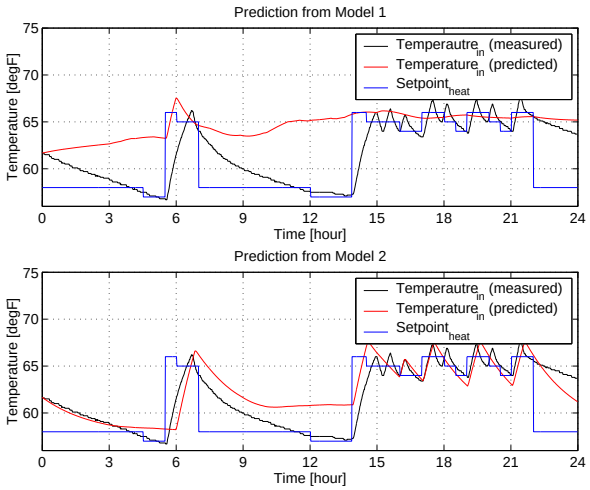
In the upper plot in Fig.4.3, the predicted indoor temperature by the MMHS algorithm and the real indoor temperature are compared. While the predicted indoor temperature did not considerably deviate from the real value (largest offset is 4.5F), the algorithm could not achieve delicate prediction. During the short-term on/off events from 3pm to 9pm, the on/off frequency of the prediction was much smaller than the actual frequency due to the relatively slow temperature decrement during off periods. Model switching between two models whose parameters are divergent resulted in unrealistic temperature behavior



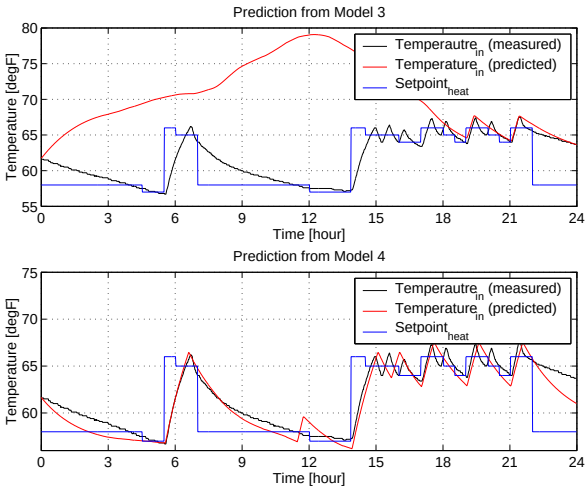
(a) Measured Data of the Target Period



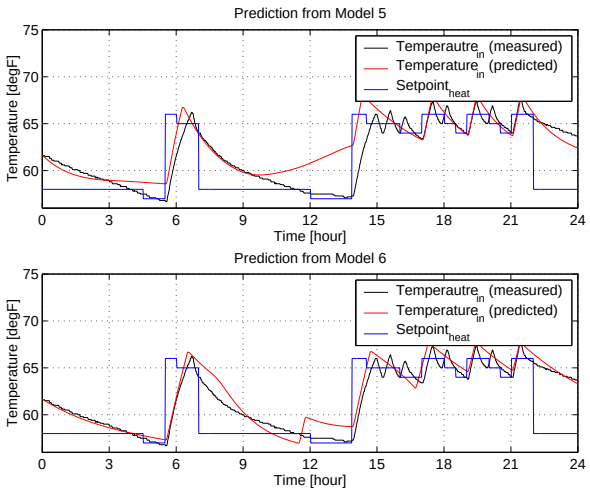
(b) Prediction from the MMHS



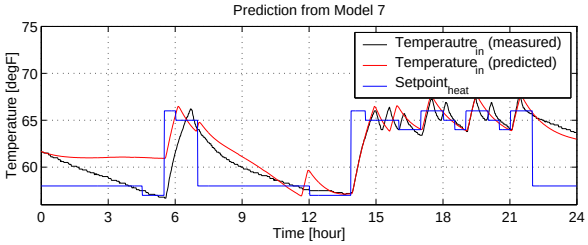
(c) Prediction from Model 1 and Model 2



(d) Prediction from Model 3 and Model 4

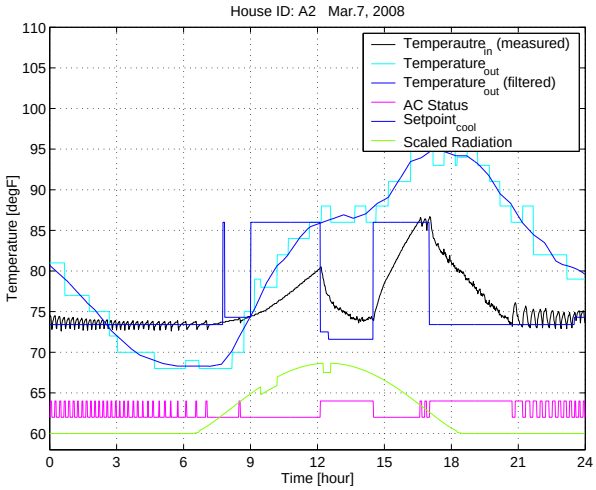


(e) Prediction from Model 5 and Model 6

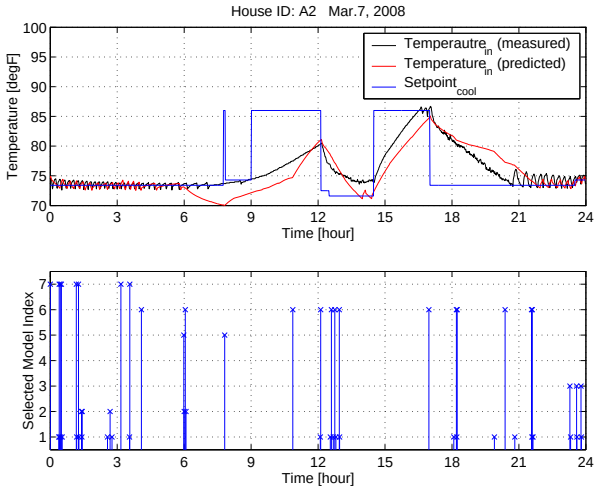


(f) Prediction from Model 7

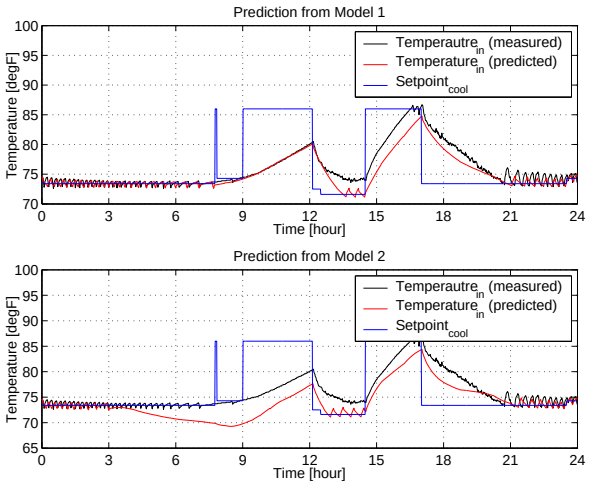
Figure 4.3: 3-Hour Ahead Prediction by the MMHS Algorithm (House M12, Feb.8, 2008)



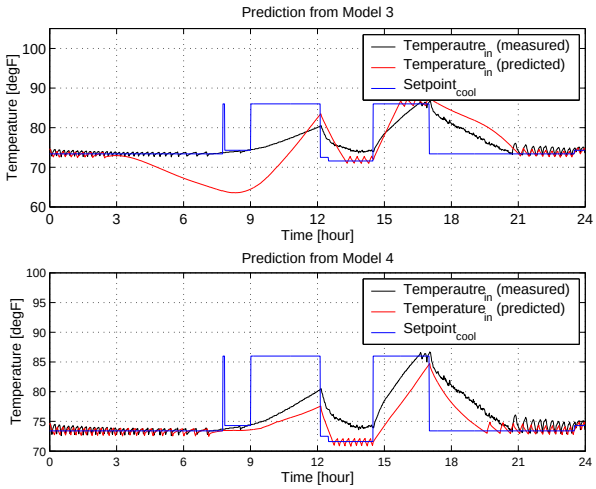
(a) Measured Data of the Target Period



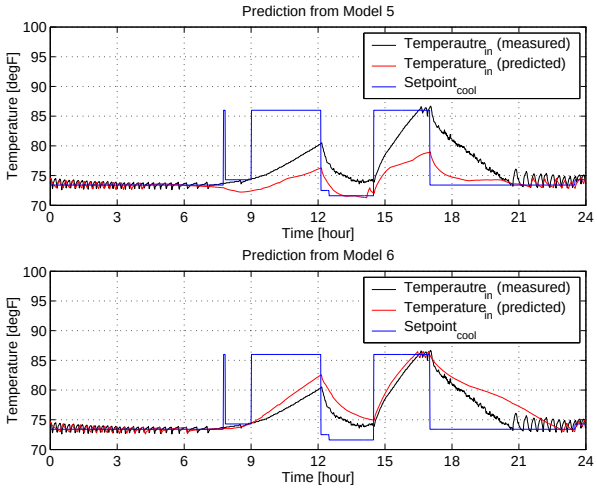
(b) Prediction from the MMHS



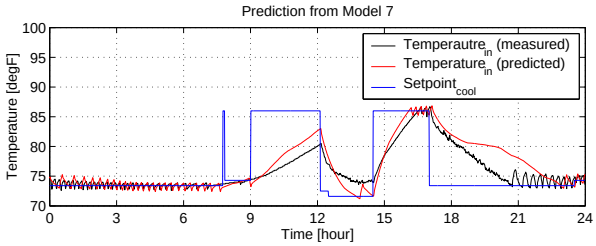
(c) Prediction from Model 1 and Model 2



(d) Prediction from Model 3 and Model 4



(e) Prediction from Model 5 and Model 6



(f) Prediction from Model 7

Figure 4.4: 3-Hour Ahead Prediction by the MMHS Algorithm (House A2, Mar.7, 2008)

around 12:10pm and 1:30pm. The model switching occurs more frequently as the size of sample data for proximity measure decreases. Frequent model switching induces abrupt temperature changes.

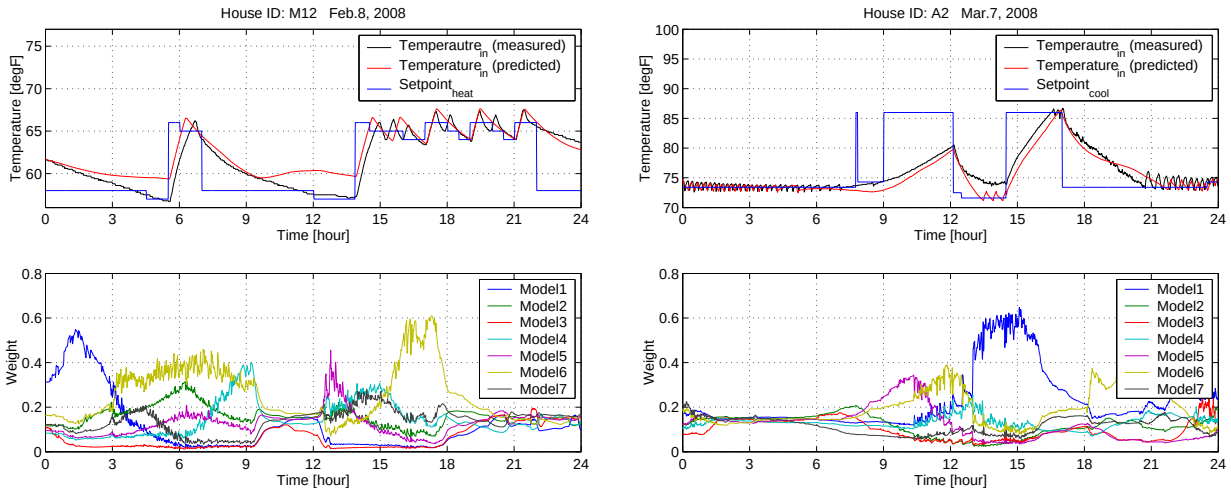
It turns out that model proximity does not provide an acceptable criterion for determining suitability of the model in prediction especially when the algorithm should predict further indoor temperature. The MMHS method, however, provides one way to select the least risky[2] model from the concrete criterion. The prediction results from the individual model (Model 1 to Model 7) are shown in Fig.4.3(c)(d)(e) and (f). The results include bad predictions from Model 1 and Model 3 and fairly good predictions from Model 4 and Model 6. When it is not certain which model would provide the best performance, the MMHS definitely gives the optimal choice.

In the case that the model proximity did not reflect future house behavior at all, the MMHS algorithm fails to provide reasonable prediction. The prediction result from the summer house A2 in Fig.4.4 is one example of poor performance. None of models provided good prediction during the entire period, but some of them partially showed good congruity with the real temperature. For example, the prediction from Model 1 was good, with the exception of three hours from 12pm to 3pm. Model 3 gave sufficient prediction quality during the exact same period when Model 1 failed to predict temperature correctly. However, the MMHS algorithm did not choose the correct model, which resulted in bad prediction for both short-term events and long-term events. (See Fig.4.4(b).)

4.2.3 Results from the Multiple-Model Soft Switching (MMSS)

Although the MMHS algorithm reduces the possibility that the prediction based on a single model would result in the worst performance, it failed to achieve accurate prediction. Discontinuity due to model switching is one of the main drawbacks in the MMHS algorithm. It is risky to use only one model with the least regression error for prediction, especially in

[2]The least risky model does not guarantee the best output but it might avoid the worst prediction performance.



(a) House M12 (Feb.8, 2008) (b) House A2 (Mar.7, 2008)

Figure 4.5: 3-Hour Ahead Prediction by the MMSS Algorithm

the case that each model has similar error distribution while parameters of the seven models are totally different. Instead of ignoring the rest of the models, weighting each model and then combining the estimations can reduce the risk of selecting a wrong model. Because all models reflect on the house thermal dynamics to some extent, the MMSS algorithm can be a reasonable approach.

In order to make it convenient to compare the performance with that of the MMHS algorithm, the identical houses on the same date were used for simulation. The lower plots in Fig.4.5 show each model's weight that was calculated from Eq.4.1. Most importantly, it is obvious that the discontinuity in the predicted temperature disappeared and the prediction quality was improved.

While the prediction from 12am to 3am by using Model 1 in Fig.4.3(b) is far from the actual temperature behavior, the prediction during the same period in Fig.4.5(a) is much closer to the actual temperature. Even though Model 1 was heavily weighted during the period, estimations from the other six models compensated the dominant effect of Model 1. The same compensation, however, also degrades the prediction in some periods. From 10am to 12pm, all models have similar weights, and Model 6 has the least error. Since the MMHS algorithm exclusively used Model 6 for prediction, and Model 6 provided great

prediction quality, the prediction was better than when all estimations from the seven models are equally weighted.

In the summer house, the MMSS did not considerably improve prediction quality except during the period from 6am to 12pm. The fact that all models have similar estimations and weights from 0am to 6am and one of the models is prominently dominant from 8pm to 10pm explains why two algorithms did not make a big difference.

4.3 Applicability-Based Model Switching Algorithm

4.3.1 Algorithm Structure and Applicability Calculation

The key of proximity-based model switching is to decide which model can generate the closest estimation to upcoming indoor temperature behavior. In order to evaluate the estimations of the models, a limited number of samples is used. When the sample data contains a single mode only, model selection tends to be biased. For example, if sample data are extracted from a long AC off period, the selected model may not provide good estimation for repetitive on and off cycles in three hours. In Fig.4.3(b), model 1 was used to predict the 12am to 3am period. The period associated with the sample data corresponds to 6pm to 12am in the previous day, which contains frequent on and off cycles. In the proximity-based model switching method, similarity of HVAC on/off patterns between a sample period and a target period is crucial to determine the prediction quality.

If some models can only represent a specific temperature behavior, it might also be valuable to take into account which model is most tolerable in typical situations. Therefore, valuation of each model requires more than one sample data. While the proximity-based model switching is an overly simplistic algorithm, a new algorithm should be rather comprehensive and less sensitive.

The applicability-based model switching algorithm also uses six models[3]. The six models

[3]The total number of models is flexible. In this paper, 6 models were used to compare the prediction

Day	-7	-6	-5	-4	-3	-2	-1	0 (Prediction Day)
Model 7	↻	e_{76}	e_{75}	e_{74}	e_{73}	e_{72}	e_{71}	$a_7 = \frac{6}{\sum e_{7i}}$
Model 6		↻	e_{65}	e_{64}	e_{63}	e_{62}	e_{61}	$a_6 = \frac{5}{\sum e_{6i}}$
Model 5			↻	e_{54}	e_{53}	e_{52}	e_{51}	$a_5 = \frac{4}{\sum e_{5i}}$
Model 4				↻	e_{43}	e_{42}	e_{41}	$a_4 = \frac{3}{\sum e_{4i}}$
Model 3					↻	e_{32}	e_{31}	$a_3 = \frac{2}{\sum e_{3i}}$
Model 2						↻	e_{21}	$a_2 = \frac{1}{\sum e_{2i}}$

e_{ni} : RMS error of n^{th} model w.r.t the data on $-i^{th}$ day
a_n : Applicability of the n^{th} model
↻ : Model creation

Table 4.1: Daily RMS Error of Models and Applicability Calculation

are represented in the exact same ARX model whose parameters are calculated from 24-hour-long data. Unlike the proximity-based model switching algorithm, the latest ARX model (Model 1) is not used. For example, the models from Feb. 6 (Model 2) to Feb. 1(Model 7) are used for prediction on Feb, 8. Applicability of a model is conceptually the same as the proximity in the proximity-based model switching algorithm in that it is based on an RMS error between measurement and estimation. The only difference is that the sample is 24 hours long and updated in a 24-hour period.

Tab.4.1 explains how the applicability of each model is updated. In order to accord the model ID number with that in the previous method, a negative day number is used. When prediction is performed on Day 0, six models that were created on Day -7 to Day -2 are considered. The model number indicates how far the day of model creation is from a prediction day. Error between measurement and estimation from the model is calculated from the next day of the model creation. In the nth model case, total n-1 RMS errors can be accumulated. That is why this method cannot use Model 1 for prediction. Applicability is defined as an average of RMS errors for at most six days.

Similarly to the previous algorithm, both hard switching and soft switching were tested.

result with that of the proximity-based model switching algorithm.

In the hard switching, the model with a maximum applicability (or minimum RMS error e) is used for prediction during a 24-hour period. In the soft switching, each model is weighted according to its applicability. The initial model switching law is given by the following equation.

$$w(n) = \frac{a(n)}{\sum a(i)}, \quad i = 2, 3, \cdots, 7 \tag{4.3}$$

$$T_{in}(t + \Delta t) = \sum w(i) T_{in}^{i}(t + \Delta t), \quad i = 2, 3, \cdots, 7 \tag{4.4}$$

where $w(n)$: weight of the nth model

$a(n)$: applicability, or reciprocal of averaged RMS error of the nth model

$T_{in}(t + \Delta t)$: final predicted temperature at $t + \Delta t$

$T_{in}^{i}(t + \Delta t)$: predicted temperature from the ith model at $t + \Delta t$

4.3.2 Results from the Multiple-Model Hard Switching (MMHS)

The applicability-based MMHS algorithm applied to two houses, House M12 and House A2. In order to predict indoor temperature in four consecutive days, RMS errors of nine models were calculated. The daily RMS error of each model should be updated at most six days. The errors of the nine models that were created from Feb.1 (Model 9) to Feb.9 (Model 1) in order are summarized in Tab.4.2. The model numbers in parentheses are the exact same numbers as those used in the proximity-based multiple model algorithm. Among the nine models, one model (Model 8) kept a small error as the day changed, while one model (Model 2) kept a large error. Some models (Model 5 and Model 6) showed large variance. According to the values in the table, Model 8 shows the highest applicability.

Fig.4.6 contains prediction results and applicability of each model. Contrary to the proximity-based multiple model algorithm, the prediction was performed ahead of 24 hours, and computational cost is very low. In the first two days (Feb. 8 and Feb. 9, 2008), the prediction was performed with Model 8(6*). When the predictions from the individual sex

Day	Feb.1	Feb.2	Feb.3	Feb.4	Feb.5	Feb.6	Feb.7	Feb.8	Feb.9	Feb.10
Model 9(7*)	↻	0.0280*	0.0288*	0.0239*	0.0293*	0.0332*	0.0282*			
Model 8(6*)		↻	0.0109*	0.0150*	0.0146*	0.0141*	0.0152*	0.0128		
Model 7(5*)			↻	0.0359*	0.0399*	0.0269*	0.0263*	0.0249	0.0360	
Model 6(4*)				↻	0.0135*	0.0292*	0.0302*	0.0121	0.0384	0.0325
Model 5(3*)					↻	0.0804*	0.1523*	0.0680	0.1430	0.1739
Model 4(2*)						↻	0.0198*	0.0154	0.0147	0.0299
Model3							↻	0.0526	0.0460	0.0322
Model2								↻	0.0817	0.0734
Model1									↻	0.0252

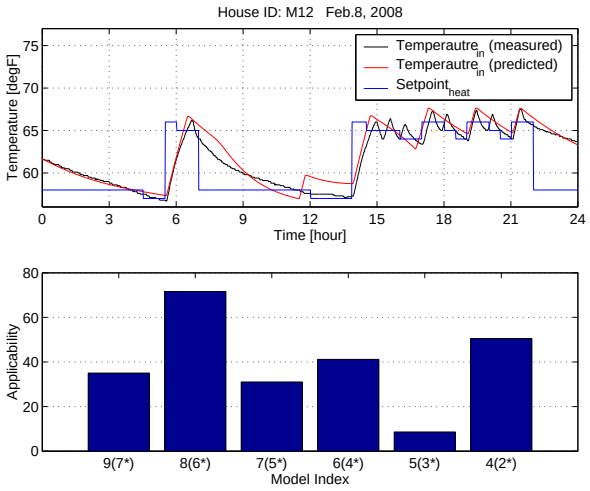
↻ : Model creation
*For prediction on Feb.8, 2008

Table 4.2: Daily RMS Errors of Models (House M12, Feb,1 - Feb.9, 2008)

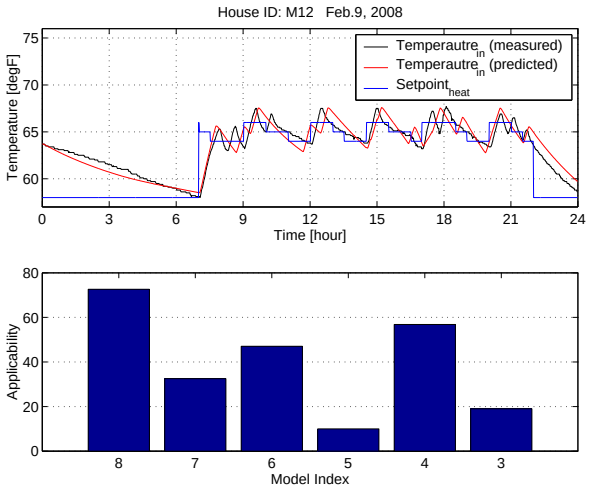
models are compared (See Fig.4.3), it turns out that Model 8 (Model 6 in Fig.4.3) provided the best prediction result. The predicted temperature on Feb. 9 also follows the actual temperature with relatively small error. Notably, the prediction for long-term events (12am to 8am and 10pm to 12am) is very accurate if the error due to initial offset can be ignored.

In the following two days (Feb. 10 and Feb. 11, 2008), Model 4 was used for prediction. Although the applicability of Mode 8 is higher than those of the other models at that time, the prediction did not follow the actual temperature, especially during the morning period. The poor prediction result, however, does not imply that this algorithm is not good without showing that the prediction results based on other models were compared with that of Mode 4.

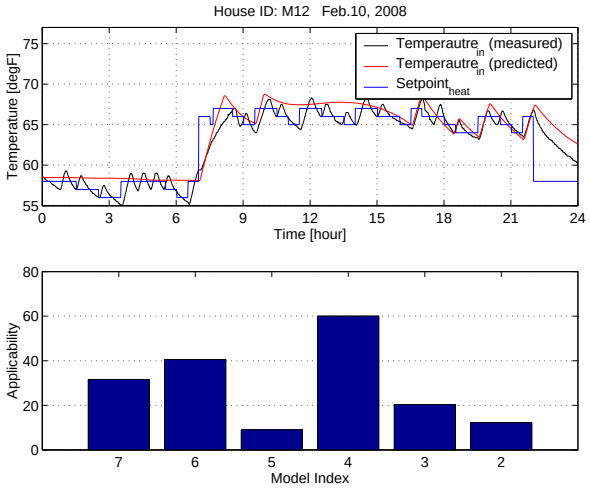
It is interesting that the applicability of Model 4 decreased by 10 when the day changed, which implies that the model may not be reliable for further prediction. On the other hand, the applicability of Model 8 did not change a lot, and the value is much higher than that of Model 4. Usually the larger model number among the six model candidates means that the model was validated for a longer period. If the selected model has a relatively small



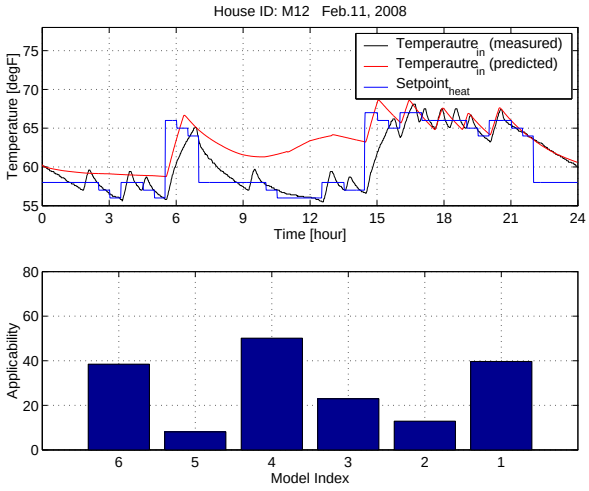
(a) Feb.8, 2008



(b) Feb.9, 2008



(c) Feb.10, 2008



(d) Feb.11, 2008

Figure 4.6: 1-Day Ahead Prediction by the MMHS Algorithm and Applicability of the Models (House M12, Feb.8 - Feb.11, 2008)

Day	Feb.28	Mar.1	Mar.2	Mar.3	Mar.4	Mar.5	Mar.6	Mar.7	Mar.8	Mar.9
Model 9(7*)	↻	0.0209*	0.0203*	0.0201*	0.1315*	0.0739*	0.0084*			
Model 8(6*)		↻	0.0349*	0.0241*	0.1158*	0.0494*	0.0161*	0.0157		
Model 7(5*)			↻	0.0136*	0.1090*	0.0602*	0.0156*	0.0274	0.0451	
Model 6(4*)				↻	0.1038*	0.0589*	0.0178*	0.0328	0.0510	0.0448
Model 5(3*)					↻	0.0505*	0.0271*	0.0199	0.0686	0.0439
Model 4(2*)						↻	0.0200*	0.0262	0.0460	0.0267
Model3							↻	0.0218	0.0425	0.0249
Model2								↻	0.0529	0.0239
Model1									↻	0.0237

↻ : Model creation
*For prediction on Mar.7, 2008

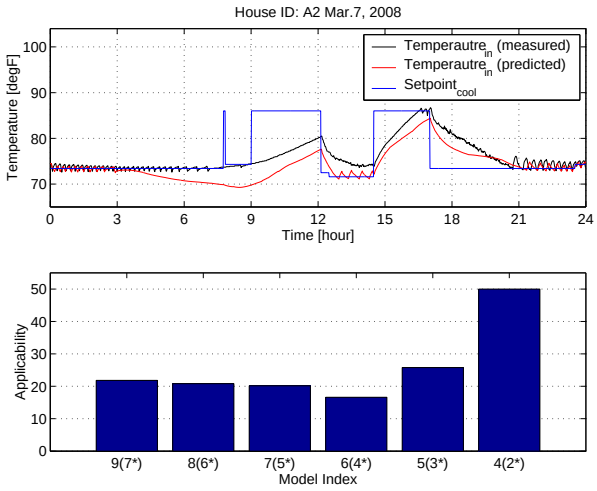
Table 4.3: Daily RMS Errors of Models (House A2, Feb,28 - Mar.8, 2008)

applicability value, the performance of the model may not be guaranteed.

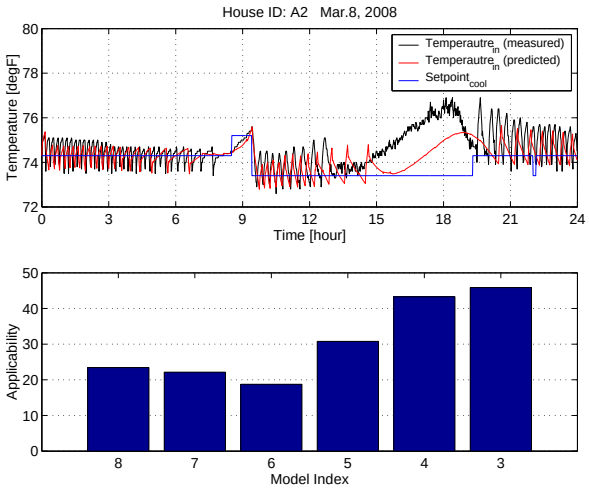
In order to avoid this, it might be useful to compare the models with the same validation period. In order for a model to have a long validation period, models close to a target prediction day cannot be used. For example, if a model should be validated for 6 days, only the model whose number is greater than 7 can be used.

Tab.4.3 summarizes the RMS errors of 9 models for House A2. While there exists the model that has a small error for all validation periods in the case of House M12, every model in House A2 has big error changes. It should be noted that all models have relatively big errors on Mar. 4. From the raw data, it turned out that there was an anomaly on Mar. 4. (Indoor temperature did not change for a long time due to system failure.) Since the error was accumulated, the anomaly on Mar. 4 affected all models that were created before Mar. 6. (Since Model 5 was based on the data from Mar. 4, this model might be wrong.) That is why the models whose numbers are greater than 4 were always excluded from the model selection.

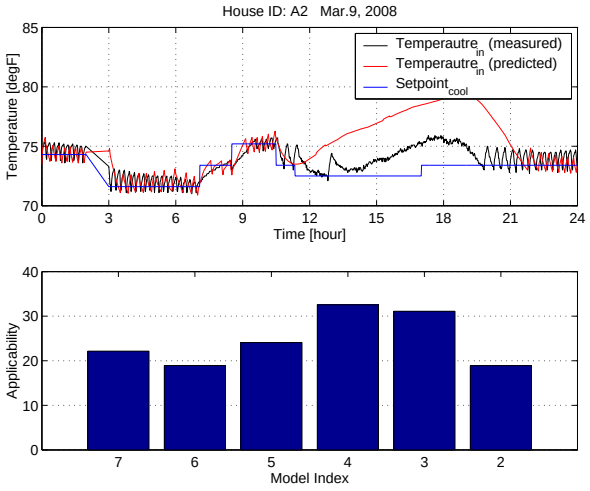
When the prediction results from other models are compared (See Fig.4.4), the appli-



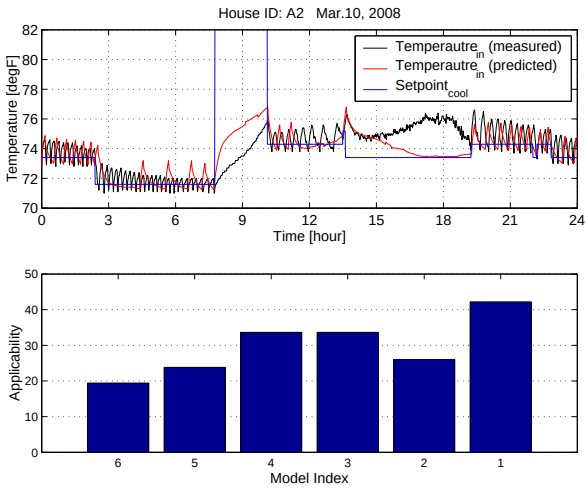
(a) Mar.7, 2008



(b) Mar.8, 2008



(c) Mar.9, 2008



(d) Mar.10, 2008

Figure 4.7: 1-Day Ahead Prediction by the MMHS Algorithm and Applicability of the Models (House A2, Mar.7 - Mar.10, 2008)

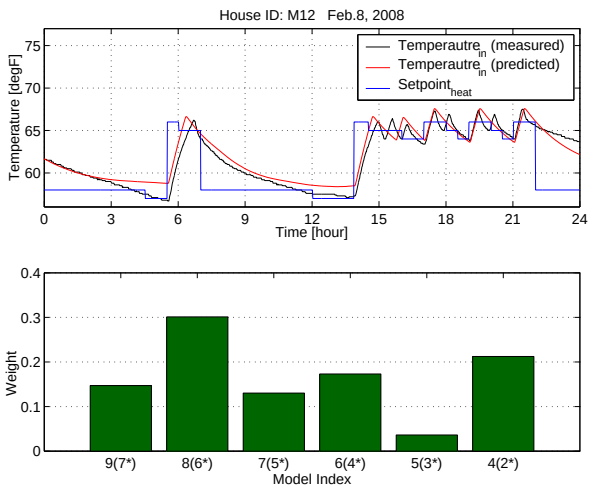
cability of Model 4(2*) in Fig.4.7 was overestimated. The result from House A2 shows how crucial the detection and exclusion of anomalies are in data analysis. Unfortunately, the prediction results in the other three days also did not satisfy the expected performance range.

4.3.3 Results from the Multiple-Model Soft Switching (MMSS)

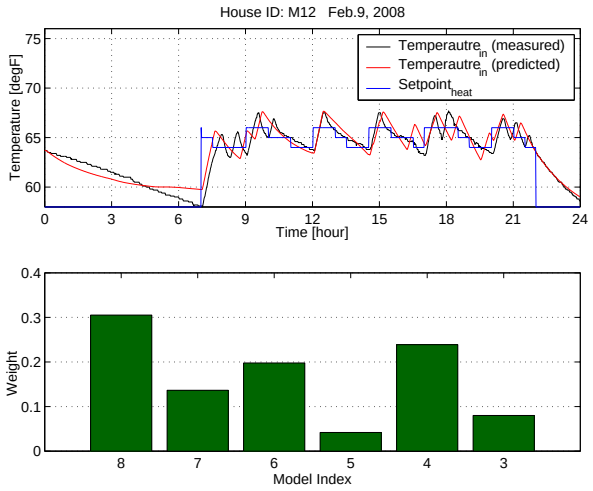
In the applicability-based method, as in the proximity-based method, the MMSS algorithm improved the prediction quality by mixing all estimations from six models. While the proximity-based MMSS algorithm was initially intended to remove the discontinuity due to frequent model changes, the MMSS algorithm in the applicability-based method is purely used to reduce the risk from excessive dependence on a single model. It was shown that the unfiltered data that contain anomalies seriously affects prediction afterward. Although the anomalies certainly degrade the quality even in the MMSS, it is expected that mixing the estimations from six models itself can play a filtering role, which attenuate the effect of anomalies.

Fig.4.9 includes temperature prediction by the MMSS in House M12. Comparison between the results in Fig.4.5, Fig.4.6, and Fig.4.9 clarifies the difference and improvement. The prediction in the first prediction day (Fig.4.9(a)) is better than those from the other six models, and also from seven models if Model 1 were also included. The prediction from a single model can be checked in Fig.4.3. The prediction in the following three days was also improved to some degree, especially in the last morning period for which the MMHS failed to predict the temperature close to the actual profile.

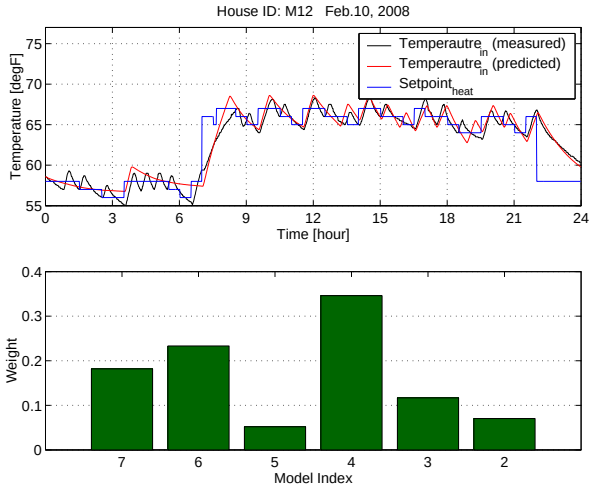
In Fig.4.9, the MMSS algorithm was applied to House A2. The overall prediction error in the first day was reduced quite a bit. The lowest prediction temperature increased by 2F (69F in the MMHS and 71F in the MMSS around 8am), and the offset between prediction and measurement from 3pm to 9pm decreased substantially. The anomalies of the data, however, were not reduced by mixing estimations. Instead, the mixing process made the



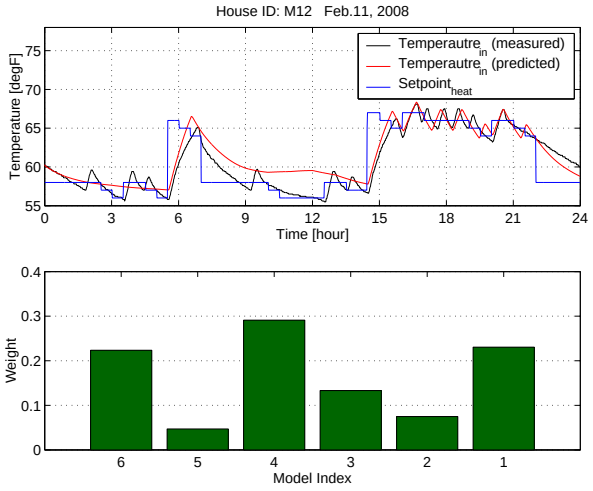
(a) Feb.8, 2008



(b) Feb.9, 2008



(c) Feb.10, 2008

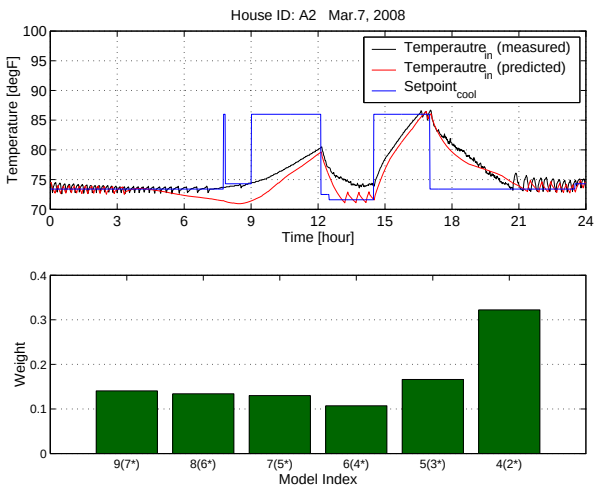


(d) Feb.11, 2008

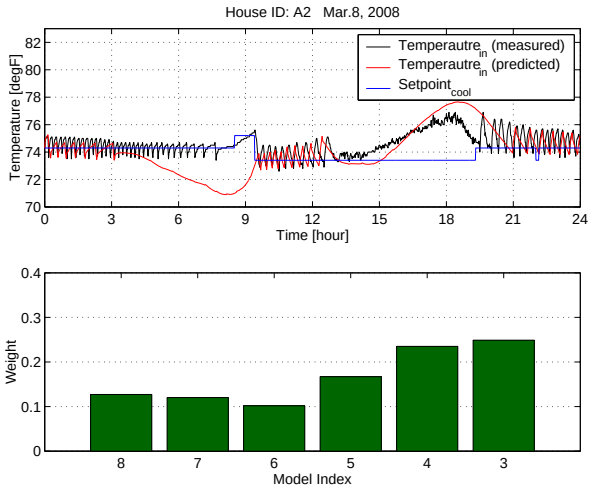
Figure 4.8: 1-Day Ahead Prediction by the MMSS Algorithm and Weights of the Models (House M12, Feb.8 - Feb.11, 2008)

prediction worse by increasing the portion of the poor models based on anomalies.

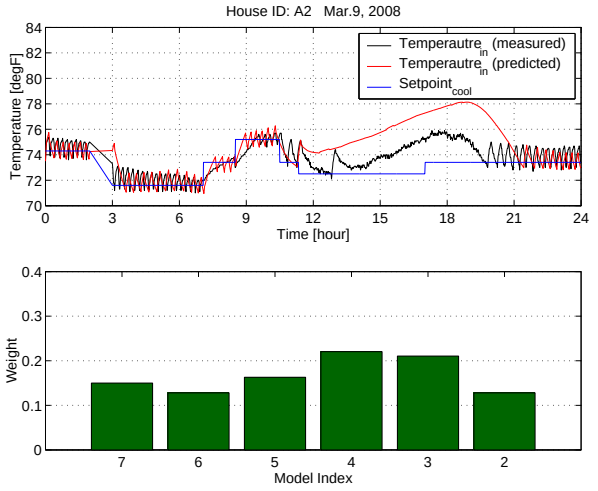
Despite the limitations, the MMSS algorithm provided the best prediction results when the data for model developing were properly filtered. However, the current applicability calculation may not be proper in that each model averaged the different number of error sets. In the case that a specific day causes considerably large error in all corresponding models due to either anomalies or the model itself, the models that are created after the day can provide an advantage. To avoid the inequity, a fixed validation period would be required, as mentioned previously.



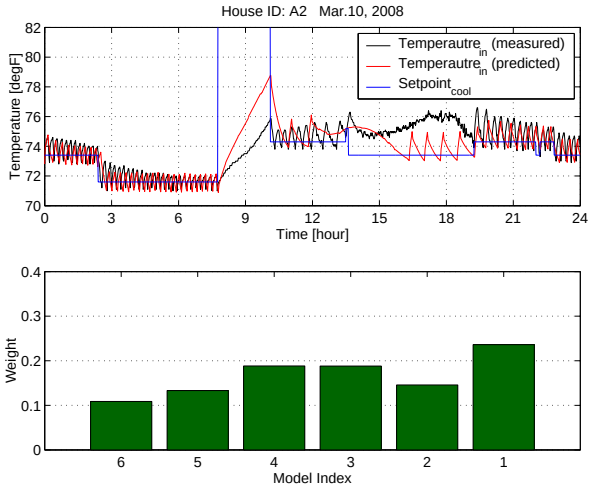
(a) Mar.7, 2008



(b) Mar.8, 2008



(c) Mar.9, 2008



(d) Mar.10, 2008

Figure 4.9: 1-Day Ahead Prediction by the MMSS Algorithm and Weights of the Models (House M12, Mar.7 - Mar.10, 2008)

Chapter 5

Summary and Conclusions

The goal of this research is to develop a new prototype of an energy management system specifically targeted for residential buildings and the learning algorithms with which indoor temperature can be predicted within tolerable error ranges. The system and algorithms should have met the following criteria:

- ◇ The system must be able to easily substitute for an existing HVAC control unit in a residential building without replacing other relevant elements.
- ◇ The system must be able to effectively respond to the dynamic utility price and minimize user discomfort.
- ◇ The control code must be able to work with any platform and any types of residential buildings.
- ◇ The control procedure inside the system must be autonomous, i.e. data analysis, decision-making, and interaction with occupants must be automated.
- ◇ Available information that the learning algorithm can use is limited.
- ◇ Computational load for the algorithm is limited.

The DREAM (Demand Responsive Electrical Appliance Manager) system, which is based on a wireless sensor network, has been successfully developed and was deployed to test the performance in the occupied houses. Several learning algorithms to identify the house thermal properties and predict indoor temperature have also been proposed, and their characteristics were analyzed with the data from both simulation houses and real occupied houses in various locations.

5.1 Demand Response Enabled Thermostat Development

Wireless sensor technology and computing technology have merged into a new generation energy management system for residential buildings. The Demand Response Electrical Appliance Manager (DREAM), based on a wireless sensor network, manages sensor data to optimize energy cost and user comfort as well as identify house characteristics and occupant patterns. This autonomous system is designed to reduce peak electrical demand and ultimately optimize energy management under the evolving electricity pricing policy.

DREAM consists of three elements: 1. wireless sensors and actuators to collect measured information and operate HVAC equipment and/or other appliances, respectively; 2. a graphical user interface to provide real-time energy use information and efficient interaction with users; and 3. a main control unit to analyze data and make the optimal decision in various situations with respect to temperature setpoint and choice of appliances. This decision is influenced by environmental factors (temperature and time of day), occupant information (thermal preference and monthly budget limit), and physical characteristics (house size and AC capacity).

According to the simulation results, optimization in the DREAM controller efficiently adjusted the setpoint to maximize user comfort according to the predefined economic index. Compared to the typical programmable thermostat, it successfully decreased the electricity use during high price periods without affecting the user comfort level. Actual functionality

and overall performance of the system was evaluated through approximately 40-day-long field tests in summer 2007. Although the system failure due to operating system freezing often occurred during the test, all functions in DREAM were properly operated. The prediction based on the default parameters was improved by using the parameters learned for 37 days in one house. Unfortunately, another house did not provide meaningful prediction results since the occupants in that house erratically opened the windows for house cooling, which substantially changes the house thermal properties.

5.2 House Dynamic Signature Learning through Data Analysis

Due to the significance of the house dynamic signature learning in the DREAM controller, identification algorithms were studied at some length. Unlike previous studies, the real data from the houses in Northern California, Minnesota, and Adelaide, Australia were used for algorithm validation and evaluation. Three different approaches, 1st order physical model method, tabular method, and 4th order ARX model method, were proposed, and their performance was evaluated.

The first method identifies the house thermal characteristics with the 1st order differential equation that is derived from the thermal influence of five heat sources (i.e. conduction, infiltration, solar radiation, internal heat gain, and either air-conditioning or heating). Through the parameters learned from two California houses as well as four simulation houses from the MZEST, it turned out that four parameters associated with conduction and infiltration, internal gain, solar radiation, and air-conditioning can be used to represent thermal characteristics of the house such as the level of insulation and air-conditioning efficiency. Despite simplicity of the model, the prediction from this method was not disappointing but this method was too simple to represent the actual temperature profile.

The tabular method was initially proposed to cover the nonlinear characteristics that

cannot be represented in simple numerical methods. The concept of the tabular method is to store indoor temperature change rate, the slope, with respect to outdoor to indoor temperature difference. By adding one more dimension, the effect of the solar radiation could be also considered. Due to the heat capacity difference between air and mass, the explicit slope change was shown from time-temperature plots. By determining the point at which temperature slope changes abruptly through linear regression error, six different tables (2 ignorable point tables, 2 off tables, 2 heater or AC on tables corresponding to both short-term and long-term periods) could be filled with slope. These tables were also used for temperature prediction, and the prediction results from most of winter houses in Minnesota were acceptable except for the long-term heater-on event. However, considering solar radiation that was obtained from the local weather station did not improve the prediction quality at all.

Identification through the 4th order ARX model was successful in that the model can follow the actual temperature profile with small error. Differently from the previous two methods, this method was very useful in predicting the long-term period events such as precooling and preheating. However, the prediction performance considerably fluctuated depending on which data were used for parameter learning.

5.3 Multiple-Model Switching Algorithm

The multiple-model switching algorithm was proposed to minimize performance inconsistency in the ARX model method. Instead of calculating the model parameters based on the predefined sample data set, it allows several models whose parameters are calculated from consecutive seven days. Among several candidates, the optimal model in a given condition can be used for prediction (Multiple-Model Hard Switching), or weighted estimations from the candidates can be combined to guarantee the continuity of output temperature as well as provide flexibility to uncertainty (Multiple-Model Soft Switching).

Depending on the criteria to select or weight a model, the algorithm is divided into proximity-based model switching and applicability-based model switching. The proximity-based model switching could not guarantee the best model selection both in the MMHS method and in the MMSS method. However, it could avoid the worst case all the time, which increases reliability of the algorithm. On the other hand, the applicability-based model switching improved the prediction quality. In particular, the applicability-based MMSS algorithm provided the better prediction than the best model among several candidates.

Bibliography

[1] S. Drewer and D. Gann, "Smart buildings," *Journal of Facilities*, vol. 12, no. 13, pp. 19–24, 1994.

[2] Z. Mo, "Intelligent buildings and intelligent agents - a human-centered framework for building controls," *Proceedings of the 41st SICE Annual Conference*, vol. 5, pp. 3151–3156, Aug 2002.

[3] S. Sharples, V. Callaghan, and G. Clarke, "A multiagent architecture for intelligent building sensing and control," *International Sensor Review Journal*, vol. 19, no. 2, pp. 135–140, May 1999.

[4] S. Buchanan and R. Taylor, "The electricity consumption impacts of commercial energy management systems," *IEEE Transactions on Power Systems*, vol. 4, no. 1, pp. 213–219, Feb 1989.

[5] E. Arens *et al.*, "Demand response enabling technology development, June 2003-November 2005," Center for the Built Environment, University of California at Berkeley, Berkeley, CA, Final Report Phase I, Dec 2006.

[6] ——, "Demand response enabling technology development, December 2005-December 2007," Center for the Built Environment, University of California at Berkeley, Berkeley, CA, Final Report Phase II, Mar 2008.

[7] T. Peffer *et al.* (2005) Sims spring 2005 dream interface design and development. http://www.sims.berkeley.edu:8000/academics/courses/is213/s05/projects/thermostat/index.php.

[8] B. Jain and A. Agrawala, *Open Systems Interconnection: Its Architecture and Protocols*. Amsterdam, Netherlands: Elsevier, Oct 1990.

[9] A. Auliciems, "Thermobile controls for human comfort," *Heating and Ventilating Engineer*, vol. 58, pp. 31–33, 1984.

[10] A. Lovins, *Air Conditioning Comfort: Behavioral and Cultural Issues*. Boulder, CO: E Source, 1992.

[11] S. Ubbelohde, G. Loisos, and R. McBride, "Advanced comfort criteria and annotated bibliography on adapted comfort. alternatives to compressor cooling," CEC Public Interest Energy Research (PIER) Program, Final Report Phase V, 2003.

[12] J. Woods, "Fiddling with thermostats: Energy implications of heating and cooling set point behavior," 2006.

[13] W. Kempton and S. Krabacher, "Thermostat management: Intensive interviewing used to interpret instrumentation data, pp. 245-262 in W. Kempton and M. Neiman (eds.) energy efficiency: Perspectives on individual behavior," Jun 1987.

[14] K. Okamoto-Mizuno, K. Mizuno, S. Michie, A. Maeda, and S. Iizuka, "Effects of humid heat exposure on human sleep stages and body temperature," *Sleep*, vol. 22, no. 6, pp. 767–773, 1999.

[15] W. Schmidt-Kessen and K. Kendel, "The influence of room temperature on night sleep in man (polygraphic night-sleep recordings in the climate chamber)," *Research in Experimental Medicine*, vol. 160, no. 3, pp. 220–233, 1973.

[16] CEC, *Residential Alternative Calculation Method (ACM) Approval Manual for the 2005 Building Energy Efficiency Standards for Residential and Nonresidential Buildings*, California Energy Commission, Sacramento, CA, 2005.

[17] D. Auslander, "Feedback control characteristics of residential, single-stage thermostats," Dec 2006.

[18] ASHRAE, *ASHRAE Green Guide: The Design, Construction, and Operation of Sustainable Buildings, 2nd ed.* Burington, MA: Butterworth-Heinemann, Nov 2006.

[19] J. Braun, "Reducing energy costs and peak electrical demand through optimal control of building thermal storage," *ASHRAE Transactions*, vol. 96, no. 2, pp. 876–887, 1990.

[20] M. Al-Homoud, "The effectiveness of thermal insulation in different types of buildings in hot climates," *Journal of Thermal Envelope and Building Science*, vol. 27, no. 3, pp. 235–247, 2004.

[21] C. P. Underwood and F. W. Yik, *Modeling Methods for Energy in Buildings*. Malden, MA: Blackwell Publishing, Oct 2004.

[22] A. LaRue, "Distributed sensing and controlling of residential hvac systems for thermal comfort, demand response, and reduced annual energy consumption," Master Thesis, University of California at Berkeley, Berkeley, CA, 2006.

[23] USCensus. (2000) 2000 census data. http://factfinder.census.gov/servlet/SAFEHousing?_sse=on.

[24] F. Porges, *HVAC Engineer's Handbook, 10th ed.* Oxford, UK: Butterworth-Heinemann, Sep 1995.

[25] W. Marion and K. Urban. (1995, Jun) User's manual for tmy2s: Typical meteorological years. National Renewable Energy Laboratory. Golden, CO. http://www.energy.ca.gov/title24/2005standards/residential_acm.

[26] R. Bras, *Hydrology: An Introduction to Hydrologic Science*. Addison-Wesley, Jul 1989.

[27] W. Burke and D. Auslander, "Thermostat based load management simulation-open loop control," Sep 2007.

[28] ASHRAE, *2001 ASHRAE Handbook: Fundamentals (P-I Edition)*. Atlanta, GA: American Society of Heating, Refrigerating and Air Conditioning Engineers, Jul 2001.

[29] R. Howell, H. Sauer, and W. Coad, *Principles of Heating, Ventilating, and Air Conditioning: A Textbook with Design Data Based on 2005 ASHRAE Handbook - Fundamentals*. Atlanta, CA: American Society of Heating, Refrigerating and Air-Conditioning Engineers, Jun 2005.

[30] T. Kuehn, J. Ramsey, and J. Threlkeld, *Thermal Environmental Engineering, 3rd ed.* Upper Saddle River, NJ: Prentice-Hall, Feb 1998.

[31] P. Martin and D. Oughton, *Faber and Kell's Heating and Air-Conditioning of Buildings, 7th ed*, ser. 7. Butterworth-Heinemann, Jul 1989.

[32] F. McQuiston and J. Parker, *Heating, Ventilating, and Air Conditioning Analysis and Design, 4th ed.* New York, NY: John Wiley and Sons, Mar 1994.

[33] K. Moss, *Heat and Mass Transfer in Buildings, 2nd ed.* New York, NY: Taylor and Francis, Apr 2007.

[34] H. A. P. Blom and Y. Bar-Shalom, "The interacting multiple model algorithm for systems with markovian switching coefficient," *IEEE Transactions on Automatics Control*, vol. 33, no. 8, pp. 780–783, Aug 1988.