

Achieving Residential Demand Response through
Predictive Evaluation of Thermal Response (ETHER)
Models: An ARIMA Approach

A Dissertation

Presented to

The Faculty of the School of Engineering and Applied Science

University of Virginia

In Partial Fulfillment

Of the Requirements for the Degree

Doctor of Philosophy in Electrical Engineering

By

Shana Chantell Craft

May 2012

© Copyright May 2012

Shana Chantell Craft

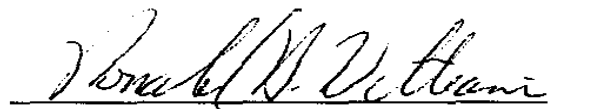
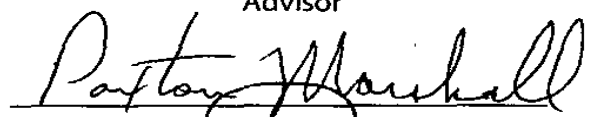


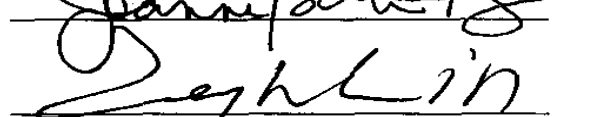

All rights reserved

APPROVAL SHEET

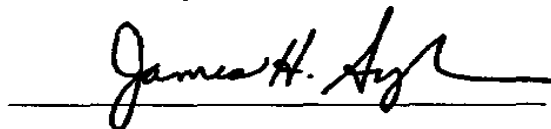
The dissertation
is submitted in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy


AUTHOR

The dissertation has been read and approved by the examining committee:


Advisor






Accepted for the School of Engineering and Applied Science:



Dean, School of Engineering and Applied Science

May
2012

Table of Contents

LIST OF FIGURES	VII
LIST OF TABLES	IX
ABSTRACT	X
ACKNOWLEDGEMENTS.....	XII
I. INTRODUCTION	1
A. MOTIVATION.....	1
B. WHY RESIDENTIAL?.....	4
C. PROBLEM STATEMENT	8
D. HYPOTHESIS.....	12
E. RESEARCH METHOD	12
F. RESIDENTIAL PROJECT.....	13
G. THEORETICAL OVERVIEW	14
H. DATA OVERVIEW	16
I. SIGNIFICANCE OF STUDY	17
J. DISSERTATION STRUCTURE	18
II. RESEARCH CONTEXT & RELATED WORK.....	22
A. EVOLUTION OF THE GRID.....	22
B. DEMAND RESPONSE MODELS	26
<i>B.1. Adjustment of Loads based on Hourly Electricity Prices</i>	<i>26</i>
<i>B.2. Electric Power Consumption based on Past Data.....</i>	<i>27</i>
<i>B.3. Energy Scheduling based on the Behavior of Consumers</i>	<i>29</i>
<i>B.4. Physical-based Residential Loads at the Appliance Level</i>	<i>29</i>
<i>B.5. Residential Energy Management based on an Optimal Stopping Method</i>	<i>31</i>
<i>B.6. Simulated Residential Energy Management System (RES).....</i>	<i>32</i>
C. INTELLIGENT DEMAND RESPONSE FRAMEWORK	32
<i>C.1. Intelligent Decision Support System (IDSS) Model</i>	<i>34</i>
<i>C.2. Intelligent Demand Response Module</i>	<i>35</i>
D. DEMAND RESPONSE PROGRAMS AND STRATEGIES	36
E. THERMAL MODELING DESIGN	38
F. OBSERVATIONAL DATA APPLICABILITY	39
III. INTELLIGENT DEMAND ENERGY ADAPTIVE SYSTEM (IDEAS)	42
A. PRE-PROGRAMMED DEMAND RESPONSE SCENARIO MODEL	45
B. EVALUATION OF THERMAL RESPONSE (ETHER) MODEL.....	48
C. FINANCIAL IMPACT MODEL	51
D. BALANCE CONSUMER PREFERENCES AND FINANCES MODEL	52
IV. DATA SOURCES.....	54
A. INTERIOR AND EXTERIOR TEMPERATURE DATA	55
B. POWER CONSUMPTION DATA.....	57
C. HUMIDITY DATA.....	58
D. CARBON DIOXIDE (CO ₂) DATA.....	60
E. WATER FLOW DATA	61

F.	SOLAR HOT WATER TEMPERATURE DATA	62
V.	VARIABLE DEPENDENCE.....	63
A.	AGGREGATE APPROACH	65
B.	NO SOLAR LOADING APPROACH.....	66
C.	SOLAR LOADING APPROACH.....	68
D.	INTERIOR HUMIDITY APPROACH.....	69
VI.	RESEARCH METHODOLOGY.....	73
A.	AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (ARIMA) METHOD	73
B.	ARIMA MODEL PARAMETERS	77
C.	ARIMA TRANSFER FUNCTION ETHER MODELS	79
C.1.	<i>Rationale for Day & Night Model Techniques</i>	80
C.1.1.	Daylight Savings Time Technique	80
C.1.2.	Equinox to Solstice Technique	81
C.1.3.	Seasonal Technique	84
C.1.4.	Entire Month Technique	85
C.2.	<i>Rationale for Outside Temperature Conditions</i>	87
D.	CASE STUDY - EVALUATION OF DAY & NIGHT ETHER MODEL TECHNIQUES	88
VII.	MODEL VALIDATION.....	92
A.	GOODNESS OF FIT MEASURES	93
B.	ASSUMPTIONS	99
VIII.	EVALUATION APPROACHES	101
A.	BASELINE APPROACH	102
A.1.	<i>Import Phase</i>	102
A.2.	<i>Model Development Phase</i>	102
A.3.	<i>Forecasting Phase</i>	104
A.4.	<i>Model Development & Forecasting Automation</i>	106
B.	REDUCED VARIABLE APPROACH	109
B.1.	<i>Data Source Correlation Method</i>	112
B.2.	<i>Variance Inflation Factor (VIF) Method</i>	113
B.2.1	Case Study: Modification of VIF Cutoff	114
B.3.	<i>Significance (p-value) Method</i>	115
C.	SELECT (REDUCED) VARIABLE APPROACH.....	116
D.	CASE STUDY – COMPARISON OF EVALUATION APPROACHES	116
E.	CASE STUDY – SELECT (REDUCED) VARIABLE APPROACH VS. TOP 5 VARIABLES.....	122
IX.	EVALUATION OF ADDITIONAL DATASETS	125
A.	PAVILION II.....	126
B.	BAYLY MUSEUM.....	127
X.	CONCLUSIONS.....	129
A.	CONTRIBUTIONS	132
B.	FUTURE WORK.....	133
C.	SUMMARY.....	134
XI.	BIBLIOGRAPHY	137
XII.	APPENDIX A - EMBEDDED ENVIRONMENT	148

A.	EVALUATION OF DATA	149
B.	MODEL IDENTIFICATION	150
C.	MODEL PARAMETER ESTIMATION	151
D.	MODEL ACCURACY.....	151
XIII.	APPENDIX B – DATA IMPORT SYNTAX CODE	153
XIV.	APPENDIX C – MODEL DEVELOPMENT SYNTAX CODE	155
XV.	APPENDIX D – FORECASTING SYNTAX CODE	155
XVI.	APPENDIX E – RE-ESTIMATION SYNTAX CODE	159
XVII.	APPENDIX F - VARIABLE REDUCTION SYNTAX CODE.....	159
XVIII.	APPENDIX G – EVALUATION TECHNIQUE PREDICTION RESULTS	163

List of Figures

FIGURE I-1: SIGNAL TRANSMISSION BETWEEN THE UTILITY AND THE CONSUMER.....	2
FIGURE I-2: GROWTH IN U.S. RESIDENCES FROM 1980 TO 2005 [15].....	6
FIGURE I-3: TOTAL PERCENTAGE OF RESIDENTIAL ENERGY CONSUMED IN 2005 [15]	7
FIGURE I-4: PROCESS FLOW BETWEEN THE UTILITY, CONSUMER AND ETHER MODELS.....	15
FIGURE I-5: FULLY AUTOMATED DEMAND RESPONSE WITH PREDICTIVE MODEL	16
FIGURE I-6: ETHER DEVELOPMENT ROADMAP.....	21
FIGURE II-1: UNIDIRECTIONAL FLOW IN THE TRADITIONAL UTILITY ENVIRONMENT [26].....	23
FIGURE II-2: BI-DIRECTIONAL FLOW ENABLED BY SMART GRID TECHNOLOGIES [26].....	24
FIGURE II-3: SIMPLIFIED THERMAL NETWORK MODEL	38
FIGURE III-1: INTELLIGENT DEMAND ENERGY ADAPTIVE SYSTEM (IDEAS)	43
FIGURE III-2: INTERNAL MODELS WITHIN THE IDEAS SYSTEM	45
FIGURE III-3: DECISION-BASED PROCESS FOR THE PRE-PROGRAMMED DEMAND RESPONSE SCENARIO MODEL	46
FIGURE III-4: MODEL SELECTION PROCESS FOR THE ETHER MODELS	49
FIGURE III-5: MODEL PROCESS TO EVALUATE THE FINANCIAL IMPACT	51
FIGURE III-6: DECISION-BASED PROCESS FOR THE BALANCING OF ECONOMIC AND COMFORT PREFERENCES	53
FIGURE IV-1: TOP 4 POWER CONSUMPTION LOAD APPLIANCES	58
FIGURE IV-2: CO ₂ READINGS DURING THE MORNING HOURS FOR THE MONTH OF FEBRUARY	61
FIGURE V-1: CONCEPTUAL DESIGN OF THE VARIABLE DEPENDENCE APPROACHES	64
FIGURE V-2: AGGREGATE APPROACH – EVALUATION RESULTS OF THE DAY MODELS	66
FIGURE V-3: NO SOLAR LOADING APPROACH – EVALUATION RESULTS OF THE DAY MODELS.....	68
FIGURE V-4: SOLAR LOADING APPROACH – EVALUATION RESULTS OF THE DAY MODELS.....	69
FIGURE V-5: INTERIOR HUMIDITY APPROACH - EVALUATION RESULTS OF THE DAY MODELS	70
FIGURE V-6: COMPARISON BETWEEN THE INTERIOR HUMIDITY APPROACH VS. THE NO SOLAR LOADING APPROACH	71
FIGURE V-7: COMPARISON BETWEEN ALL FOUR VARIABLE DEPENDENCE APPROACHES	71
FIGURE VI-1: CONCEPTUAL ORIENTATION OF THE DAYLIGHT SAVINGS TIME TECHNIQUE	80
FIGURE VI-2: CONCEPTUAL ORIENTATION OF THE EQUINOX TO SOLSTICE ETHER MODEL TECHNIQUE.....	83
FIGURE VI-3: CONCEPTUAL ORIENTATION OF THE SEASONAL ETHER MODEL TECHNIQUE.....	85
FIGURE VI-4: CONCEPTUAL ORIENTATION OF THE ENTIRE MONTH ETHER TECHNIQUE.....	86
FIGURE VI-5: EVALUATION OF THE RESEARCH TECHNIQUES FOR THE DAY ETHER MODELS	89
FIGURE VI-6: EVALUATION OF THE RESEARCH TECHNIQUES FOR THE NIGHT ETHER MODEL	90
FIGURE VII-1: SUCCESS RATIO RESULTS WITHIN 1 DEGREE F FOR DAY & NIGHT ETHER MODELS	94
FIGURE VII-2: SUCCESS RATIO RESULTS WITHIN 2 DEGREE F FOR DAY & NIGHT ETHER MODELS	95
FIGURE VII-3: SUCCESS RATIO RESULTS WITHIN 3 DEGREE F FOR DAY & NIGHT ETHER MODELS	96
FIGURE VII-4: MEASURE OF EFFECTIVENESS RESULTS FOR THE DAY & NIGHT ETHER MODELS	97
FIGURE VIII-1: CSV FILE FORMAT INTO SPSS.....	102
FIGURE VIII-2: TIME SERIES MODELER WITH DEPENDENT AND INDEPENDENT VARIABLES	103
FIGURE VIII-3: APPLY TIME SERIES MODELS	105
FIGURE VIII-4: DATA IMPORT PROCEDURE INTO SPSS	107
FIGURE VIII-5: MODEL DEVELOPMENT STEPS IN SPSS	107
FIGURE VIII-6: ETHER MODEL FORECASTING PROCESS	108
FIGURE VIII-7: VARIABLE REDUCTION FLOWCHART METHODOLOGY.....	111
FIGURE VIII-8: VIF CUTOFF EVALUATION RESULTS	115
FIGURE VIII-9: MODEL REDUCTION APPROACHES – SUCCESS RATIO FOR DAY MODELS.....	118

FIGURE VIII-10: MODEL REDUCTION APPROACHES - MEASURE OF EFFECTIVENESS FOR DAY MODELS	119
FIGURE VIII-11: MODEL REDUCTION APPROACHES – SUCCESS RATIO FOR NIGHT MODELS.....	120
FIGURE VIII-12: MODEL REDUCTION APPROACHES - MEASURE OF EFFECTIVENESS FOR NIGHT MODELS ..	121
FIGURE VIII-13: SELECT (REDUCED) VARIABLE APPROACH VS. TOP 5 VARIABLES – DAY MODELS.....	123
FIGURE VIII-14: SELECT (REDUCED) VARIABLE APPROACH VS. TOP 5 VARIABLES – NIGHT MODELS.....	123
FIGURE IX-1: PREDICTION ACCURACY RESULTS FOR PAVILION II USING ETHER MODELS	127
FIGURE IX-2: PREDICTION ACCURACY RESULTS FOR THE BAYLY MUSEUM USING ETHER MODELS	128
FIGURE XII-1: FLOWCHART OF ESTABLISHING THE ARIMA MODEL.....	149

List of Tables

TABLE IV-1: DATA COLLECTION CSV FILE ORIENTATION	54
TABLE IV-2: INTERIOR TEMPERATURE SENSOR IDS AND SENSOR DESCRIPTION	56
TABLE IV-3: EXTERIOR TEMPERATURE SENSOR IDS AND SENSOR DESCRIPTION	56
TABLE IV-4: POWER CONSUMPTION SENSOR IDS AND SENSOR DESCRIPTION.....	57
TABLE IV-5: HUMIDITY SENSOR IDS AND SENSOR DESCRIPTION	59
TABLE IV-6: CORRELATION VALUES FOR THE HUMIDITY SENSORS	59
TABLE IV-7: CO ₂ SENSOR IDS AND SENSOR DESCRIPTION	60
TABLE IV-8: WATER FLOW SENSOR IDS AND SENSOR DESCRIPTION	61
TABLE IV-9: SOLAR HOT WATER TEMPERATURE SENSOR IDS AND SENSOR DESCRIPTION.....	62
TABLE V-1: AGGREGATE APPROACH – ALL FOURTEEN INTERIOR TEMPERATURE SENSORS	65
TABLE V-2: NO SOLAR LOADING APPROACH – EXCLUDING SENSOR T11	67
TABLE VI-1: ARIMA MODEL PARAMETERS AND PARAMETER ESTIMATES [35]	78
TABLE VI-2: EIGHT TEST CASES DESIGNED TO EVALUATE MODEL PREDICTION ACCURACY	88
TABLE VII-1: EXAMPLE OF FIT STATISTICS METRICS FOR THE TRAINED NIGHT (COLD) ETHER MODEL.....	92
TABLE VII-2: PREDICTION ACCURACY RESULTS FOR DAY & NIGHT ETHER MODELS.....	99
TABLE VIII-1: BIVARIATE CORRELATION RESULTS WITH THE CO ₂ VARIABLES	112
TABLE VIII-2: MODEL REDUCTION APPROACHES – GOODNESS OF FIT MEASURES FOR DAY MODELS	117
TABLE VIII-3: MODEL REDUCTION APPROACHES – GOODNESS OF FIT MEASURES FOR NIGHT MODELS	119

Abstract

For the average electricity consumer, electricity is viewed as an invisible commodity. The consumer is not aware of their energy usage or how fast the dollars are adding up until the electricity bill arrives a month later. In the traditional utility environment, there are no mechanisms in place to alter the consumers' consumption patterns. However, with the adoption of the smart grid, consumers are given the means to track and adjust their energy usage to achieve demand response.

Demand response has been used frequently within the commercial and industrial sectors and given the emerging smart grid paradigm, it is now being expanded into the residential sector. As such, it not only promotes cost savings for the consumer, but also reduces the need for utility company generation capacity to supply peak time energy usage. Within this dissertation, the residential sector is investigated to identify gaps that need to be addressed to achieve demand response. As an integral component to help the consumer achieve demand response, a set of models are developed to predict the interior conditions within a building and provide information to the consumer identifying which loads impact the interior conditions to assist in planning of energy consumption. In particular, a set of predictive thermal response models can be embedded in an intelligent system framework which can then proffer decisions that lead to economic and environmental advantages to the consumer.

Our methodology is based on developing Autoregressive Integrated Moving Average (ARIMA) time series thermal response models from past observational data collected from a residence (i.e., smart home). A year of collected data is used for model development and analysis. Due to the nature of the time series and the objective of this analysis, 70% of the observations are used to train the models; the remaining 30% is used to test the models for forecasting accuracy.

The research shows that the ARIMA time series method can be used to obtain short-term (e.g., up to 6 hours) thermal response forecasts to not only provide consumers with insight on their comfort levels but also assist in identifying the contributing loads (e.g., heating, ventilation, and air-conditioning (HVAC), water heater) that consume energy and influence interior conditions. The advantage of this research is the ability to forecast interior conditions and identify contributing loads based on activity recognition conducted in the residence during varying outside temperature and time of day conditions. Validation of the forecasting models has been carried out by comparing the models' output with the actual data collected from the interior temperature conditions.

Acknowledgements

First and foremost, I would like to thank God for this amazing opportunity. As it is written, “For everyone to whom much is given, of him shall much be required.” – Luke 12:48. I look forward to giving back to my community to encourage our future leaders in the field of science, technology, engineering and mathematics.

I would like to thank my employer, Lockheed Martin (LM) Corporation, for selecting me to be a pilot participant of the Corporate Engineering & Technology (CE&T) PhD program that was spearheaded by Dr. Ray Johnson, LM Chief Technology Officer (CTO). I am truly grateful for this opportunity and for the support and assistance that I have received from the corporation. I look forward to expanding the breadth of knowledge obtained in the pursuit of my PhD to foster relationships, add value and win business for the corporation.

I would like to thank my church family, New Life Community Baptist Church in Gainesville, VA, for their unwavering support and prayers. As it is written, “...The effectual fervent prayer of a righteous man availeth much.” – James 5:16

I would like to thank my PhD Advisory Committee for mentoring and advising me throughout this progress. To Dr. Williams, I thank you for your willingness to advise me over these 3 years. At first, it was a little complicated getting my bearings, being back in an academic setting; however, you were patient with me and I thank you. To Joanne, I appreciate your candor and your willingness to meet with me on a weekly basis to ensure my progress and assist me with my milestones for completion. To Tariq, I thank you for taking time out of your schedule to meet with me at the Starbucks in Tyson’s Corner whenever I needed help or just to discuss matters of uncertainty. To Professor Lin, I thank you for your positivity and your willingness to listen. To John Quale, I thank you for your insight on the big picture of my research. Our dialogue helped me to focus and eventually frame the solution that has been developed. To Paxton, I thank you for

being a mentor, an advocate, and a genuine person. Our initial interaction dates back to 2002 and you have always been a champion of mine and for that I would like to say “Thank you.”

To my family, words cannot express how appreciative I am of your loving support. Being raised in Hurt, VA and looking back on my life, I can truly say that you all have been a rock in the midst of all my tears and my complaints. I thank you for your unconditional love and support.

To my friends, mentors and UVA staff members, I would like to thank you all for putting up with me through this process. This road was not easy to travel; however, you all continued to feed me with positive thoughts and uplifting words.

In closing, “Trust in the LORD with all thine heart; and lean not unto thine own understanding. In all thy ways acknowledge him, and he shall direct thy paths.” – Proverbs 3:5-6

I. Introduction

A. Motivation

The demand for energy is steadily rising; by the year 2030, a 44% increase in global demand for energy is expected [1,2]. In an effort to meet the energy demands, utility companies and consumers (end users) are partnering and investigating renewable energy sources. Sianaki et al states that as the dependence on resources to meet the energy demand shifts from non-renewable to renewable sources, the cost of upgrading the conventional electric power system, pricing and service networks is impractical. This is due, in large part, to the fact that these systems have an inadequate central capacity plan to meet the growing demand and energy shift [2].

With the heightened need to meet energy demands, the capabilities and benefits of the Smart Grid and demand response are becoming a significant research and development field of study. The Smart Grid purports to increase the connectivity, automation and coordination between suppliers, utilities and consumers. In essence, the Smart Grid encompasses a life cycle of processes from energy generation through transmission and distribution, through to consumption. The main objectives of the Smart Grid as highlighted in [3,4,5] are focused on: 1) implementing the integration of Information and Communication Technologies (ICT) into the conventional electricity grid, 2) enabling two-way transmission of information and electricity between the respective parties, 3) being able to heal itself (i.e., self-healing), 4) resisting physical and cyber attacks, 5) establishing and employing efficient methods of energy generation and energy storage and 6) providing efficient energy management processes and protocols for commercial, industrial and residential consumers.

One of the goals of the Smart Grid is to influence the demand response of the end user. Demand response can be defined as the reduction of peak demand in which end users can modify their electricity consumption patterns in response to dynamic pricing or request for demand reductions

signals provided by the utility. In essence, demand response provides the opportunity for a cost effective alternative to adding generation facilities to accommodate peak time energy usage.

In the demand response system infrastructure, two parties (i.e., the utility and consumer) are both allowed and encouraged to collaborate. The utility sends signals to the consumer to encourage a reduction in demand by a certain amount or to notify the consumer of dynamic pricing. The consumer, upon receiving the signal decides on the appropriate response to reduce demand and maximize savings while maintaining an acceptable level of comfort and convenience. Figure I-1 illustrates the interaction between the utility and the consumer.

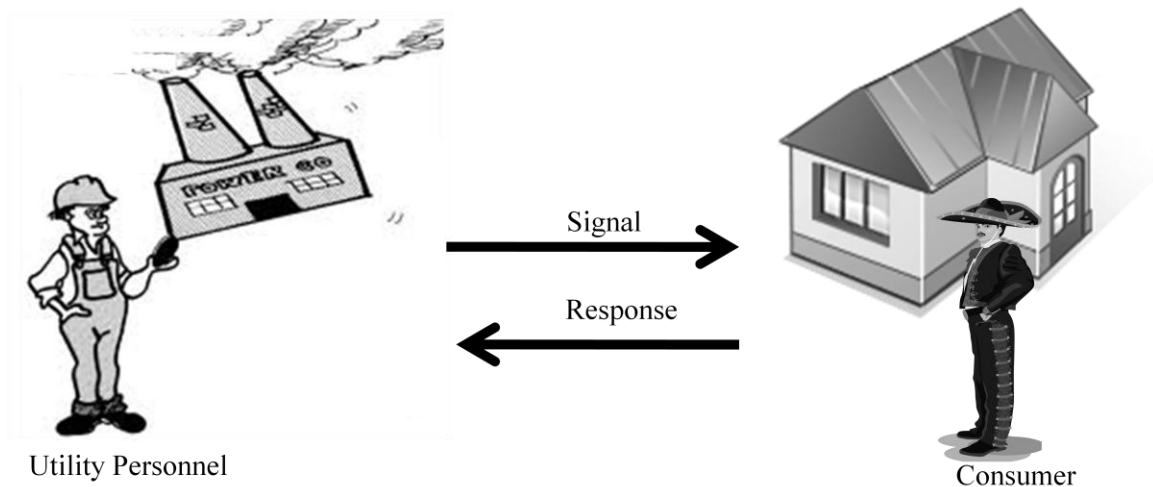


Figure I-1: Signal transmission between the Utility and the Consumer

The challenge with this scheme is that the consumer is now responsible for finding the potential loads that can be shut down or re-scheduled in an effort to reduce peak electricity demand and obtain significant savings [6]. Approaches have been proposed in literature to achieve demand response at different levels of the Smart Grid but few approaches have been developed that focus on achieving demand response on a continuous basis for the residential consumer. In a residential setting, automating the management of energy usage will obviate the need for consumers to make decisions regarding load management. Studies have shown that consumers have neither the time

nor the knowledge to make reasonable decisions regarding price variation or demand for reduction signals to be used to their advantage [7].

Intelligent demand response models and schemes have been studied in the literature to facilitate automated demand-side management of energy usage. Several different approaches focus on developing a decision-based module to make energy reduction decisions on behalf of the consumer. Sianaki et al. suggested an intelligent decision-making support system at the residential level working on behalf of the consumers by learning the consumer's preferences, based on their consumption profile, to achieve efficient utilization of energy [2]. Dam et al. suggested automated demand response from the consumer side, however, the primary target of their work was large industrial sized consumers [6]. Iwayemi et al. developed an automated residential energy management and scheduling platform to minimize peak demand without sacrificing comfort levels of the consumer [7].

Balancing energy reduction requests to achieve demand response while maintaining thermal comfort levels for the consumer is still an area of exploration. For the residential consumer, current models and schemes are insufficient to support demand response and present day schemes and models are focused more on the utility than the consumer [12]. Studies have shown that many of the demand response programs that are designed to reduce energy consumption are inconvenient for the consumer (i.e., wait times for re-scheduling, etc.) [7].

The most commonly known programs/scenarios that are available for consumers to reduce energy consumption are Direct Load Control (DLC)/Load Shut Down scenario and Price-Based Control (PBC)/Load Reduction scenario. In the DLC program, the utility sends a load shut down signal to the consumer, requesting certain loads to be shut down; this is mainly sizeable loads (i.e., HVAC). Depending on the participation agreement of the demand response program, the consumer can accept or decline the request [6]. In the PBC program, the consumer will choose

which appliance/load settings to change by a certain amount [9]. The load reduction scenario is more complicated since the consumer is now responsible for finding the loads to shut down or reduce by a certain amount [6]. Therefore, an automated agent/model that can be used as part of an intelligent demand response framework is needed to provide the residential consumer predictions on interior thermal response conditions based on the influence of contributing loads (i.e., heating, ventilation, and air-conditioning (HVAC), exterior temperature, etc.) to enable intelligent decision making in achieving demand response.

The investigation of thermal response has primarily been explored by using a thermal network design model [10], [11], [12]. However, with the deterioration of the building structure over time, the modeled system diverges from the real system operation. In addition, the predictive thermal response resistor-capacitor network (i.e., RC network) model of a building structure is difficult and quite costly to build. The set of ETHER models developed in this study alleviates the limitation of divergence by having the capability to automatically and continually update based on the building's operational behavior. That is, ETHER models are solely dependent upon the operational performance (i.e., extraction of observations from a building in operation) versus the structural framework and orientation of the building. This capability enhances the accuracy of prediction and the ability to maintain thermal comfort levels.

B. Why Residential?

In today's building industry, designers, developers and owners are searching for ways to minimize operating costs, energy usage and environmental impacts of buildings. With a push to optimize and conserve resources, residential buildings are amongst the sectors (e.g., commercial and industrial) that are under investigation. The United States household electricity report states that residential consumption of electricity has grown dramatically over the past several decades, while the usage of natural gas and fuel oil has declined [14]. Based on data published by the Energy Information Administration (EIA), it was found that the electricity consumption in 107

million U.S. residences in 2001 totaled 1,140 billion kWh, where the most significant uses within the residences included heating, ventilation and air conditioning (i.e., space heating and space cooling) and refrigerators [14]. It has been observed that hot summers and cold winters in addition to humidity impact the amount of energy consumption within a residence.

The service demands of occupants for residential buildings include lighting, water heating, heating, ventilation, and air conditioning for the winter and summer, refrigerators, electronic equipment (i.e., computers, televisions, etc.), laundry appliances (i.e., washer and dryer) and cooking appliances (i.e., range, stove, microwave, etc.), which all require significant energy usage. In 2001 as reported in [14],

- ❖ Lighting accounted for 101 billion kWh, which was 8.8 percent of the electricity consumed by U.S. residences
- ❖ Electric water heating accounted for over 100 billion kWh, which was 9 percent of the electricity consumed by U.S. residences
- ❖ Heating, ventilation and air conditioning (HVAC) accounted for 356 billion kWh, which was 31 percent of the electricity consumed by U.S. residences
- ❖ Electric space heating (i.e., portable electric space heaters) accounted for 116 billion kWh, which was 10 percent of the electricity consumed by U.S. residences
- ❖ Refrigerators accounted for 156 billion kWh of U.S. residences' electricity use
- ❖ Electronic equipment accounted for 82 billion kWh of U.S. residences' electricity use
- ❖ Laundry appliances – combined washer and dryer usage, accounted for approximately 76 billion kWh, which was 6.7 percent of the electricity consumed by U.S. residences
- ❖ Cooking appliances accounted for less than 4 percent of the electricity consumed by U.S. residences.

The number of U.S. residences has increased significantly over time due to the growth in population and economic activity; from 1980 to 2005, there was nearly a 40 percent increase. Figure I-2 shows the growth in U.S. residences from 1980 to 2005. There was a direct relationship between the growth in population and the growth in the number of residences.

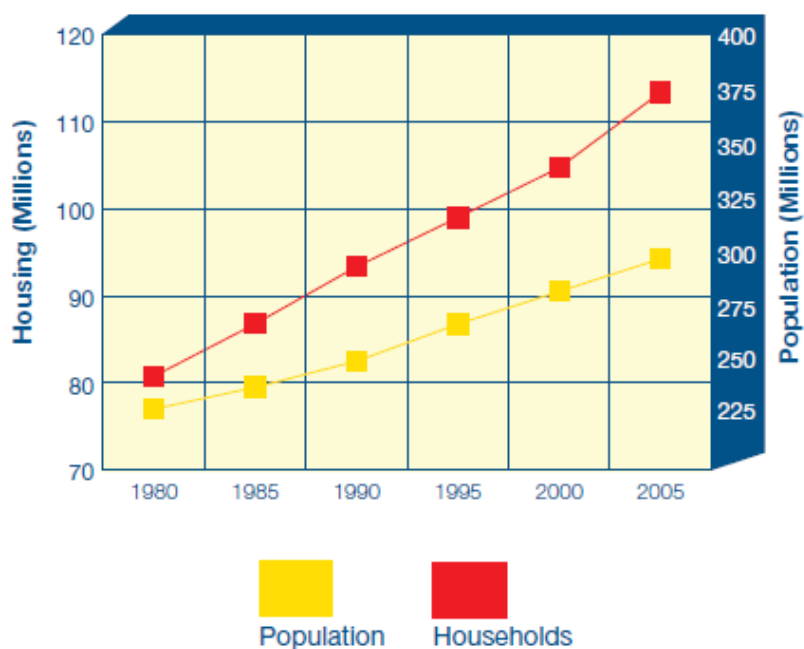


Figure I-2: Growth in U.S. Residences from 1980 to 2005 [15]

In addition to the increase of residences due to the population growth, the energy usage also increased in the residential sector. The growth in peak demand was due to bigger residences and additional loads such as more electronics (i.e., bigger televisions and computers) that exceeded the transmission growth rate by approximately 25 percent each year [16]. In Figure I-3, the pie chart illustrates the percentage of energy usage within the total residential sector for 2005. As observed, there is a significant increase in the percentage of the electricity consumed by U.S. residences compared with the reported data for 2001. However, in 2001 and 2005, space heating and space cooling is the largest usage in the residence.

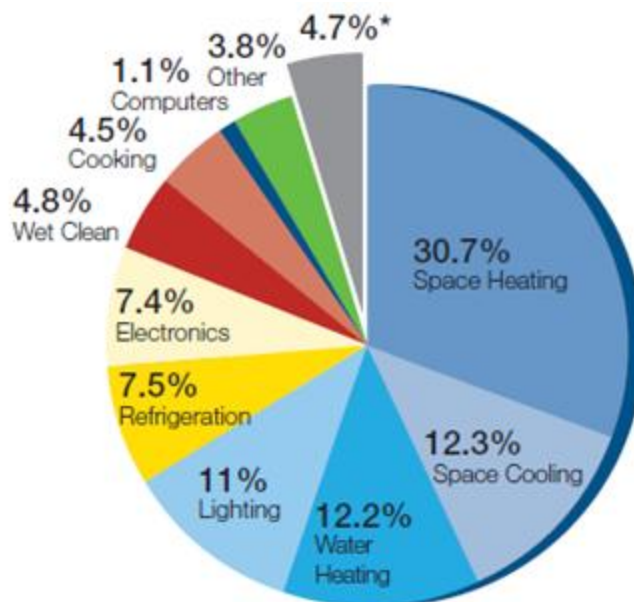


Figure I-3: Total Percentage of Residential Energy Consumed in 2005 [15]

In recent years, there has been a decrease in the growth of the residences due to the U.S. economic activity. Despite the recent decline, the residential sector has still driven an increase in electricity consumption. As a direct result of the increase in electricity consumption, energy costs and electricity bills also increase. In an effort to achieve energy efficient trends and maximize cost savings in residential buildings, designers, developers and owners are now investing in ENERGY STAR appliances, energy efficient lighting and investigating energy efficiency programs. It was reported in [16] that if every U.S. residence replaced one incandescent light bulb with a fluorescent light bulb that the country would benefit by conserving enough energy to light 3 million residences and save approximately \$600 million annually.

Along with the energy efficient techniques that are being investigated, demand response programs are also being sought to provide a platform of automation to act on behalf of the consumer and achieve cost savings. Demand response programs are voluntary schemes that are offered by the utility to their customers to curtail energy consumption during peak demand periods of time. These programs have been more frequently used within the commercial and

industrial sectors as opposed to the residential sector; this is primarily due to the fact that residential consumers are usually charged a flat or two-tiered (i.e., peak and off peak) electricity rates [7]. However, with the evolution of the Smart Grid, demand response can now be expanded into residential sectors on a larger scale [7].

The focus of this research is on the management of residential energy consumption. The Smart Grid allows the utility to send alerts to the consumer when the prices have or will increase (usually during the peak hours of the day), when there is a shortage of electricity or when there needs to be a decrease in consumption. Through the use of enabling technology such as the smart meter, it prompts users to turn off certain non-critical equipment (e.g. air conditioning unit) in their homes in order to minimize energy usage. This new system demands a framework in which people, systems and solutions are dynamic and flexible in responding to technological advances, consumer needs, prices, standards, policies and other requirements [17].

C. Problem Statement

The typical electric bill is delivered to consumers' days after the consumption actually occurs and consumers usually wonder, "Why is my electricity bill so expensive?" Consumers are literally given no visibility into the type of decisions they should make regarding energy consumption. However, the Smart Grid provides a platform to encourage consumer participation in controlling their energy consumption based on notifications of price variations of energy rates and energy reduction requests from utility providers.

As part of the smart grid paradigm, there are still technical obstacles to overcome to make the consumers' decision making and the smart grid's event response more feasible for management and control purposes. Automating consumer demand response to support the smart grid paradigm requires insight to enable intelligent decisions that consider information provided by utilities (i.e., price variation or demand reduction requests) along with consumer needs. Therefore, the central

problem we address in this research is the development of a methodology and a set of forecasting ETHER models to achieve demand response in the residential sector while also satisfying consumer needs. The automated methodology should provide a basic framework of situational awareness of the thermal response based on past performance data and add intelligence to reduce dependence on the intervention of the consumer. The methodology should also allow for the capability of predicting the interior conditions over a short duration of time (i.e., up to 6 hours) in order to manage energy consumption.

An additional challenge to be addressed focuses on the number of sensors/variables used for model development. The data used within this research is collected with sensors located within the inside and outside perimeters of a residence. A large number of sensors results in a large amount of data to be processed for model development which becomes time consuming and computationally extensive. As a result of the magnitude of data and the numerous controllable and uncontrollable variables in the system, it is uncertain how best to use this data to develop an accurate model to represent and predict the conditions within the space. If this challenge is ignored 1) over-monitoring due to an abundance of sensors will continue, 2) energy usage will continue to increase from lack of insight and 3) demand response may not be achieved. The methodology should reduce the number of sensors used for model development and provide insight into the exact appliances that impact the interior temperature.

The automation process of this methodology should bridge the existing gap between the most commonly known programs/scenarios that are available for consumers to reduce energy consumption: Direct Load Control (DLC) and Price-Based Control (PBC), and provide a sound decision making tool for reduction. In the DLC program, the utility has the ability to switch off and control appliances at times of peak demand. In the PBC program, consumers receive advanced notice of varying prices so that they can choose which appliance settings to change [9].

Our methodology should serve as a catalyst to reduce the dependence on the utility's interference and consumer's decision making to modify their energy consumption.

This body of work focuses on the achievement of residential demand response through the predictive evaluation of thermal response (ETHER) models. The ETHER models are constructed to serve as automated agents to provide additional insight to an intelligent demand response framework to enhance the consumer's decision making in an effort to achieve demand response. The set of ETHER models will focus exclusively on understanding past data for controllable and uncontrollable input variables that influence interior conditions. Automating the model development and forecasting process will lessen the need for human intervention, reduce energy consumption of appliances during demand periods, and maintain occupant comfort levels.

There are four main goals to be achieved in this research:

- ❖ **Goal 1: Model Development Methodology and Approach** - Develop a set of multivariate time series Autoregressive Integrated Moving Average (ARIMA) ETHER models based on the collected data from the residence. The parameterization of the predictive ETHER models will be based upon both controlled (i.e., HVAC activity) and exogenous variables (i.e., exterior temperature). The ETHER models will be used to predict the response of the space up to 6 hours in the advance in order to reduce energy consumption based on pricing and/or energy minimization signals. With this forecasting mechanism, reducing energy consumption should prove to be beneficial to the consumer.
- ❖ **Goal 2: Variable Reduction Procedure** - Design a test procedure that will check for highly correlated variables to reduce the excessive use of sensors within the building structure. The reduction of the number of variables can lessen the computation time for model development as well as reduce the occurrence of over-monitoring by decreasing the amount of data collected. This goal will prove to be beneficial if some sensors can be

eliminated from all models. For further analysis, an investigation of the top five variables that impact the interior conditions out of the current number of variables (i.e., 31 independent variables) will also be explored.

- ❖ **Goal 3: Automation of Model Development** - Automate the modeling process to produce self-constructing and self-parameterizing models to predict interior conditions within the building structure. Given a significant amount of data, ETHER models can be generated using the multivariate time series analysis and parameter values can be continuously refined with each forecast as actual observational data becomes available. Within the refinement process, the modeling structure and statistically significant independent variables (inputs/predictors) will be updated based on their influence of the dependent variable series (i.e., interior temperature).
- ❖ **Goal 4: Identification of Contributing Loads** – Identify the contributing loads of the modeled system so that appropriate modifications can be made to achieve the optimal settings for the controlled independent variables. Each ETHER model will be made up of both controllable and uncontrollable loads that greatly influence the interior temperature conditions. Each modeled system can be investigated by running several experiments based on different conditions (i.e., time of day, environmental conditions, seasonal influence, etc.) to predict interior conditions within the residence to test the optimal setting for the controlled independent variables.

To validate the adaptability and practicality of our applied methods, we tested our methodology on additional datasets from the baseline residence to include a public service building and a faculty residence building, both located on the campus grounds of the University of Virginia. Throughout the dissertation, case studies have been designed to evaluate the performance of our methodology comparing the developed ETHER models' output with the actual data collected for the interior temperature conditions.

D. Hypothesis

In hypothesis testing experiments, the researcher creates a scenario which involves one or more factors of interest (e.g., dependent variables) that can be observed systematically under varying conditions. The factors of interest are then thought to be impacted or influenced by factors that cause change (e.g., independent variables). The underlying premise behind hypothesis testing is to discover limiting conditions and refine understanding of the domain of interest [18].

The hypothesis in this research focuses on using data collected from a specially instrumented residence to develop a set of forecasting ETHER models to identify relationships or patterns of usage between independent variables that impact the interior conditions. Based on outside temperature and time of day conditions, each developed ETHER model will be able to predict interior conditions over a short duration of time (i.e., up to 6 hours in advance) to provide insight and information to the consumer on how to plan for energy consumption. This type of experiment will provide justification and proof that the proposed approach is valid and establishes a degree of credibility of its applicability to assist in achieving demand response for the residential consumer [19].

E. Research Method

Time series analysis is a promising research method to model the operational behavior and interior thermal response conditions of a residence in an effort to maintain thermal comfort and identify loads for reduction in response to utility signals (i.e., reduction in demand or dynamic pricing). The work described here applies a forecasting time series Autoregressive Integrated Moving Average (ARIMA) method to residential data to predict the interior thermal response conditions and identify the contributing loads for energy reduction.

The overall sequence of the ARIMA methodology used for the prediction of the interior thermal response conditions is organized into the following steps: 1) identify a hypothesis to be tested in

order to develop a model, 2) develop a model using observed data, 3) estimate model parameters, 4) check validation of the hypothesis and 5) forecast using the developed model [16]. If the hypothesis is found to be untrue, the model should be refined. Using the ARIMA method, predictions can be categorized into three categories (i.e., short term (minutes to weeks), medium term (a month up to several years), and long term (a year up to twenty years)) [17]. For purposes of the work presented in this dissertation, the methodology requires reasonable predictions of interior thermal response conditions over a short term forecast horizon (i.e., up to 6 hours) to support the development and assessment of control strategies in response to utility signals.

A set of eight Evaluation of Thermal Response (ETHER) forecasting models have been constructed based on outside temperature and time of day conditions to predict interior conditions up to 6 hours in advance. The ETHER models were built using historical data collected over a period of three years from an occupied smart home built by the School of Architecture and the School of Engineering students at the University of Virginia.

F. Residential Project

The data used in this research was collected from the ecoMOD 1 (smart home) located in Charlottesville, VA [22]. The ecoMOD 1 project is the first of four prototype houses that were built by students from the School of Architecture and the School of Engineering under the supervision of Professor Paxton Marshall and John Quale, ecoMOD Project Director.

The ecoMOD 1 home consists of two floors and a basement which includes 3 bedrooms, 2 bathrooms with a combination living/dining and kitchen area. The finished area of the home (excluding the basement) is approximately 1,244 square feet. The housing project was designed to achieve energy efficient protocols in affordable residential housing. Some of the features that attribute to the energy efficient design are high R-value wall construction (i.e., structurally

insulated panel system (SIPS), energy efficient or energy saving technologies and appliances (i.e., Energy Star appliances) and particular window placement to take advantage of solar loading [23].

Sensors were strategically distributed throughout the home to gather observational data of the building's performance. The sensors used were prefabricated and prepackaged to communicate their measurements to the National Instrument's FieldPoint system. The sensors are viewed as the lower level interface that has the ability to monitor and capture changes in the environment [24]. The collected observational data includes water flow rates, power consumption, humidity, carbon dioxide measurements; interior, solar hot water and exterior temperatures, which are housed in an Energy Monitoring System database on a FTP server [22].

G. Theoretical Overview

Figure I-4 shows an illustrative view for understanding the interaction and process flow between the utility, the consumer and the forecasting ETHER models. Pricing and demand reduction signals are sent from the utility to the consumer. Consumers who participate in demand response programs either shift their loads to off-peak hours or reduce their loads during peak hours. In turn, the utility provides a financial incentive to consumers to alter their consumption pattern so that they can meet their generating capacity demands [24].

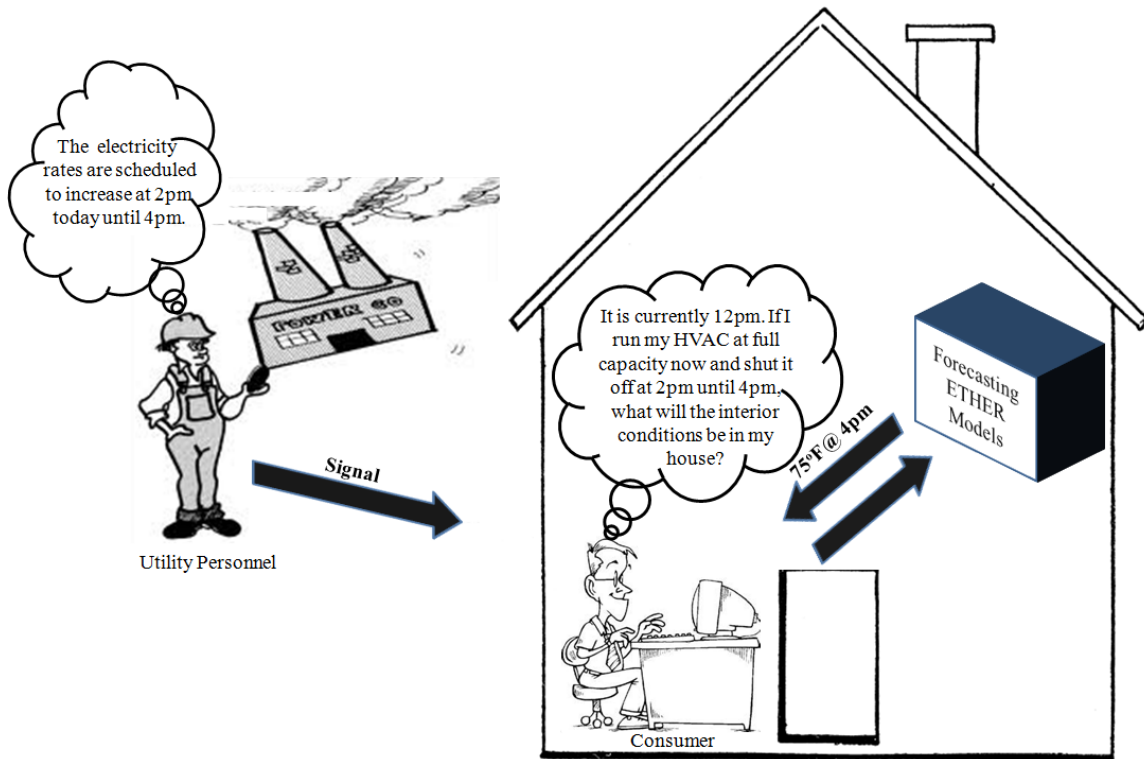


Figure I-4: Process flow between the Utility, Consumer and ETHER models

The demand response interface between the utility and consumer in most cases is still semi-automated, meaning that there is human intervention in deciding which loads to modify based on the signal. Ultimately, the goal is to fully automate the interaction between the utility and the consumer, where a fully automated demand response system involves no human intervention (see Figure I-5). With utilization of smart meters, the Smart Grid uses two-way communication between the consumer and the utility in order to manage and monitor the cost, delivery and demand of energy usage. Upon the smart meter receiving the receipt of the external signal, information (i.e., load profiles, consumer's preferences, time of day, exterior conditions, etc.) is supplied to the forecasting ETHER models to predict the building's thermal response. The ETHER models will serve as automated agents to assist in reducing energy consumption of non-critical equipment.

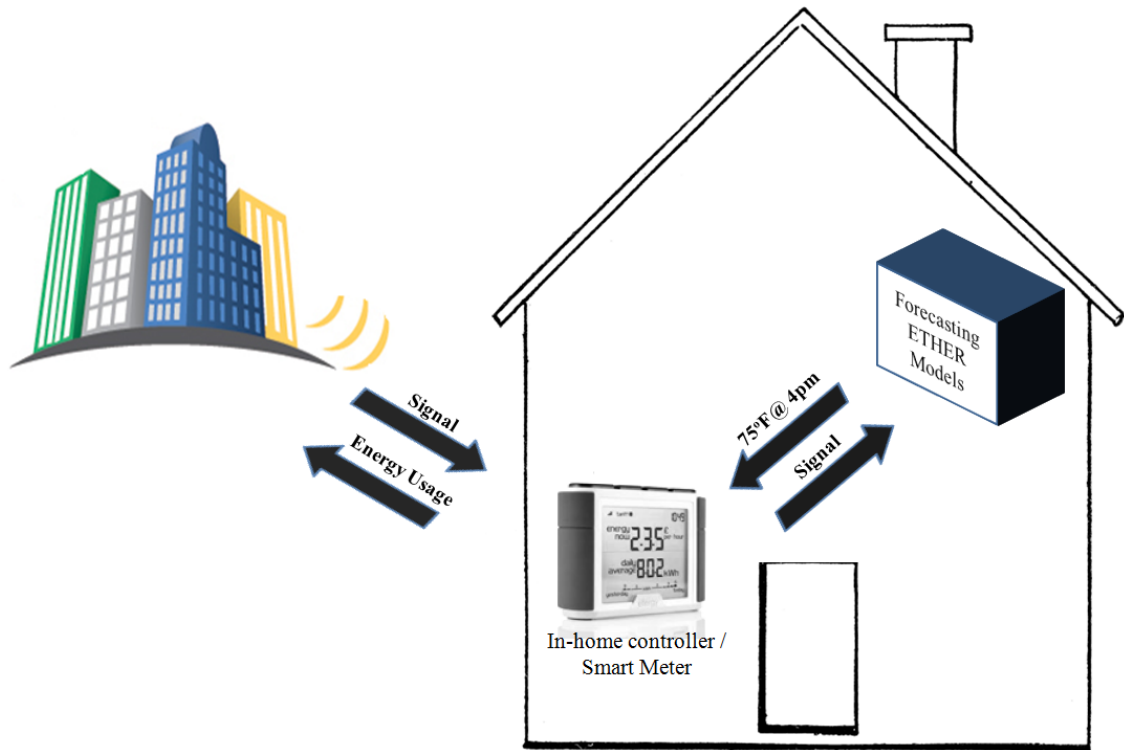


Figure I-5: Fully automated demand response with predictive model

Our goal is to concentrate on the automation of the predictive ETHER model process, where local input sensors collect data (i.e., power profiles, temperature, etc.) to be used to create a model for forecasting interior conditions up to 6 hours in advance. In an automated demand response framework, decisions would be made automatically based on consumer preferences and other factors concerning the shutdown of certain loads. The ETHER models and methodology should reduce the dependence on the utility's interference, obviate the continuous human intervention and enhance decision-making opportunities to achieve demand response in the future based on dynamic pricing and demand reduction signals.

H. Data Overview

Over three year's worth of data (i.e. interior temperature, exterior temperature, power consumption (lighting, water heater, cooking appliances), humidity, CO₂, water flow rates and solar hot water temperature, etc.) was collected from the residence; however, only one year's

worth of data is being used due to the excessive number of missing values or values outside of the protocol range for the specified sensors. The data used were sampled on a minute-to-minute basis to primarily observe the cycling activity of certain load appliances (e.g., HVAC).

There are a total of 45 parameters used to build the multivariate time series ETHER models, where 31 of the parameters will be used as independent variables. The response or dependent variable used in the model is the interior temperature. The interior temperature variable is a composite function of the thirteen interior sensors due to high correlation amongst the sensors. An additional interior temperature sensor is highly impacted by solar loading conditions and is excluded from the data analysis process due to its discrepancies in interior temperature readings. The composite interior temperature variable provides insight into the interior conditions within the residence. The inputs or independent variables used in the model are exterior temperature, power consumption (i.e., HVAC, water heater, etc.) humidity, CO₂, water flow rates and solar hot water temperature. In this dissertation, initially, humidity will be used as an independent variable to obtain performance results and with further investigation; it will be used as a dependent variable along with the interior temperature to observe the prediction accuracy of the interior condition.

I. Significance of Study

The significance of this study is to increase decision making opportunities within an intelligent framework and achieve accurate predictions of thermal comfort levels. This study adds to the functionality of the previously related work in the area of intelligent demand response models by predicting the interior thermal response conditions in response to a demand response scenario being considered. This information will better enhance the autonomous decision-making on behalf of the consumer to align with their comfort preferences as well as identify the contributing loads that influence the interior conditions to be modified for optimization and control purposes. Therefore, the current work adds intelligence to facilitate residential demand-side management of

energy usage in order to achieve demand response and assist in maintaining thermal comfort levels.

J. Dissertation Structure

Chapter I provides a brief overview to identify the motivation for this study and the specific problem to be addressed by developing the set of multivariate ARIMA time series ETHER models. In this chapter, the targeted goals are identified to set the stage for execution of our research methodology and evaluation approaches and techniques that will be applied to additional datasets for validation of its significance and originality to maintain accurate model predictions.

Chapter II is the background chapter and provides the foundational research context and related work that is imperative to the understanding of our work. The background has been divided into six major areas (sections) that provide a basis of relevance to this field of study. The first section provides a brief summary of the evolution of the grid, which discusses the progression from the conventional electricity grid to the notion of a smarter grid to achieve interoperability and enhance communication opportunities amongst all the connected parties. The second section details the current demand response models that have been proposed in literature. The applicability and practicality of each model is detailed while identifying the gaps and areas of improvement. The third section specifically targets the intelligent demand response frameworks that are designed to act autonomously on behalf of the consumer. A brief discussion is provided to explain how our ETHER models can enhance its functionality for achievement of residential demand response. The fourth section highlights the demand response programs and scenarios that are currently being implemented as pilot programs to varying customer sets. Based on the scenario that is selected by the consumer to align with his/her preferences, electric appliances and devices are modified to achieve a reduction in energy consumption. The fifth section discusses the thermal network design which is the typical method that is used to predict thermal response

conditions of a building space. The sixth section provides an overview of how observational data has been applied in various disciplines to achieve acceptable results for model development.

Chapter III introduces the Intelligent Demand Energy Adaptive System (IDEAS) framework, which is needed to facilitate automated residential demand-side management of energy usage. The basic idea behind this framework is to bridge together all the demand response models that have currently been proposed or introduced in literature to achieve demand response for the end-user. This framework includes scenario models, financial impact models, ETHER models, etc. to work collectively together to serve as an intelligent agent for the consumer.

Chapter IV focuses on the individual data sources collected from our baseline residential project. There are a total of forty five sensors located on the inside and outside of the residence. This information provides insight into the conditions and behavioral activity of various devices and appliances within the home which are important to understanding their impact on the interior temperature of the residence.

Chapter V highlights the approaches that were designed to achieve an accurate calculation and prediction of the interior temperature of the residence. There are a total of fourteen temperature sensors strategically placed throughout the residence. In this chapter, four approaches were investigated. The first approach combined all fourteen temperature sensors. The second approach eliminated the uncorrelated variable and averaged the remaining thirteen temperature sensors. The third approach evolved from the second approach by averaging the remaining thirteen temperature sensors; however, instead of eliminating the uncorrelated variable, the uncorrelated variable was used as an independent variable. The fourth approach focused on the prediction of the interior temperature by using the interior humidity as a dependent variable rather than an independent variable.

Chapter VI provides the research methodology used to develop the ETHER models. The general premise behind this methodology is to use an ARIMA process to develop the outside temperature and time of day conditional models based on the observational data collected from the residence. The time of day refers to daytime and nighttime (i.e., Day (6:00 AM – 7:59 PM) and Night (8:00 PM – 5:59 AM)). The outside temperature conditions are based on exterior temperature ranges (i.e., Hot (85°F +), Warm (70°F – 84°F), Cool (54°F – 69°F) and Cold (0°F – 53°F)). This method is used to characterize behavior within the residence during these specified conditions to take advantage of identifying the various impacts to the interior temperature.

Chapter VII details the goodness of fit measures used to validate the developed models and serves as evaluation measures for the prediction accuracy. The validation of the models is conducted by comparing the models' output to the actual data collected from the residence. For purposes of our research, we desire to maintain the prediction accuracy with 1 degree F of measurements.

Chapter VIII highlights the evaluation approaches that were investigated to increase the model prediction accuracy based on the number of variables and magnitude of data used within this study. Case studies are provided to compare each evaluation approach and as a process of elimination, determine which approach will be used for further analysis and model development purposes.

Chapter IX provides a basis of validation for our methodology and forecasting ETHER models using two additional datasets collected from different buildings. The buildings vary in size, structural materials and year of development; however, the observational data is used to provide insight into activity within each space. The same process and procedures that were conducted on our baseline dataset is enforced on the additional datasets. Figure I-6 shows a conceptual illustration of the process that was rendered to frame the methodology and approaches that were

used to develop the ETHER models and validate the methods to achieve acceptable results for prediction accuracy of the interior temperature within the building spaces.

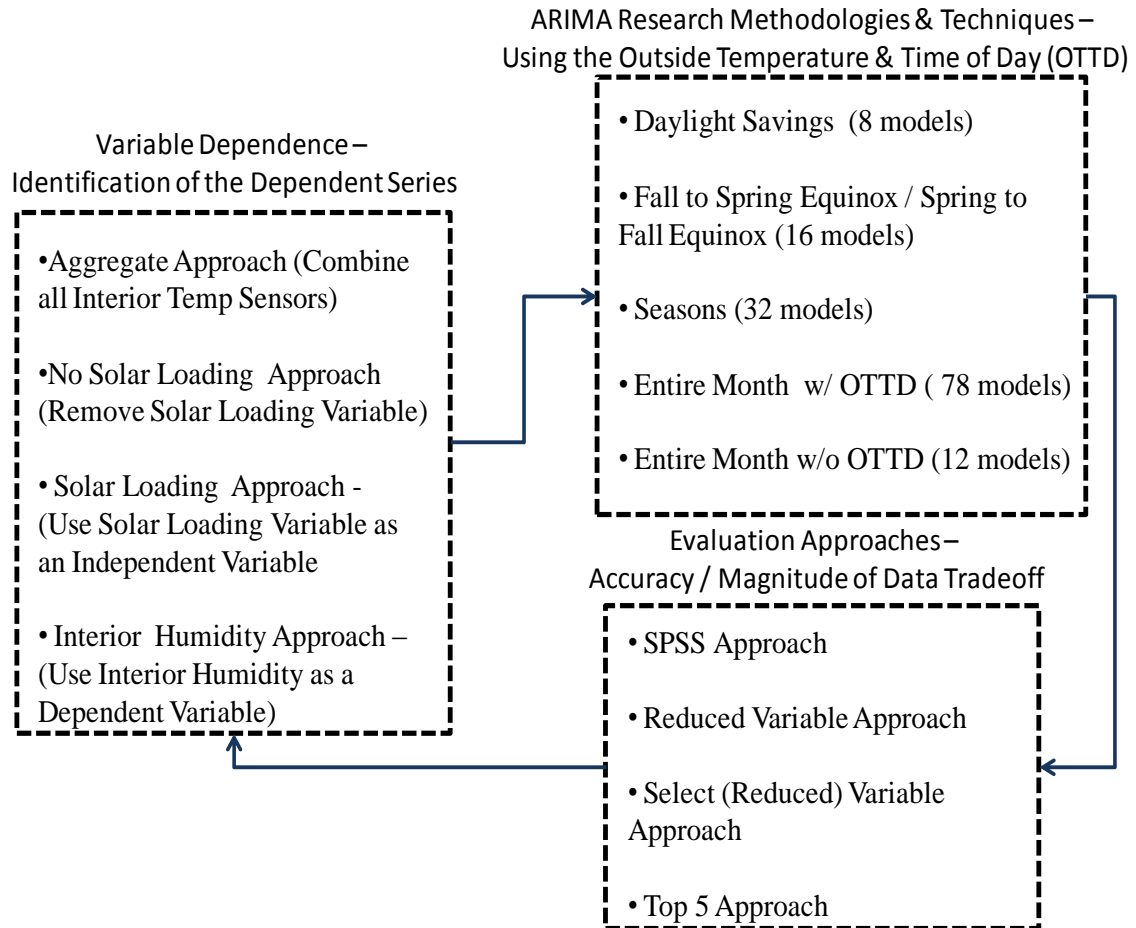


Figure I-6: ETHER Development Roadmap

Chapter X provides the conclusions of the work conducted and achieved within this dissertation. The contributions of our work are restated based on the initial set of goals that were put in place at the beginning of our research. Potential ideas of future work are also provided for expansion of this body of work. Two manuscripts have been submitted for journal publications and are currently under review.

II. Research Context & Related Work

This chapter discusses the gaps within prior research and identifies the methodologies that are related to the work presented in this dissertation. Previous literature has been reviewed within this subject matter to provide a foundational basis for further discussion of the problem that needs to be addressed. The main problem addressed by this body of work is the development of a methodology and a set of forecasting Evaluation of Thermal Response (ETHER) models to achieve demand response in a residential setting while also satisfying consumer needs. Methodologies and approaches that span over a range of disciplines have proven to be essential to understanding the existing gaps and areas for improvement within the demand response area of efficient energy consumption.

The context of the research is discussed by starting with the existing state of the electricity power grid to the evolution of the advanced technologies and techniques that encompass the concept of the Smart Grid. Primarily, the focus is centered on the benefits of advanced metering infrastructure and the capabilities that are afforded by its implementation. In addition, the previous work conducted in this area is highlighted to span across demand response models, demand response scenarios, thermal response models, observational data applicability and intelligent demand response models and schemes that are used to achieve energy reduction.

A. Evolution of the Grid

Traditionally, the utility environment only provides unidirectional power and information flow from the power plants to the demand-side and from the demand-side to the utility, respectively. As a result, there is no mechanism in place to alter the consumers' consumption patterns due to this one-way pipeline of information. Figure II-1 shows the traditional unidirectional flow of information.

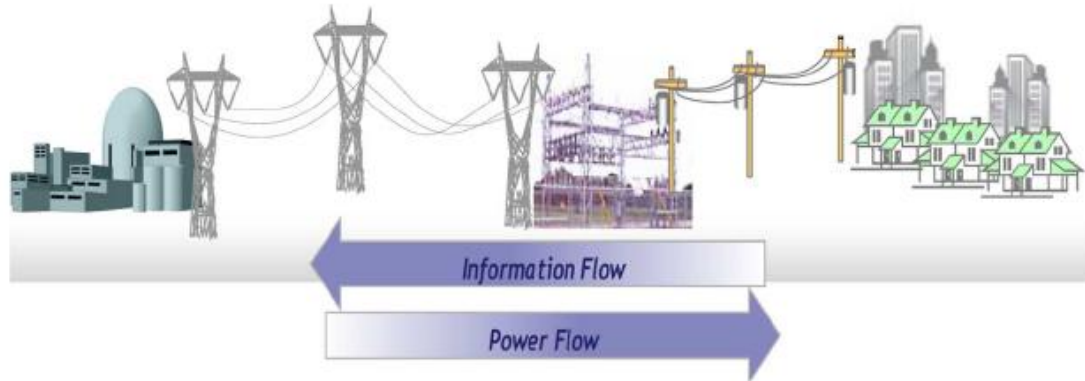


Figure II-1: Unidirectional Flow in the Traditional Utility Environment [26]

Adoption of the Smart Grid will, however, enhance every facet of electricity delivery to include generation, transmission, distribution and consumption. The Smart Grid encompasses the advanced technology (i.e., advanced metering infrastructure) to improve demand-side management, energy efficiency, and a self-healing electrical power grid [26]. The Smart Grid is an electricity network that is combined with Information and Communication Technologies (ICT) and techniques to deliver sustainable, economic and environmental-friendly resources to the end user [6]. The benefits of the advanced metering infrastructure include the establishment of bidirectional communication and interaction between all participating parties and seeking out available distributed resources to be used as energy providing products. The Smart Grid will aid in the integration of alternative sources of energy to be used and enable the utilities to make more efficient use of their existing assets. Figure II-2 schematically shows the bidirectional flow of information through utilization of the Smart Grid.



Figure II-2: Bi-directional Flow enabled by Smart Grid Technologies [26]

Most importantly, the Smart Grid allows us to integrate, interface with and intelligently control our resources on the consumption level. Research has been initiated in studying the emerging use of thermal storage for peak shifting, the anticipated growth and cost reduction of using renewable sources for generation at the residential levels, and the availability of bi-directional secure communications network which are anticipated to significantly alter the nature of future power system operations, as well as consumer behavior [8]. Enabled by the Smart Grid technology, consumers will have the opportunity to manage their energy consumptions during peak periods, resulting in reduced energy costs.

With the emerging smart grid environment, demand response is viewed as a critical capability to aid in shaping the electricity industry. Demand response is defined as a capability that allows for voluntary reduction of energy during peak periods of the day. The benefits of demand response are the achievements of system stability and reliability when the electric power grid is overly taxed. Various demand response pilot programs have been initiated throughout the U.S. to engage consumers in participating to reduce energy consumption and provide situational awareness to the utilities as well [28], [24]. The average consumer is not concerned with the amount of electricity that is being used until the electricity bill finally arrives. Being able to account for and have

knowledge into the varying electricity costs can help consumers be mindful of how to best optimize usage of appliances and devices within each household.

In our research, the utility signals that are sent to consumers in an effort to achieve demand response are request for reduction signals and price signals. The request for reduction signal is a determinant of usage that is sent when there is a detection of instability within the power system. The real time electricity price signals are also provided to consumers through Smart Grid technology. Based on certain demand response programs, consumers are able to receive notification of a price increase several minutes prior to the corresponding hour to adjust consumption usage accordingly. The price signal is an effective deterrent for wasteful consumption especially when the prices are significantly higher. Both signals are informative indicators to the consumers thereby allowing the consumers to change their consumption behaviors; however, the average consumer does not have the time to respond to every signal nor do they want the inconvenience of searching for appliances to render inoperable.

Therefore, automating consumer demand response to support the smart grid paradigm alleviates the dependence on the consumer and enables intelligent decisions to be enforced that consider information provided by utilities (i.e., price variation or demand reduction requests) to benefit the consumer and their thermal comfort needs. The air-conditioning unit (i.e., HVAC system) is the most sizeable load in a residence. Typically, during peak demand periods, utilities will want to reduce usage of the HVAC system to minimize energy consumption. This, in turn, may impact the consumers' comfort levels. Maintaining thermal comfort levels while aiding in the achievement of demand response is the centralized focus of the ETHER models.

The work presented in this dissertation adds to the functionality of the previously related work in the area of intelligent demand response models by predicting the interior thermal response conditions in response to a demand response scenario being considered. This information will

better enhance the autonomous decision-making on behalf of the consumer to align with their comfort preferences as well as identify the contributing loads that influence the interior conditions to be modified for optimization and control purposes. Therefore, the ETHER models add intelligence to facilitate demand-side management of energy usage in order to achieve demand response and assist in maintaining thermal comfort levels.

B. Demand Response Models

In literature, various demand response models have been developed all for the common purpose of adjusting the electrical demand in response to utility signals. Demand response models enable the adjustment of electrical loads to be shifted to off-peak or lower demand periods as well as rendered inoperable if there is no impact to the consumer. The benefits of demand response models include the improvement of operational efficiency, risk reduction of system instability, and reduction of energy costs.

This section summarizes the most relevant demand response models related to our research area. The brief synopsis of each body of work provides the basis for the developed methodologies and helps to frame our research contribution. In addition, the comparisons and differences between the scopes of work are highlighted to enhance the understanding of the various methods that were developed.

B.1. Adjustment of Loads based on Hourly Electricity Prices

Conejo et al. [27] developed a real-time demand response model using a linear programming algorithm to adjust loads on an hourly basis in response to hourly electricity prices. Enabled by the Smart Grid, the concept of the rolling window model was used as the basis of their research to derive optimal consumption levels for the current hour to be transmitted to the supplier. The electricity prices were forecasted over a 24 hour period by using an Autoregressive Integrated Moving Average (ARIMA) based model while adjusting the price bounds as real data became

available. The price signals are assumed to be received ten minutes ahead of each hour. Thus, allowing a consumer the opportunity to adapt its hourly load consumption in response to electricity price signals. This research requires manual human intervention in which the consumer sets and resets changes in their energy consumption patterns. The researchers envision that the rolling window linear programming algorithm can be integrated into the Energy Management System within a residence to allow for the minimization of energy cost and the maximizing of the consumer and utility interaction.

Our research used ARIMA-based models as our core modeling technique. In comparison, an assumption was made that price signals and request for reduction signals would be received ten minutes ahead of each hour to allow for prediction over the 6 hour forecast horizon. With this notion, optimization of the electrical loads could be processed and evaluated to determine their impact on the interior conditions within the residence. Conejo et al.'s work did not focus on how optimizing energy consumption on an hourly basis based on electricity prices impacted the consumer's thermal comfort levels. Our ETHER models are static models composed of demand parameters that greatly impact the interior conditions based on past activity patterns that occur during different outside temperature and time of day conditions. However, there is a mode of refinement of the ETHER models as actual data for the demand parameters becomes available, post the specified forecast horizon. In addition, we envision our work to be embedded into an intelligent framework to act autonomously on behalf of the consumer to limit the need for human intervention.

B.2. Electric Power Consumption based on Past Data

Hobby developed a methodology to construct a demand response model for electric power consumption [28]. The model uses past observational data from the U.S. government's Residential Energy Consumption Survey (RECS) to determine the necessary parameters needed to fit the demand function to achieve real time demand response. The survey contains energy

consumption data from several thousand randomly selected households. The data covered housing unit characteristics (i.e., insulation, structural materials, etc.), kitchen appliances, electronic devices, space heating, water heating, space cooling, lighting, household income status, electricity prices and climate data, etc. Hobby used a least-squares fit method to determine which parameters from the RECS database best represented the formulation of the equation to calculate electric power consumption. Similarly to other databases, the RECS database had missing values in its data files. To mitigate this issue, the missing values were replaced by information from outside sources which in turn decreased the accuracy of some of the information. A resultant simplified demand function was developed to fit all the data from the thousand randomly selected homes, which consisted of numerous unknown constants to be augmented with information. Based on the outcome of the model, Hobby discovered that several different demand functions are needed since demand parameters vary between households. For further exploration, the RECS household data will be clustered based on income and the same methodology will be applied.

By comparison, our work uses past observational data from an Energy Monitoring System extracted from one residence. This data is used to construct a thermal response model to predict interior conditions within the household over a 6 hour forecast horizon. The collected data includes kitchen appliances, electronic devices, space heating and space cooling, environmental conditions (i.e., temperature and humidity), lighting, etc. Income and electricity prices were not factored into our model development as independent variables. In addition, we were presented with a similar constraint with missing values. Our mitigation to this issue was to remove all missing values from the data during the data preparation phase and process the data as is for simplification of model development and analysis. As a result, eight simplified low-fidelity ETHER models were developed to assist in an effort to achieve demand response, whereas Hobby developed one demand function and concluded that several demand function models are required to represent different household specifications.

B.3. Energy Scheduling based on the Behavior of Consumers

O'Neill et al. [29] developed a demand response model to reduce residential energy costs by learning the behavior of residential consumers and automatically making decisions regarding energy scheduling and allocation. The pricing signal is the primary indicator for activation of the energy reduction model. This body of work presents an adaptive and optimal automated Energy Management System to be used at the consumption level for the residential consumer. The consumer can make an energy reservation for the devices or appliances within the home that needs to be fully functional. The model, in turn, decides the optimal time to run the desired devices to minimize long term costs for the consumer. The model uses an online learning application, Q-learning, that allows the Consumer Automated Energy Management System (CAES) algorithm to learn the behaviors of the consumer and make energy allocation decisions, accordingly [30]. The CAES algorithm relies heavily on the statistical behavior of the consumers and energy pricing and factors in long term financial costs to schedule energy usage.

Our ETHER models are also focused on the residential consumer (e.g., single residence). The models are not designed to learn the behavior of consumers; however, they are intended to work as a part of an intelligent framework that functions and responds to utility signals on behalf of the consumer. We consider both price signals and request for reduction signals as activation signals to engage our ETHER models. O'Neill et al.'s model learns the behavior of the consumer based on the consumer reservations and statistical relationships among the consumers' selection. In our work, an assumption has been made that consumer preferences will be an input into the framework which will add value to the evaluation of our ETHER models based on the statistical relationship of the controllable and uncontrollable loads that impact the interior conditions.

B.4. Physical-based Residential Loads at the Appliance Level

Shao et al (i.e., Virginia Tech researchers) developed a set of demand response-enabled load models within a residential setting at the appliance level [33]. The loads are categorized as

controllable and critical loads. The controllable loads (e.g., HVAC, water heating and the dryer) are defined as the loads that have minimal impact on the consumer's standard of living and the critical loads (e.g., cooking, lighting, refrigerator) have a noticeable impact on the consumers when modifications are made to their settings [31]. The purpose of their research is to develop load models that can facilitate the study of changes in electricity consumption in response to customer behavior and/or signals from a utility [31]. Specifically concentrating on the developed space cooling and space heating load models, these models required the thermodynamic principles of the building structure as built-in parameters in the load models in addition to the heating unit characteristics. The validation of the space heating and cooling models were performed by using data (i.e., house structure parameters, HVAC unit size, etc.) from the ecoMOD 1 prototype house (i.e., the residence used in our research) as input values and comparing the model output with the actual measurements.

Our ETHER models take into consideration both controllable (i.e., HVAC, water heater, washer/dryer, etc.) and uncontrollable components (i.e., exterior temperature, humidity, etc.) which impact operational conditions within the residence. In our body of work, the HVAC system would be considered a critical load by the definition presented by Shao et al [33]. Evaluation of the impact on the consumer comfort levels based on modification of the appliance level loads was not a part of their scope; however, this is part of the scope of our work to include changes in environmental conditions. The ETHER models are constructed based on past observational data which eliminates the complexity, assumptions and speculations on how to calculate the area of the ceiling, walls, and windows, the solar heat gain coefficient values, and the thermal resistance of the insulation, etc. within the residence. Our research fosters the opportunity to predict interior conditions over a short duration of time to align with customer preferences for thermal comfort as well as study the impact of controllable components as changes in energy consumption are adjusted.

B.5. Residential Energy Management based on an Optimal Stopping

Method

Iwayemi et al. developed an automated residential energy management and scheduling platform using an optimal stopping rules method to achieve demand response [11]. Through the use of a home energy controller (HEC), the optimal scheduling method determines the best time to activate a household appliance based on the observed series (i.e., energy profiles) of those devices in order to minimize costs for the consumer [32]. Their work is focused on the pricing signal for activation of the energy reduction platform. The HEC uses electricity prices to determine the mode (i.e., active or standby) of each device and decide upon the strategic scheduling to balance energy usage within the residence. Studies have shown that loads operating in standby mode can account for approximately ten percent of the household energy usage. The work proposed by Iwayemi et al. eliminates the opportunity for energy wastage and automatically makes decisions that will benefit the consumer. The overall objective of the optimal stopping rule model is to minimize the total power consumption and reduce the cost of the consumer's electricity bill.

In comparison, our work uses the energy profile data of the appliances as well as the observational data of the temperature, humidity, number of people in the household, etc. to act as part of an intelligent automated framework to achieve residential demand response. Validation of our ETHER models is conducted by performing an evaluation on the optimal settings of the devices within the household to predict interior conditions of the residence over the 6 hour forecast horizon. In addition, the ARIMA-based method used to construct our ETHER models identifies the loads that influence the dependent variable series. The ETHER models are designed specifically to obtain information on the thermal comfort levels of the consumers while identifying the contributing loads which can benefit and add value to an automated energy management platform.

B.6. Simulated Residential Energy Management System (RES)

Roe et al. proposed a residential energy management system (RES) to achieve automated demand response on behalf of the consumer. The purpose of their work is to provide a control system that provides visibility of energy usage and planning for smart residential buildings [33]. The framework of the RES testbed is composed of primary actors and secondary actors. The primary actors consist of the consumer, the utility, and the home energy controller. These three actors can initiate the system events for the RES testbed. In addition, the controller manages the system components to meet the requirements of the consumer and the utility. The secondary actors consist of the controllable (i.e., electric appliances, HVAC, refrigerator, etc.) and uncontrollable (i.e., small residential plug loads) appliances and devices in the residence. These actors respond to the commands provided by the primary actors. As an input into the RES testbed, the demand response service times are assumed to be provided at least 24 hours in advance. This information provides the management system time to prioritize appliance activation based on the consumer's request and the utilities' energy reduction request.

Our ETHER models are designed to be an automated agent to provide insight to an intelligent framework, like a RES testbed, based on optimization settings of household devices and customer preferences. Our work assumes that pricing and the request for demand signals are provided minutes before the proceeding hour for forecast evaluation which is a lessened time interval compared to the notification timeframe needed for the RES testbed. Thereby, the ETHER models can support the planning of energy usage and provide information that can aid in meeting energy demand reductions.

C. Intelligent Demand Response Framework

Intelligent models and schemes have been proposed on the consumer-side in literature for the achievement of demand response and efficient utilization of energy. These models act on behalf of the consumer to make decisions regarding the rescheduling of the appliance usage and the

optimization of the settings, etc. The primary objective is to lessen impositions upon the consumers to respond to the initiation signals from the utility while meeting the requirements from the utility as well as the consumer.

As described in the previous sections, numerous demand response models have been developed for the residential consumer. Some models required manual human intervention and many others functioned as an automated control system on behalf of the consumer to meet energy demand requests. As highlighted in literature, there are three types of demand response: fully automated, semi-automated and manual demand response [34], [35], [10]. 1) Fully automated demand response involves no human intervention. It entails an automated intelligent framework autonomously acting on behalf of the consumer. 2) Semi-automated demand response involves human intervention in which a demand response strategy or scenario is initiated by the consumer. Once the scenario has been initiated, the intelligent framework abides by planned strategic policies to achieve demand response protocols. 3) Manual demand response involves receiving an external signal from the utility and the consumer manually reduces the demand of the loads to meet the request for reduction.

In this section, we will highlight two intelligent fully automated demand response frameworks on the consumer-side which are focused on reducing energy usage, achieving cost savings and eliminating human decision making. Both frameworks utilize external interfaces and data to generate an effective strategy to achieve demand response. The primary target of one of the intelligent frameworks is the residential consumer; the other specifically targets the industrial consumer. The methodologies and approaches presented for the industrial end user, however, can be adaptable to cater to the residential consumer.

C.1. Intelligent Decision Support System (IDSS) Model

Sianaki et al. developed an intelligent system known as the Intelligent Design Support System (IDSS) which captures the residential consumer's preference and consumption profiles to assist the system in changing the consumption behavior according to the dynamic pricing signals as a means to achieve demand response [6]. This framework connected with Smart Grid enabled technologies (i.e., smart meter) is designed for real time demand response operations. The IDSS system, developed as a fuzzy multi-criteria decision-making and neural network based model, acts autonomously on behalf of the consumer to facilitate load management on a continuous basis when the consumers modify their preferences in different scenarios (i.e., dynamic pricing, several occupants in the house, etc.) via an external portal. The IDSS model is processed in four steps: 1) identify the different types of variables needed to acquire information on the consumer's preferences and energy profiles of various devices within the residence, 2) develop a user interface (e.g., web portal) that stores the consumer's inputs regarding the variables and preferences, 3) capture the exogenous variables (i.e., price signals, exterior temperature, availability of renewable sources, etc.) as a basis of information for the decision making model and 4) develop a neural network model that learns the consumer's preferences and energy profiles based on data stored in the user interface [6]. The decision making model operates to accommodate to two different objectives. If the objective is to achieve cost reduction, the rendered system behavior is to maintain the standard of living for the individual consumer. If the objective is to satisfy the needs of various occupants in the residence, the system behavior finds the best medium to meet the temperature and economical consumption requirements to achieve demand response in the midst of pleasing all the occupants. Different changes in the behavior and preferences of the consumer are used to develop the learning model to act on behalf of the consumer(s).

The IDSS system is tracking the consumer's behavior and developing a model that mimics their behavior. ETHER, in contrast, is using data that reflects the physical attributes of the system (i.e., sensor data recording temperature, energy profiles, etc.) to predict physical results (e.g. interior temperature) from decisions made by the consumer. The ETHER models are focused on forecasting the impact of the interior temperature based on variations of controllable and uncontrollable factors (i.e., energy consumption and environmental condition variations) to assist the consumer in making real time decisions to achieve demand response. In addition, we envision our ETHER models to be a part of an intelligent framework system, similar to the IDSS system, to act as an automated agent to add value and needed information to manage the thermal response comfort levels thereby aiding in the achievement of demand response.

C.2. Intelligent Demand Response Module

Dam et al. developed an automated demand response module for the industrial consumer; however, the underlying premise of the work can potentially be used for the residential consumer as well. This approach has been proposed in a semi-automated structure, meaning there is manual human intervention to initiate a pre-programmed demand response strategy within a decision module. The demand response module, designed as an expert and fuzzy logic system, makes decisions on behalf of the consumer based on load management policies and potential financial benefits to decide whether to follow the recommendation (i.e., shut down or reduce load) by the utility or to decline the acceptance of the signal [10]. Declining the acceptance of the utility signal can result in penalties for the consumer based on the demand response program.

Expert systems consist of a precise set of statements and a look-up table of values based on pre-defined rules of inputs present in the system. On the other hand, fuzzy logic systems consist of a set of decision rules to deal with vague or ambiguous information. The combination of the expert systems and the fuzzy logic systems presents the case of dealing with the best of both worlds where an evaluation can be conducted on precise and imprecise information in a logical way.

The load management policies are a set of rules that cannot be violated by the demand response module. These rules consist of requirements pertaining to the maximum number of interruptions, identification of critical loads, maximum amount of load reductions, etc. [10]. In addition, an importance factor is taken into consideration which determines how crucial a load is to the meet the needs of the consumer. The importance factor serves as a prioritized ranking of which loads should not be impacted unless deemed necessary for energy reduction. The scope of Dam et al.'s work is focused on obtaining information from external interfaces (i.e., utility signal, load data, electricity prices, etc.) while abiding by the load management policies to make the best decision on whether to accept or reject the demand response request signal on the consumer-side.

Industrial consumers are concerned with making decisions pertaining to the operation of sizeable motors of different production lines; where the residential consumers are concerned with reducing lighting and air conditioning demands to achieve demand response. Regardless of the primary target, demand response refers to short-term energy consumption modifications by all customer sets. The scope of the ETHER models is not centered on the financial incentives or penalties that may be inflicted based on the acceptance or decline of the utility signal; however, the ETHER models provide information related to interior conditions which can aid in the evaluation process to arrive at the best decision.

D. Demand Response Programs and Strategies

Demand response programs and strategies have been offered by utilities in order to encourage consumers to use their energy efficiently and reduce energy consumption during peak periods to result in financial cost savings and increase on the system stability [36]. The advanced metering infrastructure (i.e., smart meter) is the core component to enable the capability for information flow to benefit both the consumer and the utility. Through participation of the demand response programs enabled by Smart Grid technologies, consumers are offered financial incentives and credits for their voluntary participation. However, if a utility signal is declined or ignored, a

penalty is invoked on the consumer. Most demand response programs have been conducted as part of pilot programs to test their effectiveness in performance as it pertains to the utility and the consumer.

Consumers participating in demand response programs can either shift their loads to off-peak periods or simply reduce the load during peak periods to achieve demand response [25]. Research has been conducted to develop decision models in a semi-automated and fully automated environment to aid in making a calculated decision to meet the demand of the utility and the consumer. Automated demand response programs and strategies present the ability to reduce congestion and demand on the electric power grid in a more efficient way [37].

In this dissertation, the demand response strategies (i.e., Load Shut Down and Load Reduction) represent the acceptable actions to take in response to the request for reduction signals. In Chapter III, these strategies will be discussed in greater detail as we introduce the intelligent framework known as the Intelligent Energy Adaptive System (IDEAS) and all the individualized demand response models that contribute to its practicality to function on behalf of the consumer. Under the load shut down scenario, the utility sends a load shut down signal to the consumer to disconnect the most sizeable loads (i.e., HVAC system, hot water heater, etc.) from the network or connect the loads to a renewable generating source for operation. Under the load reduction scenario, the utility sends a signal to the consumer to reduce its energy consumption by a certain amount [10]. Our ETHER models do not incorporate pre-programmed demand response strategies for energy reduction; however, the demand response strategy model discussed in Chapter III shows that based on the outcome of the scenario selection, optimization settings for the appliances are determined, which serves as a required input into the ETHER models for evaluation of the thermal response conditions.

E. Thermal Modeling Design

The primary focus of our research is to predict the thermal response conditions within a residence using data specifically related to its physical attributes to gain insight into the behavioral aspects of energy use in an effort to achieve demand response. The investigation of thermal response has primarily been explored by using a thermal network design model. The thermal network model, shown in Figure II-3, uses an electric analog RC network circuit to predict and analyze conditions within a building's structure for optimization through simulating processes of conduction, convection and radiation in the building under specific environmental conditions to achieve energy reduction [38], [39].

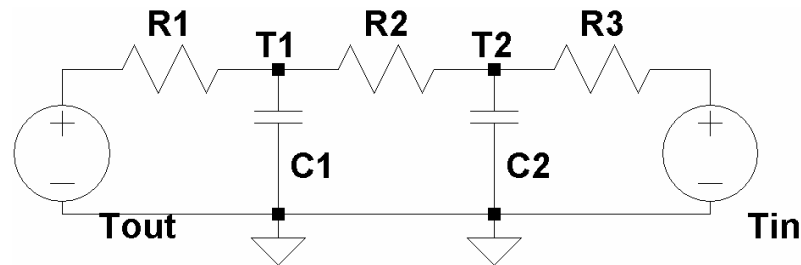


Figure II-3: Simplified thermal network model

The resistors, i.e., R1, R2 and R3, represent the thermal resistance describing the walls of the building or the obstruction of air flow through the structure. The capacitors, i.e., C1 and C2, represent the thermal capacitance or heat storage within the structure. The sources, i.e., Tout and Tin, represent the outside air temperature and the mechanical ventilation (i.e., heating and cooling) source used to control the interior air temperature, respectively. The nodes, i.e., T1 and T2, represent the outdoor surface temperature node and the indoor surface temperature node, respectively. The following parameters which can be additives to the thermal network schematic above are most often used to help simulate the thermal response of the building structure: e.g. 1) the thermal effects of the outdoor air, 2) solar radiation as well as shading, 3) color and orientation of different exterior surfaces, 4) solar penetration and window shading devices, 5)

internal convective and radiative heat sources as well as natural and mechanical ventilation [38]. The challenge with this predictive thermal response modeling technique is that it is highly unlikely that the complete details of the building structure (i.e., thermodynamic properties of all the materials) are that well known or that the thermal characteristics can easily be predicted. With, these uncertainties, combined with the uncertainties of environmental condition predictions and the behavior of the consumers; prediction accuracy is strictly limited [40].

Our objective is to develop an approach that will permit the extraction of a set of ETHER models from observations of the building in operation (i.e., the building's performance data); resulting in a set of self-constructing and self-refining models. The thermal network models are static in their formation. ETHER models, by contrast, can be constantly refined as operational and structural conditions change. Utilizing the ARIMA-based approach to develop the ETHER models represents a simple methodology using observational data to arrive at highly comparable and in most cases better prediction accuracy of the thermal response conditions.

F. Observational Data Applicability

Observational data has been used in many cases as a basis for inverse modeling. Inverse models are derived from empirical behavior based on a set of empirical parameters. In essence, the observational data is used to “back out” parameters that provide the best representation for the chosen model [41]. Compounding the inverse modeling methodology with time series analysis, various models have been developed that produce accurate predictions of their desired dependent variable. Our ETHER models use observational data to develop a time series ARIMA-based model structure for forecasting interior conditions up to 6 hours in advance to achieve demand response.

Kusiak et al. developed a multivariate time series model to predict power ramp rates of wind farm power over 10-minute intervals using observational data [41]. These researchers sought to extract

models based on observational historical data stored in the Supervisory Control and Data Acquisition (SCADA) system to enhance the management of the electric power grid and perform analysis to optimize the power ramp rate (i.e., dependent variable). Data mining approaches were used for evaluation to identify independent variables or predictors that greatly influence the power ramp rate. In comparison to our work, Kusiak et al. eliminated datasets that contained an excessive number of missing values for simplification and in an effort to maximize the prediction accuracy of the models.

Contreras et al. developed ARIMA models to forecast day-ahead electricity prices to develop strategies for negotiation in the energy market [20]. These models were constructed using an ARIMA-based time series analysis methodology. This scheme was executed in five steps: 1) develop a hypothesis, 2) use observed data to develop a model, 3) estimate the demand parameters, 4) validate the hypothesis and 5) use model for forecasting. The same sequence of steps was used to develop the ETHER models. Obtaining accurate predictions of the electricity prices can provide valued insight into the scheduled allocation and modification of energy consumption on the utility and consumer-side. This information can better equip an intelligent framework in acting autonomously on behalf of the consumer.

Razak et al. developed a seasonal ARIMA model to conduct load forecasting using data mining techniques [21]. The researchers used data supplied by a utility to observe the behavior and load patterns based on seasonal influence and the variation of weekday versus weekend. Environmental conditions and consumer preferences were not factored in as inputs into the development of the models. As a result, the model accuracy was lowered. Validation of these models was conducted by using a similar goodness to fit measure (i.e. mean absolute percentage error (MAPE), shown in our work, to compare the measured data to the actual data provided by the utility. Our ETHER models factor in the environmental conditions and the seasonal influence to achieve an acceptable set of models.

Each referenced model was built using historical observation data and focused on the prediction of a dependent series using time series analysis. Compounding each model could add to the enhancement of achieving demand response through utilization of an intelligent framework. In the next chapter, an intelligent framework is referenced to include various demand response models for the common purpose of ensuring financial savings, thermal comfort and energy reduction. However, this framework can easily be expanded to encompass the innovative work that has been proposed by various researchers in literature to aid in the achievement of demand response for the residential consumer.

III. Intelligent Demand Energy Adaptive System (IDEAS)

Residential demand-side management (i.e., demand response) has been overlooked in some instances due to the reluctance or inability of the consumer to modify their energy consumption patterns. In most situations, residential consumers have neither all the required knowledge nor sufficient time to search for appliances to reduce and/or meet energy reduction demands. However, involvement in demand-side management by the residential consumer can assist the utility from bringing expensive peak plants on line to supply power, which would increase generation costs. Benefits related to demand-side management include shifting demand away from peak demand periods, reducing overall energy consumption and reducing the need for additional power plants [32]. The net results are cost savings and the reduction of energy wastage for both parties.

The Smart Grid paradigm allows residential consumers to actively participate in demand response by enabling access to real-time information on energy usage through the deployment of new technologies (i.e., metering devices, etc.). However, technologies that are deployed to residential consumers provide limited information (i.e. pricing and energy usage), leaving consumers to determine which modifications need to be made. In other words, these technologies do not provide a level of information that identifies the particular load appliance(s) that needs to be reduced to comply with the reduction signals. Therefore, the need exists for the implementation of an intelligent demand-side management system to promote active involvement and achievement of energy reduction, to identify particular load appliances for reduction and to benefit the consumer (i.e., financial incentive).

An intelligent demand energy adaptive system (IDEAS), shown in Figure III-1, is needed to facilitate automated residential demand-side management of energy usage. An automated system, such as this, is necessary to avoid penalties that can be enforced upon the consumer when they do not respond to a load reduction signal. For example, in the curtailable load demand

response program, consumers participate by committing to modify their load consumption in response to a request signal. The benefit for the consumer can be represented in different forms such as lower electricity prices, bill credits, etc.; however, a penalty is administered if the consumer fails to comply [32]. Automation of an intelligent system is also needed to act as an autonomous agent on behalf the consumer. For example, in the direct load control demand response program, the utility determines the actions and implements various instructions to certain sizeable appliances within the residence [33]. The benefit to the consumer is a compensation fee for the inconvenience; however, the downside is the inconvenience factor (i.e., potential discomfort, the need to utilize an appliance that needs to be shutdown, etc.). It has been noted that 35 percent of consumers do not want utilities to control the appliances within their residence regardless of the cost savings and financial incentives [35], [32].

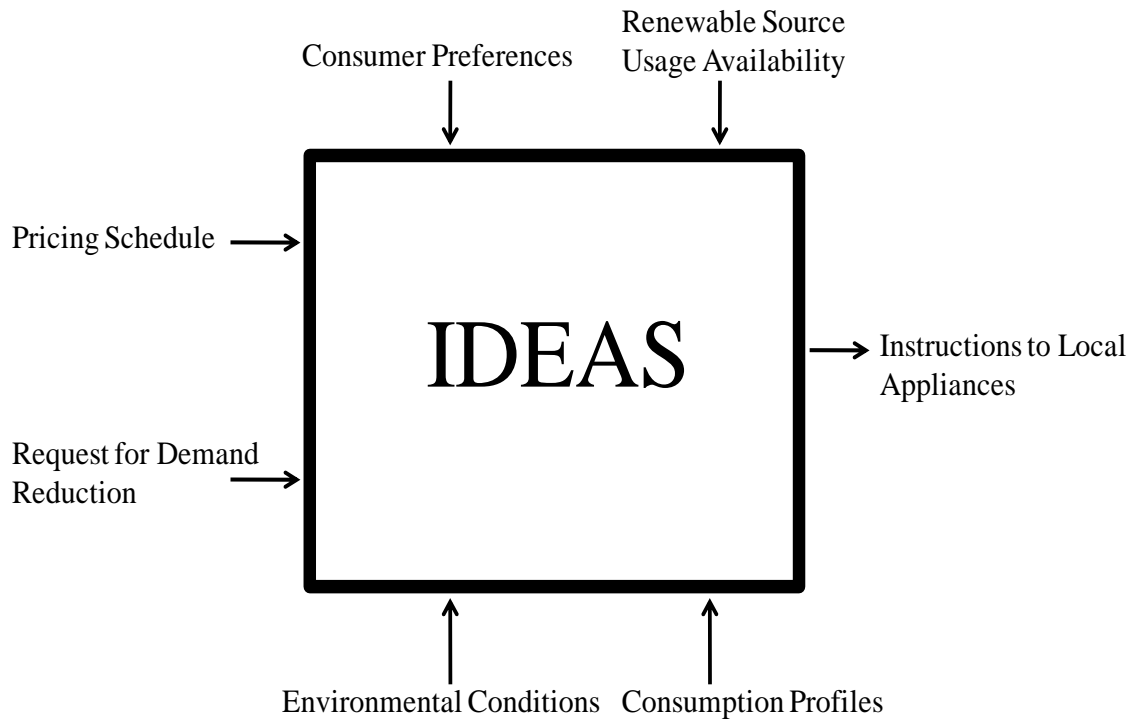


Figure III-1: Intelligent Demand Energy Adaptive System (IDEAS)

IDEAS, as shown in Figure III-1 would receive information from the utility regarding price schedules, needs for demand reductions, consumer incentives, etc. Additional information to IDEAS would include the consumer's preferences (i.e., comfort levels, etc.) and physical conditions (i.e., temperature, humidity, number of people within the home, etc.). Assumptions are made that the environmental conditions (i.e., predicted exterior temperature) are provided over a 24 hour period and the dynamic pricing notification is given an hour in advance. Based on the information received, IDEAS would use this collected information to control (i.e., shut down or reduce) the controllable appliances in the residence.

The IDEAS system is made up of internal models, shown in Figure III-2, that combine inputs to control appliances in order to achieve the consumer and utility goals. The overall IDEAS system has not been implemented as part of the research for this dissertation; however, the practicality of the system will be discussed by detailing the functionality and purpose of each model. Our contribution to the notion of this intelligent framework is the Evaluation of Thermal Response (ETHER) models, which will be highlighted in this section; moreover, the methodology and structure of the ETHER models (i.e., Autoregressive Integrated Moving Average (ARIMA) time series methodology) will be discussed in a later chapter.

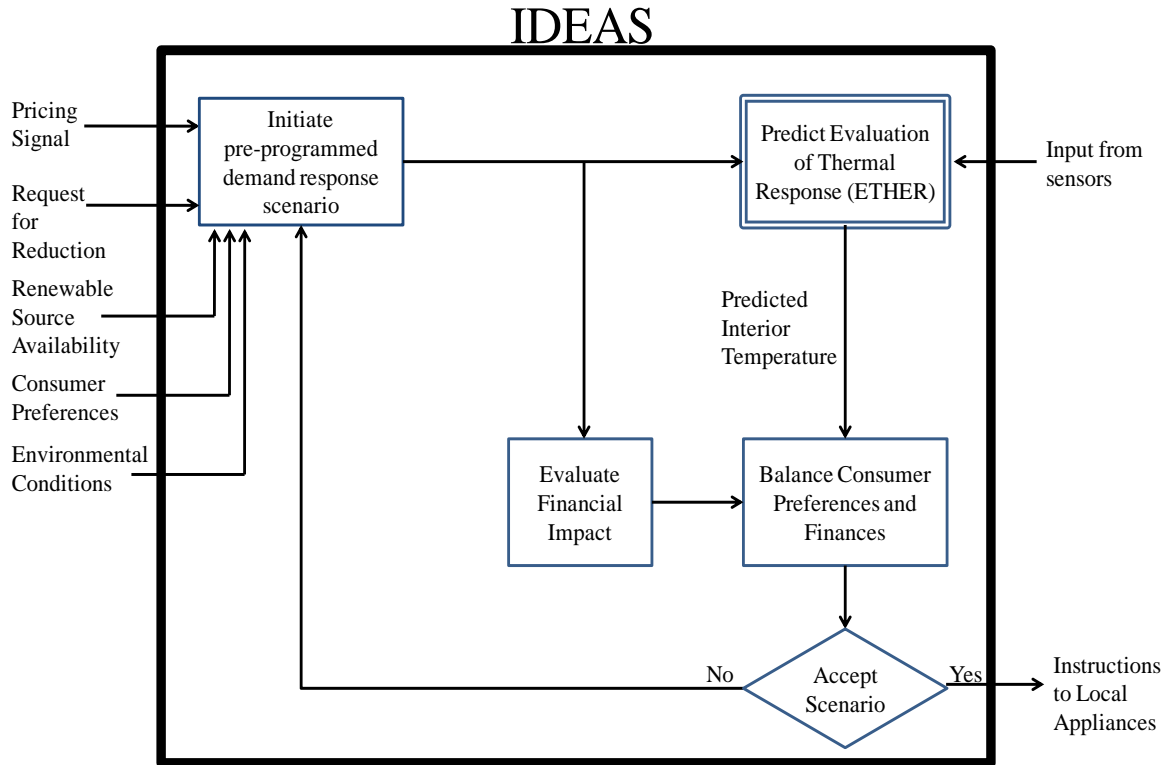


Figure III-2: Internal Models within the IDEAS system

A. Pre-Programmed Demand Response Scenario Model

The pre-programmed demand response scenario model receives all the external inputs to make an informed decision regarding the actions needed to reduce energy consumption. There are two demand response scenarios that have been primarily discussed in this dissertation: Direct Load Control (DLC) and Price-Based Control (PBC). In the DLC scenario, the utility has the ability to switch off and control appliances at times of peak demand. In the PBC scenario, consumers receive advanced notice of varying prices so that they can choose which appliance settings to change [13]. The intelligent system will allow for modifications to both scenarios to achieve the same result which is the reduction of energy consumption. Based on the DLC and PBC scenarios, the intelligent system will eliminate the intrusion of the utility company from

controlling certain appliances and eliminate the need for consumers to seek appliances to reduce in response to utility signals.

The demand response scenario model automatically decides which load appliances should be reduced or eliminated to meet the goals of the consumer and the utility without human intervention. The model is based on two scenarios: 1) reduction of appliance settings and/or 2) the shutdown of certain appliances. With these scenarios, optimization of the load profiles for the load appliances should also be investigated to target the amount of setting reduction. Figure III-3 illustrates a conceptual design of the decision-based process for selecting the most appropriate scenario for energy reduction by utilizing the external input information.

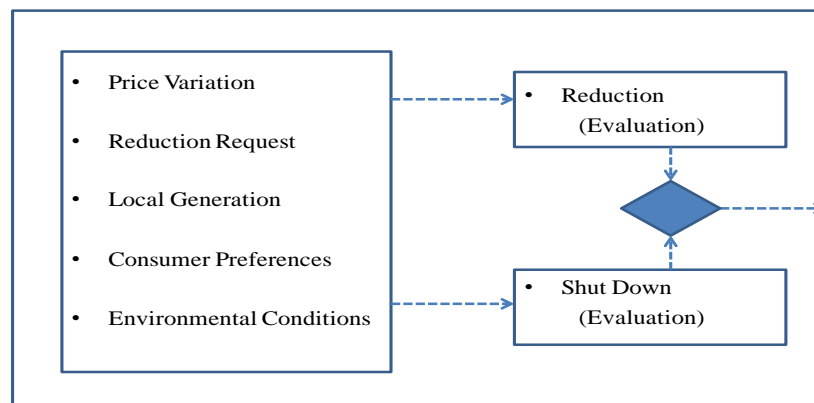


Figure III-3: Decision-based Process for the Pre-programmed Demand Response Scenario Model

The varying price signals are sent from the utility based on electricity prices. The price signal is used to serve as an indicator to control energy demand. Research has been conducted to look into a real-time pricing scheme which serves as a demonstration of autonomous control of consumer loads using a market mechanism at the distribution level [34]. In real-time and dynamic pricing schemes, the market, the utility, and the residence are connected to achieve demand response. The residence is connected to the market dispatch system at the distribution level, using Smart Grid information and technology communications, to receive bids and reflect hourly prices on

electricity use. Consumers, in turn, are able to view this information on an advanced metering device (i.e., smart meter) and are able to make load modifications based on economic preferences.

The request for reduction signals are sent from the utility when there is a demand supply imbalance in the grid (i.e., reliability of the grid); this can be caused by a shortage of energy generation based on the demand of energy. In addition, these signals are sent to the consumer to prevent the need to bring peak power plants on line to accommodate for peak demand periods. A request for reduction signal can make the difference between a reliable operation of the electric power grid and the potential of experiencing rolling blackouts [36].

The local generation sources are evaluated for useable purposes to assist with the minimization of energy consumption from the electric power grid. These sources can be represented as photovoltaic panels, wind turbines, etc. The availability of the energy generation from these renewable sources are processed by the model to determine the necessary actions to take based on the scenario selection.

The preset consumer preferences are also taken into consideration. The consumer input is essential to meeting their goals while selecting a scenario that will meet the goals on the utility as well. These preferences can span from personal comfort and convenience to economic preferences. This information assists in selecting a scenario that will tailor to the needs and wants of the consumer.

The environmental conditions, mainly concentrating on exterior temperature and exterior humidity, are needed to provide additional information for evaluation in the decision-based process for scenario selection. For example, this information will assist in the scenario bartering process of selecting a reduction in the settings of the HVAC system vs. a complete shutdown of the HVAC system. The weather forecast parameters can be extracted from the National Weather forecast database and can be updated based on the real-time weather conditions.

After the evaluation of the inputs, a scenario (i.e. reduction or shutdown) is automatically selected to enable demand response. The selection of the scenario activates two additional internal models within the IDEAS system: 1) Predict Evaluation of Thermal Response (ETHER) Model and 2) Evaluate Financial Impact. The ETHER model will be discussed in the next section.

B. Evaluation of Thermal Response (ETHER) Model

The purpose of the set of ETHER models is to predict the residence's thermal response using local input sensor data and output information from the pre-programmed demand response scenario (i.e., environmental conditions, time of day, load profiles, etc.). The ETHER models add value to the IDEAS system by providing information on the interior temperature conditions (i.e., thermal response) within the residence. This information can better assist the intelligent framework to align with the comfort preferences of the consumer.

The baseline residence (i.e. ecoMOD1) used to develop the ETHER models is equipped with a total of 45 sensors. The data collected by the input sensors includes: interior temperature, exterior temperature, power consumption (lighting, water heater, electronic devices and cooking appliances), humidity, CO₂, water flow rates and solar hot water temperature, etc.). In the case of residential consumers, studies have proven that developing models to achieve demand response is quite complicated due to the uncertain and inconsistent behavior of the consumer [37]. Therefore, by utilizing the data collected from the local sensors and the optimization of the load profiles based on the selection of the scenario, the set of ETHER models are used to predict the interior conditions.

A set of eight ETHER models were constructed based on outside temperature and time of day conditions (shown in Figure III-4). For contextual purposes of this research, time of day refers to day time and night time (i.e., Day (6:00 AM – 7:59 PM) and Night (8:00 PM – 5:59 AM)). The outside conditions are based on exterior temperature conditions (i.e. Hot (85°F +), Warm (70°F –

84°F), Cool (54°F – 69°F) and Cold (0°F – 53°F). A year’s worth of data was filtered and split into 8 datasets based on time of day and outside temperature ranges, where 70% of the data in each subset was used to build a training model and the remaining 30% was used to test each model structure. The outside temperature and time of day conditional models take into account the interior behavior of the residence based on the specified conditions (i.e., reduced setting of space heating and space cooling, etc.). Therefore, based on the time of day and outside temperature conditions, an ETHER model structure is selected that best fits the input conditions and used to predict the interior conditions in the residence up to 6 hours in advance.

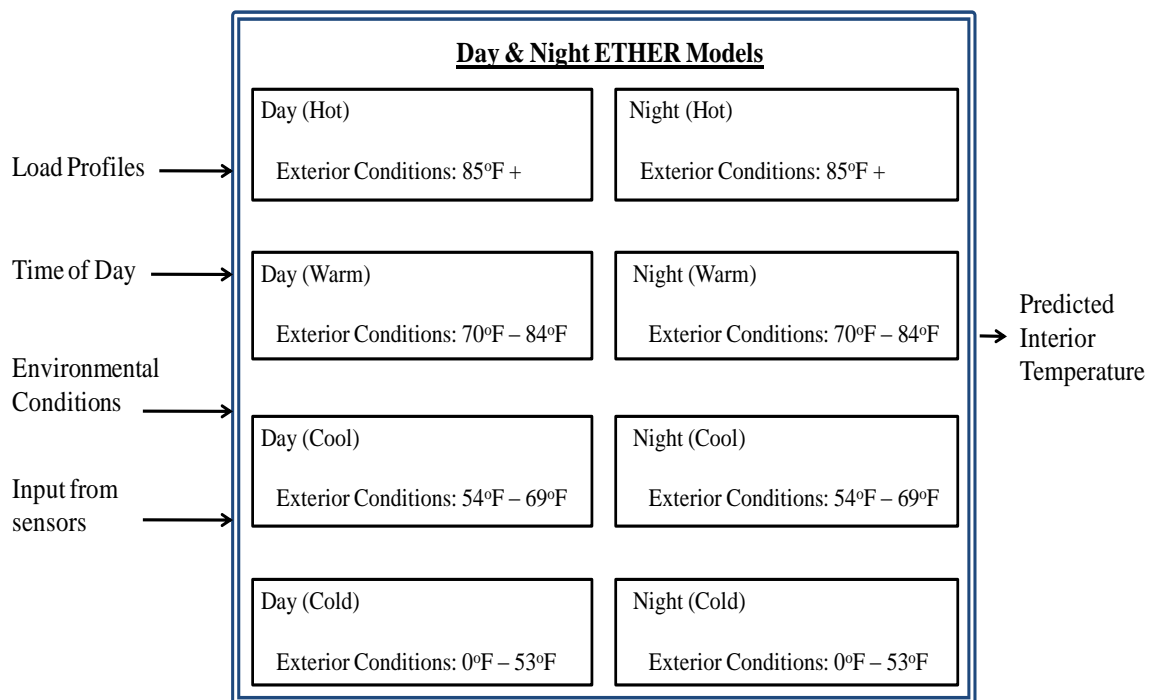


Figure III-4: Model selection process for the ETHER Models

The time range for the time of day selected for this research is based on the daylight savings’ sunrise and sunset times. Daylight savings time accounts for nine months out of the year (i.e. from the second Sunday in March until the first Sunday in November). During daylight savings time, the clocks are moved forward an hour which essentially means an hour of daylight is moved from the morning to the evening [37]. With a longer duration of daylight, the need for artificial

lighting within the residence is reduced. Our research assumes that the selected time bounds do not reduce the quality of models or the integrity of the data to predict the interior temperature conditions when the clocks revert back to standard time.

The exterior temperature ranges are selected based on the similarities of performance for interior space cooling and space heating within the residence. The descriptors for the temperature ranges (i.e., hot, warm, cool and cold) are loosely subjective to each individual; however, they are needed for model selection. In addition, the history weather website for the Charlottesville area [4] was used to verify additional environmental conditions that were needed for development of the models.

Different loads (i.e., cooking appliances, electronic equipment, space heating and space cooling, etc.) are being monitored within the residence by the sensors. The profiles for each monitored load are provided to the ETHER models to aid in the process of interior temperature prediction. These profiles are an essential input needed to guide the preparation and continuous refinement of the models.

The output of the ETHER models is the predicted interior temperature. Predictive interior temperature error bounds of +/- 1 degree F have been selected for this research. The predictive error bounds are based on the setpoint swing of the digital thermostat located in the residence. Typically, the setpoint swing or deadband for a residential thermostat is between 1 degree F or 2 degree F; however, a larger setpoint swing (i.e. 4 degree F) will reduce the constant cycling and power consumption activity of the HVAC system. The setpoint swing is used to maintain the interior temperature at a thermal comfort level within the residence; thus, the same bounds were set for the model predictions. In general, there are a number of factors that must be addressed when defining conditions for human thermal comfort which include: metabolic rate, clothing insulation, radiant temperature, air temperature, humidity and air speed [33]. For the purposes of

this research, we are focused on defining conditions for human thermal comfort through the investigation of interior air temperatures.

C. Financial Impact Model

After the selection of the scenario, a financial model is also activated to evaluate the financial impact of the chosen scenario. Figure III-5 illustrates the process-oriented evaluation for the financial impact model. The purpose of this model is to provide financial knowledge to the intelligent framework to financially benefit the consumer and provide economical details that will aid in meeting energy demands. For instance, the shutdown scenario has been selected from the demand response scenario model; however, lighting is identified as a must run service [25] during the peak demand period of the day. The financial model will assess the elevated cost for energy usage during the peak period respective to the residence as well as provide the result of the cost savings which are generated from shutting off non-essential appliances. Evaluation of the financial impact of demand modifications can provide an opportunity to better manage energy costs, balance and optimize demand usage, and improve energy efficiency [40].

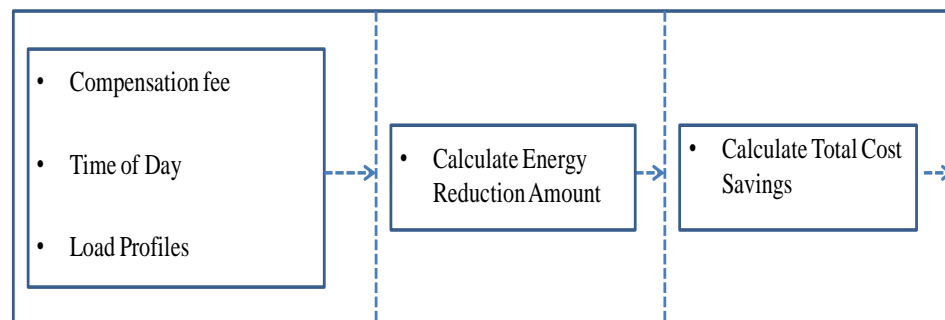


Figure III-5: Model Process to evaluate the financial impact

The following information will serve as inputs into the financial impact model: compensation fee, time of day and load profiles. The compensation fee is part of the consumer incentive package for reducing energy consumption in response to the utility signals (i.e., price variation and request for

reduction). The time of day indicator is needed to determine if the energy consumption is occurring during the peak hours of the day and to determine the varying prices of electricity during the specified time interval. Thirdly, the load profile information is needed to identify which particular appliances were modified. The next step to determine the financial impact is to calculate the total energy reduction amount based on the modifications of the load settings. All this information is required to calculate the total cost savings for the consumer.

D. Balance Consumer Preferences and Finances Model

The purpose of this model is to balance the comfort preferences with the economic preferences of the consumer. Figure III-6 illustrates the decision-based process of aligning the consumer's desired cost savings with their anticipated level of comfort. The level of thermal comfort is subjective amongst each individual consumer and there are many environmental, behavioral and personal factors that affect their comfort level [41]. Due to the fact that there are many variables and factors that affect the sensation of thermal comfort which makes it difficult to determine the measure of precision using models in an actual environment, an assumption was made that the actual interior temperature data collected from the buildings is a good indicator for the occupants' comfort range and characterizes the thermal perceptions of the occupants. As validation of this assumption, Wan et al [59] stated the variation in the interior temperature affects the indoor thermal comfort levels. However, it is noted that humidity and interior temperature are two main environmental factors that impact the thermal comfort of an occupant. For this study, the interior temperature will be investigated as a considerable variable for comfort to later be coupled with the interior humidity for future work.

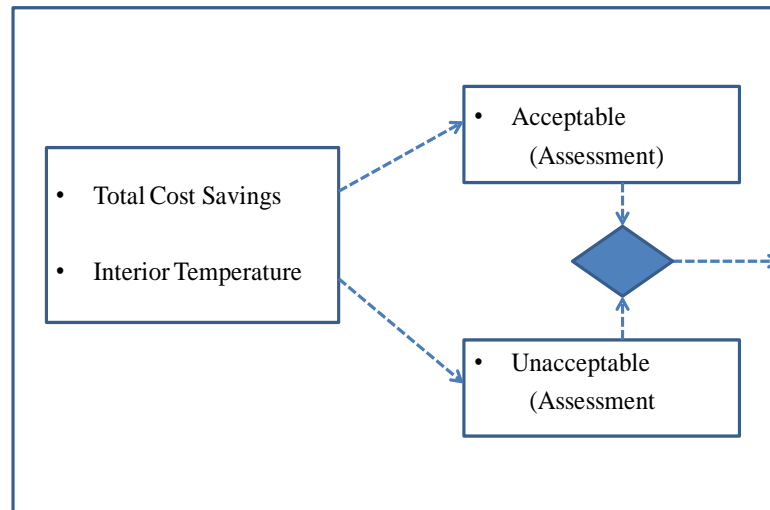


Figure III-6: Decision-based Process for the Balancing of Economic and Comfort Preferences

The inputs into the balance consumer preferences and finances model are the outputs from the previously discussed models: ETHER and Financial Impact. The output of the ETHER models is the interior temperature of the residence which is predicted over a 6 hour forecast horizon based on load profiles, exterior influences and outside temperature and time of day conditions. The forecast of the interior temperature will aid the intelligent framework in determining how the reduction in certain load profiles impacts the thermal comfort levels of the occupant. The output of the financial impact model is the total cost savings based on the reduction of energy usage while factoring in the electricity price during the time interval specified for reduction. If the financial and thermal comfort conditions are acceptable, instructions are sent to the local appliances (i.e., HVAC, washer and dryer, dishwasher, etc.) to be reduced or shutdown. If the conditions are unacceptable, a new demand response scenario is requested for consideration.

Combining the functionality of all the internal models presented in the IDEAS system increases the likelihood of energy reduction to meet the demands of the consumer and the utility. This system provides an autonomous decision-based framework to achieve the reduction of demand and energy costs and maintain thermal comfort levels for the consumer.

IV. Data Sources

Over three year's worth of data (i.e., interior temperature, exterior temperature, power consumption, humidity, CO₂, work flow rates and solar hot water temperature, etc.) was collected from the ecoMOD 1 prototype house; however, only one year's worth of data (i.e., 2009 data) is used for model development and evaluation purposes. This is primarily due to the excessive number of data points that were missing from the sensor recordings. In addition, some recorded data points were outside of the protocol range for the specified sensors; this was an indicator that the data was not accurate in nature and perhaps the sensors were malfunctioning.

The observational data used to construct the multivariate time series model is stored in the Energy Monitoring database on a FTP server via the internet. The data can be extracted as a CSV file or a Tab delimited file; for the purposes of this research, we chose to extract the data as a CSV file which obviates any transformational procedures for the purpose of data preparation. Table IV-1 shows the orientation of the CSV file. Each column of data is labeled based on the reporting sensor and the rows correspond to a particular time, observed in the leftmost column [23].

Table IV-1: Data Collection CSV File Orientation

<u>Date & Time</u>	<u>T1</u>	<u>T2</u>	<u>T3</u>	<u>T4</u>	<u>T5</u>	<u>T6</u>
1/1/2009 0:00	73.6	66.3	65.2	73.3	70.6	71.0
1/1/2009 0:01	73.6	66.3	65.3	73.2	70.6	71.0
1/1/2009 0:02	73.6	66.3	65.2	73.3	70.6	70.9
1/1/2009 0:03	73.6	66.3	65.2	73.2	70.6	71.0
1/1/2009 0:04	73.7	66.3	65.2	73.2	70.6	70.9
1/1/2009 0:05	73.7	66.2	65.2	73.5	70.6	70.9
1/1/2009 0:06	73.7	66.3	65.2	73.6	70.6	71.0
1/1/2009 0:07	73.8	66.3	65.2	73.5	70.6	70.9

The frequency of data collection can be recorded on a minute to minute or hourly basis. For the purpose of our research, we used the minute data to primarily characterize the cycling activity of certain load appliances (e.g., HVAC compressor). Since the setpoint swing or deadband on the thermostat within the residence has a narrow band of variation, the HVAC tends to cycle more

frequently from an on to off status to accommodate the comfort level of the consumer. Understanding the operational behavior of sizeable loads within the home can provide insightful information and identify sensible factors that play an important role in determining which loads greatly impact the interior conditions of the residence. Consequently, one of the goals of this research is to determine which load appliances highly impact the interior temperature based on pattern recognition so that optimal settings can be achieved for these controlled independent variables.

The output or dependent variable desired in the Evaluation of Thermal Response (ETHER) models is the interior temperature. The interior temperature variable is a composite function of the interior temperature sensors which assist in providing insight into the interior conditions within the residence. The inputs or independent variables used in the ETHER models are exterior temperature, power consumption (i.e. receptacles, lighting, appliances, etc.), humidity, CO₂, work flow rates and solar hot water temperature. At this time, we will take a detailed look into each data source parameter to get a better understanding of the type of data that needs to be transferred into information for the purposes of this research.

A. Interior and Exterior Temperature Data

There are fourteen interior temperature sensors used for measurement collection of the interior temperature (shown in Table IV-2). These measurements provide information to evaluate the design of the house's envelope in terms of energy efficiency [23]. Through the use of multiple sensors, we are able to determine how much variation in air temperature exists throughout the house. Table IV-2 shows the interior temperature sensor identification numbers and descriptions. The interior temperatures are measured in degrees Fahrenheit (°F). The temperature span for the interior sensors is 40°F to 90°F.

From our preliminary analysis and cross-referencing previous research [42] conducted on the ecoMOD 1 house, it was determined that 13 interior temperature sensors are highly correlated with one another. For instance, Sensors T1 through T10 and Sensors T12 through T14 are all highly correlated.

Table IV-2: Interior Temperature Sensor IDs and Sensor Description

Interior Temperature Sensors (°F)			
<u>ID</u>	<u>Description</u>	<u>ID</u>	<u>Description</u>
T1	1 st floor next to bay window	T8	2 nd floor bathroom
T2	1 st floor near kitchen area	T9	2 nd floor middle bedroom
T3	1 st floor beside front entrance	T10	2 nd floor corridor
T4	1 st floor above kitchen cabinets	T11	Landing (Stairway) Skylight
T5	1 st floor in laundry/bath area	T12	2 nd floor east bedroom
T6	2 nd floor master bedroom in far corner	T13	1 st floor at the base of the stairs
T7	2 nd floor master bedroom behind door	T14	2 nd floor at the top of the stairs

There are four exterior temperature sensors used for measurement collection of the exterior temperature conditions. Table IV-3 shows the exterior temperature sensor identification numbers and descriptions. The four sensors are oriented on the four facets of the residence. For our analysis, Sensor T21 is used as the true exterior temperature measurement. Sensors T18 through T20 tend to be impacted by solar loading and other environmental conditions which decreases the reliability for obtaining true measurements. The exterior temperatures are measured in degrees Fahrenheit (°F). The temperature span for the exterior sensors is 0°F to 120°F.

Table IV-3: Exterior Temperature Sensor IDs and Sensor Description

Exterior Temperature Sensors (°F)	
<u>ID</u>	<u>Description</u>
T18	West wall next to bay window
T19	East wall on 1 st story porch floor
T20	South wall on galvalumn surface
T21	North wall near 2 nd story porch

B. Power Consumption Data

There are fourteen power consumption sensors used for measurement collection (shown in Table IV-4). Each sensor is connected to an element (e.g., receptacles, appliances) in the main circuit breaker panel. These sensors are used to determine the approximate electrical consumption for the various loads within the residence. The power consumption of the loads is measured in Watts. One of the benefits in collecting power consumption data is to determine if the ecoMOD 1 house is achieving its goal of reducing residential energy consumption. Research has been previously conducted to investigate if the ecoMOD 1 house is more energy efficient than a typical residence [42].

Table IV-4: Power Consumption Sensor IDs and Sensor Description

Power Consumption Sensors (Watts)			
<u>ID</u>	<u>Description</u>	<u>ID</u>	<u>Description</u>
CT1	2 nd floor lighting and receptacles	CT8	Conventional Hot Water heater
CT2	Water Systems and Fieldpoint Power	CT9	1 st floor lighting and receptacles
CT3	1 st floor lighting and receptacles	CT10	Range Hood and Dishwasher
CT4	Kitchen receptacles and microwave	CT11	1 st floor lighting and receptacles
CT5	2 nd floor bathroom GFI receptacles	CT12	Refrigerator
CT6	Washer and Dryer	CT13	HVAC Outdoor (Compressor)
CT7	Range (Stove)	CT14	HVAC Indoor (Air Handler)

From preliminary analysis of the data, it is shown in Figure IV-1 that the maximum top 4 power consumption loads are the HVAC air handler, conventional hot water heater, range (stove) and washer and dryer. Due to this finding, we would expect that the top 4 power consumption loads would be identified in the structure of the ETHER models as sensible factors that play an important role in impacting the interior conditions of the residence. With the identification of the specific load appliances in the structure of the ETHER models, the intelligent framework can explicitly target to achieve optimal settings of these appliances in response to the utility signals.

In previous research, the HVAC compressor and air handler sensor values were combined for analysis purposes. For the purposes of our research, we decided to evaluate each appliance individually, initially, to later be combined to be referenced as a HVAC component. The power consumption readings shown in Figure IV-1 gives an illustration of the maximum amount of power that was consumed by each appliance; however, it does not give a depiction of the total annual energy use for each appliance; this is due to the excessive number of missing values.

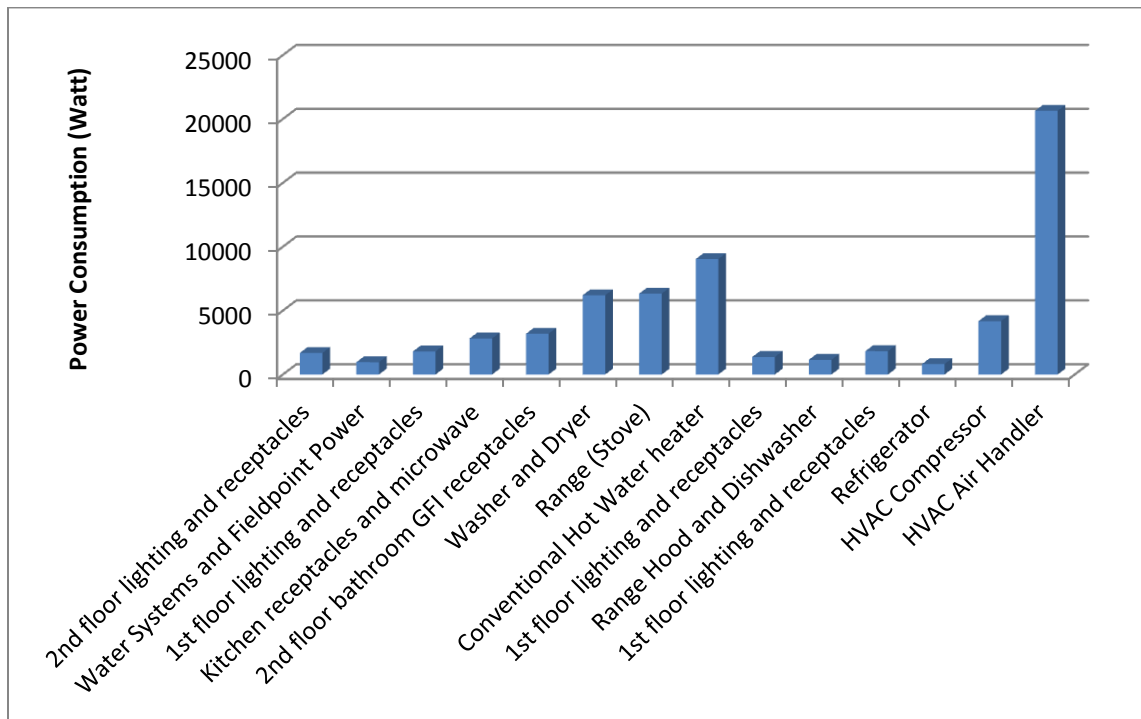


Figure IV-1: Top 4 Power Consumption Load Appliances

C. Humidity Data

There are three humidity sensors used for measurement collection at the ecoMOD1 house; two are located indoors (i.e., one is on the main level of the residence and the second one is on the 2nd level) and one is located outside of the home (shown in Table IV-5). The human comfort level is greatly influenced by the humidity level of air. Depending on the level of humidity, the same temperature setting may feel hotter or cooler for higher and lower humidity levels, respectively. Thus, it is important to be able to track, and potentially, control humidity within actively

conditioned indoor spaces. The humidity is measured in percentage (%). The protocol humidity span for the sensors is 0% to 100%.

Table IV-5: Humidity Sensor IDs and Sensor Description

Humidity (%)	
<u>ID</u>	<u>Description</u>
T15	1 st floor at the base of the stairs
T16	2 nd floor at the top of the stairs
T17	North wall near 2 nd story porch

From preliminary analysis of the humidity data, it is shown in Table IV-6 that the indoor humidity sensors (i.e., Sensors T15 and T16) are highly correlated with one another. For the purposes of our research, a high correlation amongst variables is defined as a resultant correlation value of .85 or higher. From this finding, we can create a composite function for Sensor T15 and Sensor T16 to represent the interior humidity within the residence. The sig (2-tail) value represents the significance factor of the correlation and the N represents the total number of observations for each variable.

Table IV-6: Correlation Values for the Humidity Sensors

		Correlations		
		T15	T16	T17
T15	Pearson Correlation	1	.925**	.650**
	Sig. (2-tailed)		.000	.000
	N	48240	48240	48240
T16	Pearson Correlation	.925**	1	.549**
	Sig. (2-tailed)	.000		.000
	N	48240	48240	48240
T17	Pearson Correlation	.650**	.549**	1
	Sig. (2-tailed)	.000	.000	
	N	48240	48240	48240

** . Correlation is significant at the 0.01 level (2-tailed).

D. Carbon Dioxide (CO₂) Data

There are two carbon dioxide (CO₂) sensors used for measurement collection at the ecoMOD1 house (shown in Table IV-7). One sensor is located on the main level of the residence and the second sensor is located on the 2nd level. Carbon dioxide levels are largely dependent on the number of people breathing in a particular space. The carbon dioxide is measured in parts per million (ppm). The protocol carbon dioxide span for the sensors is 0-2000 ppm.

Table IV-7: CO₂ Sensor IDs and Sensor Description

CO ₂ (ppm)	
<u>ID</u>	<u>Description</u>
T22	1 st floor near bay window
T23	2 nd floor corridor

For our analysis purposes, the CO₂ data provides insight into the number of occupants that are in the residence during different times of the day. Carbon dioxide can also provide a reasonable measure of air quality. It is generally accepted that when levels exceed approximately 1000 parts per million (ppm), additional ventilation is required. Figure IV-2 illustrates the carbon dioxide readings within the residence during the morning hours for the month of February.

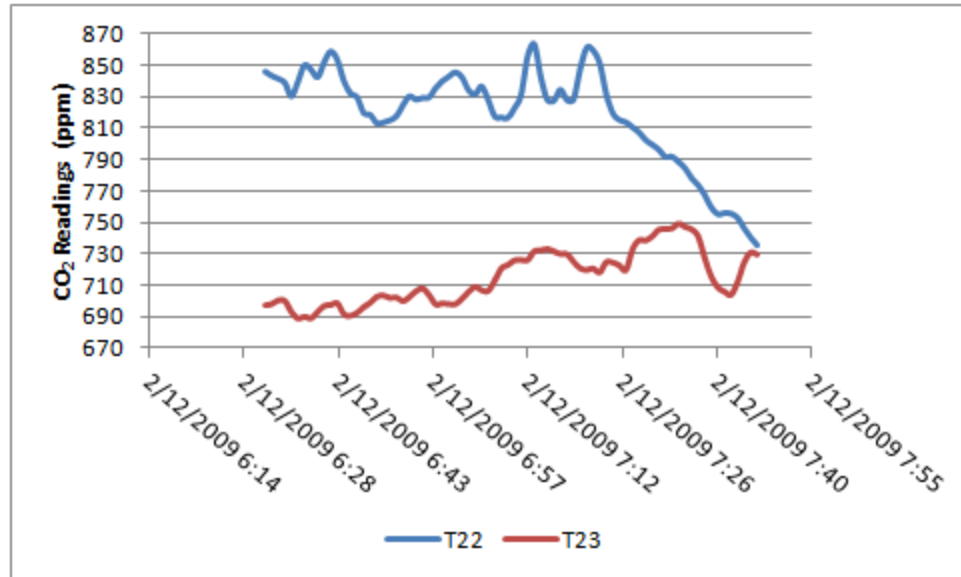


Figure IV-2: CO₂ Readings during the morning hours for the month of February

E. Water Flow Data

There are four water flow sensors used for measurement collection at the ecoMOD1 house (shown in Table IV-8). The water flow data is collected by sampling the status of the flow meter's reed switches. The reed switch opens and closes for every gallon of water that passes through the meter. The measurements are binary (i.e., 0 or 1) in value, which represent a gallon of water usage. Due to the nature of the data recorded for the water flow rate, an engineering decision was made that the water rate sensor data was not a sensible factor and provided minimal to no influence to the rate of change of the interior temperature.

Table IV-8: Water Flow Sensor IDs and Sensor Description

Water Flow (gal)	
<u>ID</u>	<u>Description</u>
F1	Rainwater
F2	City Water
F3	Cold Water to Solar Heat Exchange Tank
F4	Solar Loop

F. Solar Hot Water Temperature Data

There are four solar hot water temperature sensors used for measurement collection at the ecoMOD1 house (shown in Table IV-9). The solar hot water heating system functions as a pre-heat for the conventional electric water heater, reducing the amount of energy needed to heat water in the house. Two of the solar hot water temperature sensors are highly correlated with one another. To reduce the redundancy of information, a composite function is computed between the sensors. The solar hot water temperature is measured in degrees Fahrenheit (°F). With further analysis, it became apparent that the solar hot water temperature recordings were not sensible factors and provided minimal to no influence to the rate of change of the interior temperature.

Table IV-9: Solar Hot Water Temperature Sensor IDs and Sensor Description

Solar Hot Water Temperature (°F)	
<u>ID</u>	<u>Description</u>
T24	City Cold Water Input
T25	Heated Water Output to Conventional Tank
T26	Solar Loop Out of Heat Exchange Tank
T27	Solar Loop Return to Heat Exchange Tank

Through detailed analysis and observation, it can be observed that the model output (i.e., interior temperature) is influenced by changes in the recorded model parameters (i.e., data sources). However, there are some immeasurable factors not listed in this chapter, including open windows or doors, electric fans for cooling, electric heaters for heating, etc., that could also provide some influence to the rate of change of the interior temperature. As a result, mapping the relationship of the interior temperature to its influence factors is a key component in achieving model accuracy of prediction.

V. Variable Dependence

Variable dependence pertains to understanding the relationship between the dependent and independent variables. In this section, we will investigate the importance of understanding the type of data collected from the residence and how to turn this data into information to increase model predictions. Accuracy of model predictions can improve decision making on behalf of the consumer by having them select which appliance settings to reduce, shut down or reschedule in response to price variation or request for reduction signals.

The eight Evaluation of Thermal Response (ETHER) models used for forecasting are developed by observing the operational behavior inside the residence based on the outside temperature and time of day conditions over a full year; seasonal influence has been taken into consideration. The structure of each model is made up of both controllable and uncontrollable variables. The controllable variables are all the load components (i.e., lights, electronic devices, HVAC system, range (stove), microwave, etc.). Examples of uncontrollable variables are exterior temperature, exterior humidity, etc. Different identified variables in each model structure represent the loads that greatly influence the interior temperature conditions based on activity patterns and its respective statistical significance. As a result, each modeled system can be investigated so that appropriate modifications can be made to achieve the optimal settings for the controlled independent variables. The premise of establishing optimal settings is to allow for energy reduction during peak demand periods. Several experiments have been set up with observation data from the residence to test the accuracy of the forecasting models under varying conditions (i.e., different times of the day, environmental conditions, seasonal influence, modification of controlled load consumptions, etc.).

Normally, in a demand response program, the utility looks to reduce energy consumption of the sizeable loads in the residence (i.e., HVAC, Hot water heater, etc.). However, having the insight into which loads impact the interior conditions can help to pinpoint which loads to decrease

during demand response periods. The capability to predict interior conditions within the residence will allow optimal settings to be selected for the load appliances based on the contributing controllable variables. The methodology used for model development identifies the significant controlled variables that highly impact the interior temperature. In addition, the interior temperature is the primary dependent variable in this research; thus, having the appropriate method to assess the interior conditions can definitely affect the quality of the model's prediction accuracy.

There are four variable dependence approaches that were evaluated to increase the model prediction accuracy in our research, specifically targeting the daytime models. Figure V-1 shows a conceptual design of variable dependence approaches. Each daytime model is evaluated using all four of the variable dependence approaches.

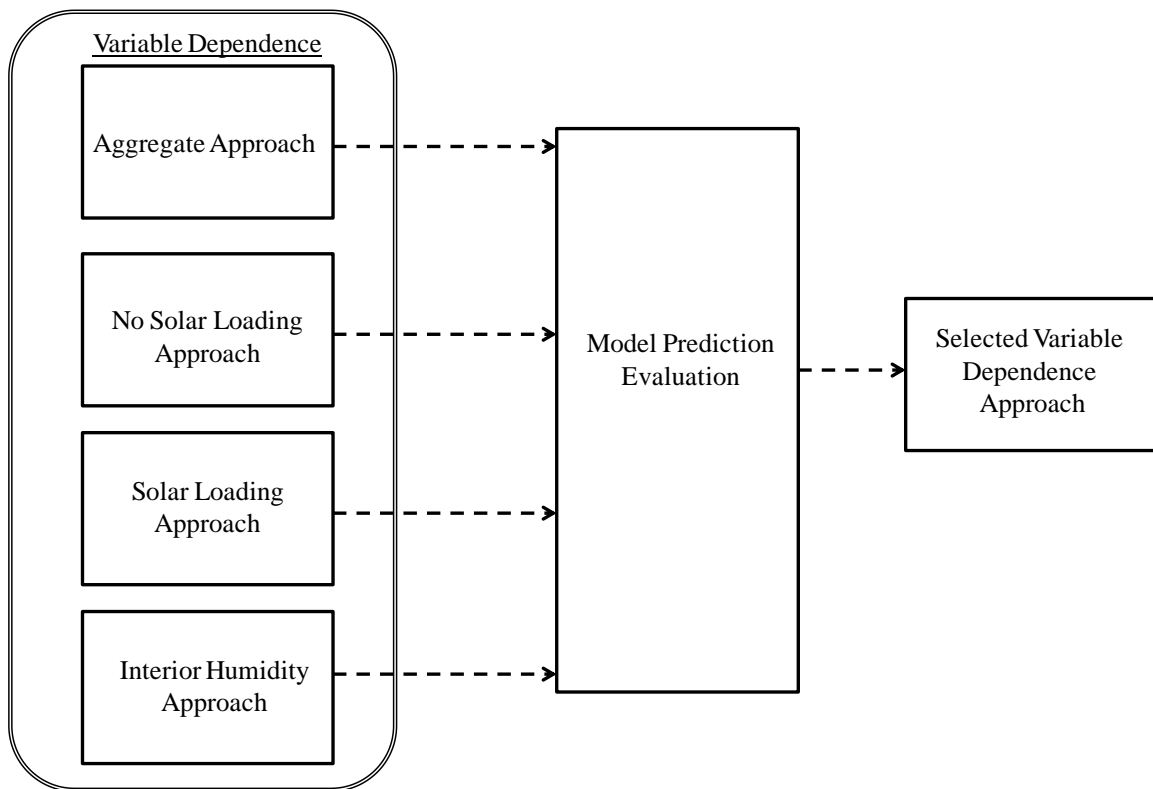


Figure V-1: Conceptual Design of the Variable Dependence Approaches

Due to the affect of the sunlight during the daytime hours, the model accuracy seemed to decline compared to the nighttime models. Thus, different approaches were chosen to manipulate the assessment of the interior temperature and as a result, there were variations in the accuracy of predictions that were obtained. The approach that produces the best prediction accuracy results is selected as the primary approach to be used for further model development and forecasting. Example results for each variable dependence approach used for the Day Evaluation of Thermal Response (ETHER) models are shown.

A. Aggregate Approach

The aggregate approach consists of creating a composite function for the interior temperature variable with all fourteen interior temperature sensors (i.e., Sensor T1-T14), shown in Table V-1. In other words, all fourteen interior temperature sensors are averaged together to create the IntTemp variable that is used for model development and forecasting.

Table V-1: Aggregate Approach – All Fourteen Interior Temperature Sensors

T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12	T13	T14	IntTemp
75.9	72.2	72.8	75.5	75.4	74.8	75.7	74.9	79.2	77.2	92.3	74.2	74.0	77.9	76.6
75.9	72.3	72.9	75.4	75.5	74.8	75.8	74.8	79.2	77.3	92.1	74.2	74.1	77.8	76.6
75.9	72.3	72.9	75.5	75.4	74.7	75.8	74.8	79.2	77.3	92.2	74.2	74.0	78.0	76.6
75.9	72.3	73.0	75.5	75.4	74.8	75.8	74.8	79.2	77.3	92.2	74.1	74.0	77.8	76.6
75.9	72.4	73.0	75.5	75.4	74.7	75.8	74.7	79.1	77.3	92.2	74.2	74.0	78.0	76.6

From Table V-1, it can be observed that Sensor T11 is reporting a temperature reading that is approximately 16 degrees F higher than the maximum temperature reading reported by the other interior temperature sensors. An assumption was made to consider Sensor T11 readings as the true solar loading influence into the interior envelope of the residence. For that reason, this approach aims to evaluate whether the solar loading information combined with the other thirteen interior temperature variables can increase the model prediction accuracy. Figure V-2 shows the evaluation results of the Day Models using the Aggregate Approach.

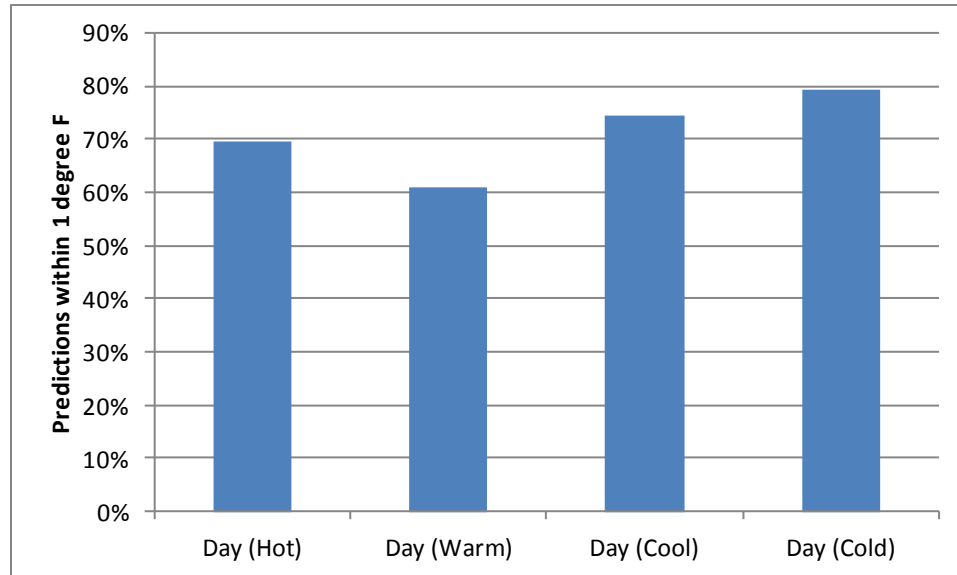


Figure V-2: Aggregate Approach – Evaluation Results of the Day Models

As observed, the Day (Warm) model performed below the level of prediction accuracy of the remaining models. Typically, prediction measurements are within 2 degree F of measurements in literature, resulting in 80% or higher for prediction accuracy. As a method of simplification used within this study, missing values or values that were recorded outside of the protocol range of the sensors are removed from the datasets. Thus, an assumption was made that with predictions within 1 degree F of measurements, a prediction accuracy of more than 65% is desired for all resultant models.

B. No Solar Loading Approach

The interior temperature variable is a composite function of thirteen interior temperature sensors (i.e., T1-T14, excluding T11) due to high correlation amongst the sensors. Table V-2 shows the individual interior temperature sensors excluding Sensor T11 to calculate the IntTemp variable.

Table V-2: No Solar Loading Approach – Excluding Sensor T11

T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T12	T13	T14	IntTemp
75.9	72.2	72.8	75.5	75.4	74.8	75.7	74.9	79.2	77.2	74.2	74.0	77.9	75.4
75.9	72.3	72.9	75.4	75.5	74.8	75.8	74.8	79.2	77.3	74.2	74.1	77.8	75.4
75.9	72.3	72.9	75.5	75.4	74.7	75.8	74.8	79.2	77.3	74.2	74.0	78.0	75.4
75.9	72.3	73.0	75.5	75.4	74.8	75.8	74.8	79.2	77.3	74.1	74.0	77.8	75.4
75.9	72.4	73.0	75.5	75.4	74.7	75.8	74.7	79.1	77.3	74.2	74.0	78.0	75.4

As observed in the Aggregate Approach, Sensor T11 is highly impacted by solar loading conditions and is excluded from the data analysis process in the No Solar Loading Approach due to its discrepancies in interior temperature readings. Therefore, Sensor T11 is not factored in a dependent variable or an independent variable. In addition, excluding Sensor T11 results in an IntTemp variable that is approximately 1 degree F lower in magnitude. The aim of this approach is to evaluate if excluding the solar loading information from the IntTemp composite function can increase the model prediction accuracy. Figure V-3 shows the evaluation results of the Day Models using the No Solar Loading Approach.

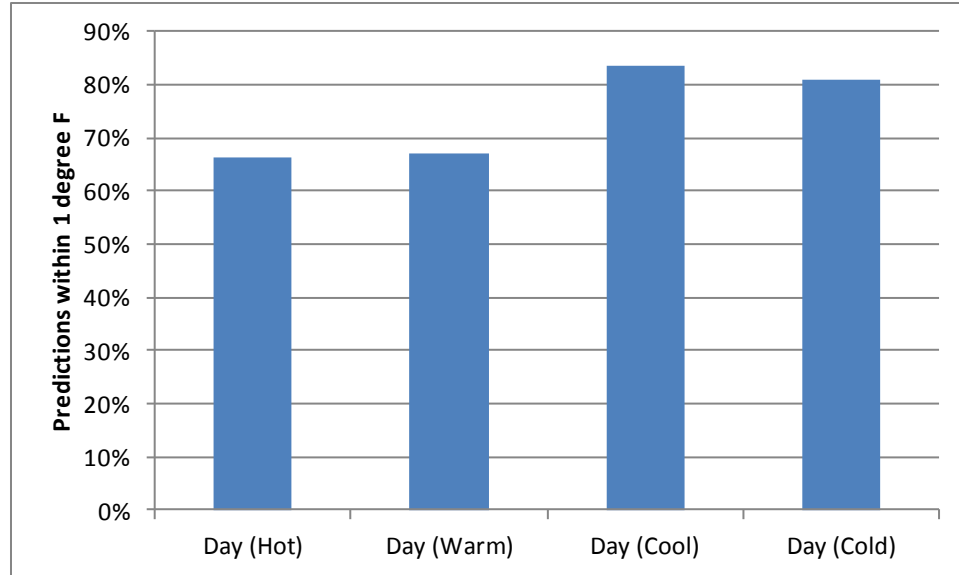


Figure V-3: No Solar Loading Approach – Evaluation Results of the Day Models

All of the resultant models produced a prediction accuracy of more than 65%. Compared to the results obtained with the Aggregate Approach, the Day (Hot) model accuracy reduced by approximately 3%; however, the remaining models had an increase in model prediction accuracy by more than 9%. Based on the results produced by using the No Solar Loading approach, this approach is used for all further model development, test and evaluation and data analysis procedures.

C. Solar Loading Approach

The interior temperature is a composite function of thirteen interior temperature sensors (i.e., T1-T14, excluding T11), where Sensor T11 is evaluated as an independent variable. Intrinsicly, the Solar Loading Approach uses the premise of the No Solar Loading Approach with the same calculation of the IntTemp variable; however, instead of obviating Sensor T11, this approach tests its influence of Sensor T11 as an independent variable. Figure V-4 shows the evaluation results of the Day Models using the Solar Loading Approach.

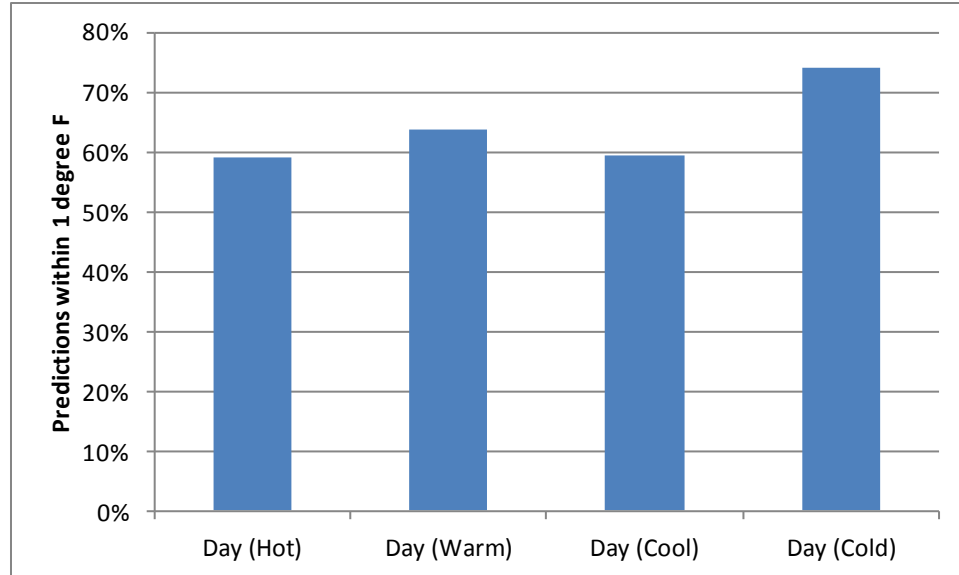


Figure V-4: Solar Loading Approach – Evaluation Results of the Day Models

The aim of this approach is to examine if the solar loading information has an impact on the dependent variable series to increase the prediction accuracy of the ETHER models. As observed in Figure V-4, there is a declining effect in the model prediction accuracy compared to the resultant models produced in the Aggregate Approach and the No Solar Loading Approach. As a result, we can conclude that the Solar Loading Approach may not be the best method to use in an effort to increase model prediction.

D. Interior Humidity Approach

In the Interior Humidity Approach, the composite interior humidity (i.e., Sensor T15 and Sensor T16) and the composite interior temperature variable (i.e., T1-T14, excluding T11) are both evaluated as dependent variables. Due to the identified assessment of the various approaches above, Sensor T11 is not being considered as a dependent variable or an independent variable. Therefore, the approach uses the basis of the No Solar Loading Approach to calculate the IntTemp variable. In addition, due to the high correlation between Sensor T15 and Sensor T16, the interior humidity is calculated by taking the average of these two sensors to create a

composite function. The aim of this approach is to observe the rate of change of the interior temperature conditions with having the interior humidity as an output parameter versus an input parameter. Figure V-5 shows the evaluation results of the Day Models using the Interior Humidity Approach.

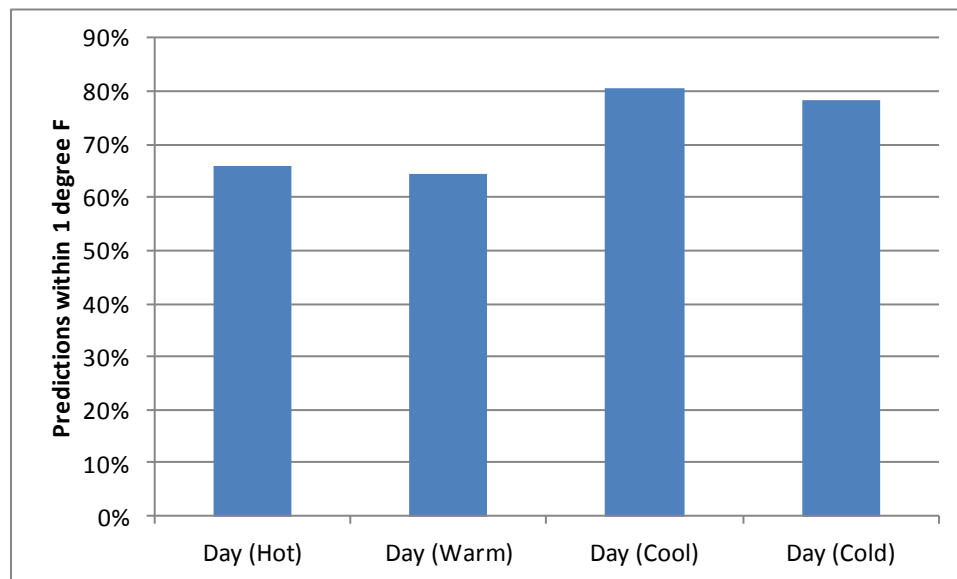


Figure V-5: Interior Humidity Approach - Evaluation Results of the Day Models

The interior humidity is a factor that impacts the comfort temperature of an individual or a group [43]. The combination of temperature and humidity where occupants report their level of comfort is known as the comfort zone. In the previous mentioned approaches, the interior humidity has been evaluated as an independent variable; however, in the Interior Humidity Approach both interior temperature and interior humidity has been examined to assess the model prediction results. All of the resultant models, except for the Day (Warm) models, produced a prediction accuracy of more than 65%. As observed in Figure V-6, the Interior Humidity Approach produces comparable results to the prediction accuracy of the No Solar Loading Approach; however, the No Solar Loading Approach still out-performs the other variable dependence approaches.

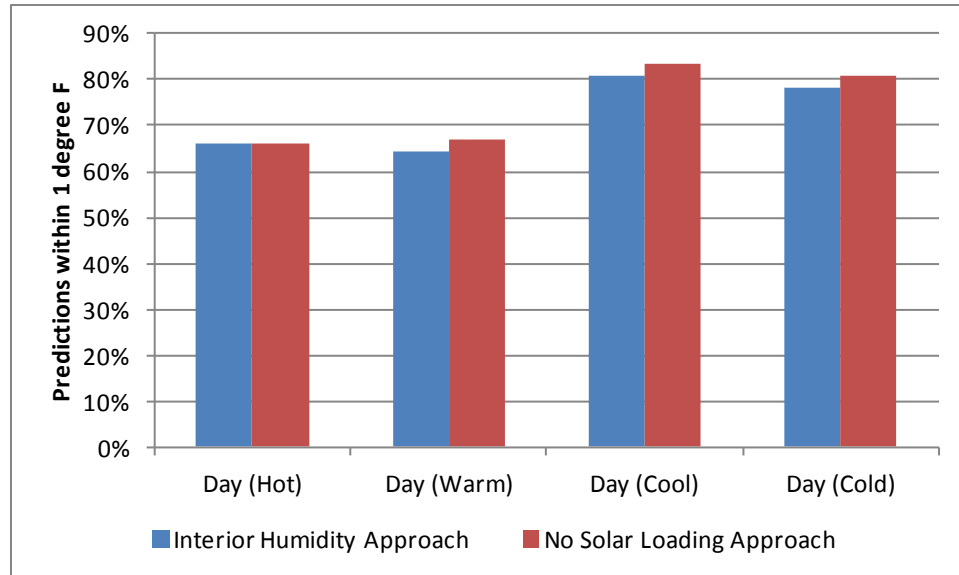


Figure V-6: Comparison between the Interior Humidity Approach vs. the No Solar Loading Approach

The evaluation of all four variable dependence approaches, shown in Figure V-7, contributed to mapping the relationships between the dependent and independent variables while manipulating the effect of the solar loading conditions.

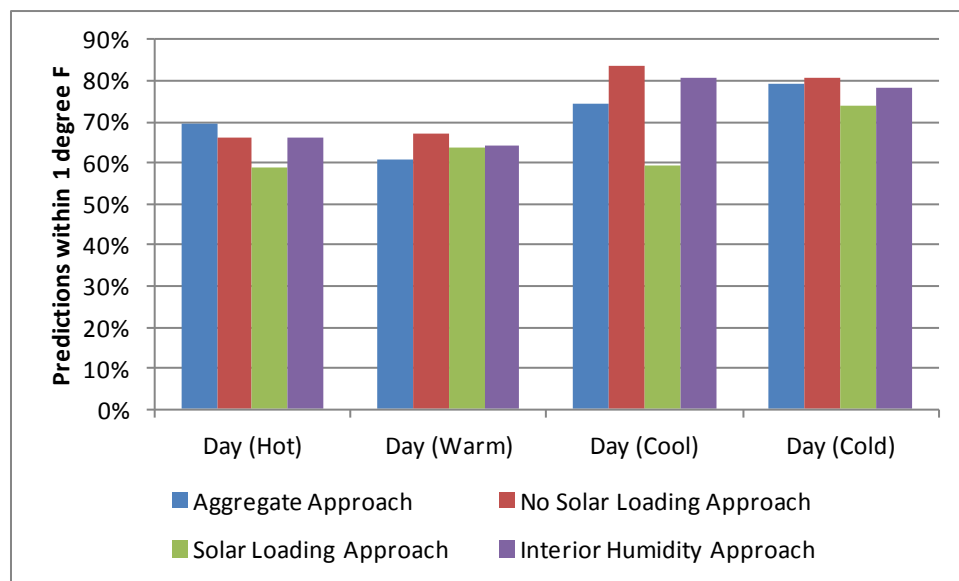


Figure V-7: Comparison between all Four Variable Dependence Approaches

As a result, the No Solar Loading Approach produced the best model prediction results of the interior temperature conditions out of the four variable dependence approaches that were examined. Now that an approach has been selected for further model development and evaluation, we will look at the research methodology used to develop each model.

VI. Research Methodology

This chapter describes the methodology for developing the Autoregressive Integrated Moving Average (ARIMA) time series Evaluation of Thermal Response (ETHER) models used to predict interior conditions in an effort to achieve demand response within a residence. These models are based on time series analysis and provide reliable and accurate forecasts of interior temperatures. A good forecast of the interior temperatures can better equip an intelligent framework to choose the appropriate demand response scenario to minimize energy consumption without impacting thermal comfort levels of the consumer.

Time series analysis focuses on determining future values based on known observations. A time series is a set of observations generated sequentially in time which are typically measured in uniform time intervals [42], [43], [21]. For instance, Khasnabis et al. [42] developed a multivariate time series model using observational data to predict power ramp rates of wind farm power over 10-minute intervals. Many models, detailed in literature, were built using historical observation data through time series analysis. Yet, these particular models were focused on prediction of wind speeds, electricity price and load forecasting, pattern recognition and power ramp rates. Based upon the forecasting capability observed from these models, our research uses the observations of multiple parameters which were recorded sequentially over equal time intervals to develop the multivariate time series ETHER models.

A. Autoregressive Integrated Moving Average (ARIMA) Method

In our work, the time series of the thermal response of a residence is analyzed using the Autoregressive Integrated Moving Average (ARIMA) method. As documented in other literature, ARIMA methods are normally used to analyze time series due to their accuracy and mathematical soundness [20]. ARIMA represents a class of time series models that describes both stationary and non-stationary time series data where it combines several time series techniques

such as autoregressive models, differencing or integration, and moving average models to identify an acceptable model [44], [45]. Stationary time series data assumes statistical equilibrium that does not change over time which is fixed at a constant mean level and with constant variance. In contrast, non-stationary time series data has no constant mean level over time.

In general, an ARIMA model is fitted to time series data to gain a better understanding of the collected data or to predict future data points in the series [46]. The overall sequence of the ARIMA methodology [20] used for the prediction of the thermal response is organized into the following steps:

- 1) **Identify a hypothesis to be tested in order to develop a model** - The hypothesis in this research focuses on using data collected from a specially instrumented residence to develop a set of forecasting ETHER models to identify relationships or patterns of usage between independent variables that impact the interior conditions.
- 2) **Develop a model using observed data** – Initially, a set of eight ETHER models were developed using the collected data from the residence. The time series data was used to observe the relationships between the independent variables and the response variable.
- 3) **Estimate model parameters** – Model parameters were estimated and statistically significant independent variables were selected by using the Statistical Package for Social Sciences (SPSS) software application.
- 4) **Check validation of the hypothesis** – The hypothesis was validated by comparing the predicted interior temperature to the actual interior temperature conditions within the residence. The time series data was separated into a training dataset and a testing dataset; the fitted models were tested using the data from the testing dataset.
- 5) **Forecast using the developed model** – Using the SPSS application, the developed time series ETHER models were applied to new datasets for forecasting. In turn, the forecasting accuracy of the models was validated by obtaining the prediction error.

If the hypothesis is found to be untrue, the model should be refined and model parameters can be re-estimated. Using the ARIMA method, predictions of interior thermal response conditions are conducted over a short term forecast horizon (i.e., up to 6 hours) to support decision making opportunities in adjusting the electrical demand in response to utility signals.

The ARIMA model parameters (i.e., ARIMA (p,d,q)) are identified to provide insight into the number of autoregressive orders (p) in the model, the order of differencing (d) applied to the series before estimating models and the number of moving average orders (q) in the model. The autoregressive orders specify which previous values from the series are used to predict current values. For example, the autoregressive (AR) process of order p can be represented as [45]:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + a_t \quad (1)$$

y_t represents the response or dependent variable (i.e., interior temperature), a_t represents the model residuals (i.e., $y_t - y_{t-p}$), ϕ are parameter estimates. The differencing becomes necessary when trends are present in the data and is used to remove their effect; this method is used to obtain stationary time series data. By using the differencing operator, ∇ , the first order continuous difference can be expressed as [45]:

$$\nabla y_t = y_t - y_{t-1} = (1 - B)y_t \quad (2)$$

y_t represents the present interior temperature of the time series data, y_{t-1} represents the previous (i.e., t-1) time series interior temperature data and B is the backward shift operator. The moving average orders specify how deviations from the series mean for previous values are used to predict current values [46]. For example, the moving average (MA) process of order q can be represented as:

$$y_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} \quad (3)$$

y_t represents the interior temperature time series data, a_t represents the present model residual, a_{t-q} represents t-q time series model residuals and θ is the parameter estimate which in practice is estimated from the data.

To achieve greater flexibility in fitting the actual time series data, a combination of the autoregressive, differencing and moving average terms are advantageous to achieve an acceptable model [45]. Mathematically, the general form of the ARIMA model process is shown as a linear difference equation defined by (4), (5) and (6):

$$\varphi(B)y_t = \theta(B)a_t \quad (4)$$

where

$$\varphi(B) = \phi(B)\nabla^d = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p \quad (5)$$

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q \quad (6)$$

y_t represents the interior temperature time series data, a_t represents the model residuals (i.e., $y_t - y_{t-q}$), B represents the backward shift or lag operator defined by $B^m y_t = y_{t-m}$, ϕ and θ are parameter estimates, and ∇^d represents the differencing operator where d represents the differencing order (i.e., 1, 2, etc.) and $\nabla = 1 - B$ [32]. For example, an ARIMA (2,1,4) model would be represented as the following equation: $(1 - \phi_1 B - \phi_2 B^2)(1 - B)y_t = (1 - \theta_1 B - \theta_2 B^2 - \theta_3 B^3 - \theta_4 B^4)a_t$.

The above processes show the methodology for the dependent variable (i.e., interior temperature) time series data being affected by previous values of itself; however, the dependent variable can also be influenced by other time series data (i.e., independent variables). Moreover, each time series of the collected data from the residence is connected with each other. The input or independent variables (i.e., exterior temperature, humidity, power consumption, CO₂, etc.) used

in this research are represented as a transfer function model. The sequence of steps for developing an ARIMA transfer function model involves the same steps as described earlier: hypothesis, model development, parameter estimation, fitting check and forecasting [47].

The general form of the ARIMA method for the fitted resultant transfer function models described in [45], [47] is used for this research is defined in (7), (8) and (9):

$$Y_t = \delta^{-1}(B)\omega(B)X_t + N_t \quad (7)$$

where

$$N_t = \phi^{-1}(B)\theta(B)a_t \quad (8)$$

$$(1 - \delta_1 B - \delta_2 B^2 - \dots - \delta_r B^r)Y_t = (\omega_0 - \omega_1 B - \dots - \omega_s B^s)X_t \quad (9)$$

Y_t represents the interior temperature time series data, $\delta(B)$ represents the time series of the dependent variable, $\omega(B)$ represents the time series of independent variables, $\phi(B)$ represents the autoregressive operator, $\theta(B)$ represents the moving average operator, B represents the backward shift operator, N_t is represented as an ARIMA model, a_t represents the error term, X_t represents the input or independent variables.

B. ARIMA Model Parameters

The ARIMA Transfer Function method can be viewed as an effective and mature technique to identify the relevant independent variables that impact interior conditions within the residence. For model development, significant predictors (i.e., independent variables) are identified based on its influence of the time series data of the dependent variable. Using the SPSS software application, model parameter estimates are generated based on the time series data of the dependent and independent variable series. The software application uses maximum likelihood estimation to obtain all the estimated values of the parameters.

An example of the ARIMA model parameters and parameter estimates is shown in Table VI-1 which highlights the output of the dependent variable (i.e., IntTemp) and the identified independent variables (i.e., CT1, CT3, etc.).

Table VI-1: ARIMA Model Parameters and Parameter Estimates [35]

ARIMA Model Parameters					Estimate
IntTemp-Model_1	IntTemp	No Transformation	Difference		1
	CT1	No Transformation	Numerator	Lag 0	0.001
			Difference		1
	CT3	No Transformation	Numerator	Lag 0	0.001
			Difference		1
	CT4	No Transformation	Numerator	Lag 0	0
			Difference		1
			Denominator	Lag 1	-0.63
				Lag 2	-0.81
	CT6	No Transformation	Numerator	Lag 0	0
			Difference		1
	CT7	No Transformation	Numerator	Lag 0	-3.95E-05
			Denominator	Lag 1	1.105
				Lag 2	-0.699
			Difference		1
	T16	No Transformation	Numerator	Lag 0	0.008
				Lag 1	-0.002
			Difference		1
	T17	No Transformation	Numerator	Lag 0	-0.004
			Difference		1
	T21	No Transformation	Numerator	Lag 0	-0.008
			Difference		1

Understanding the output data (i.e., model structure, lag values, parameter estimates, etc.) is essential in obtaining the fitted ARIMA Transfer Function ETHER model and forecasting model equations. The sensor identification numbers discussed in Chapter IV are being used to represent the independent variables; each variable has its own model parameter estimate. The lag operator (i.e., Lag 0, 1 and 2) is equivalent to the backward shift operator which was discussed in the previous section. From evaluating the information presented in Table VI-1, the ARIMA Transfer Function model (i.e., ARIMA (0,1,2)) is identified as being the best fit model for the sample training dataset, where the numbers in the parentheses represent the autoregressive (AR) order

(p), the differencing (I) order (d) and the moving average (MA) order (q), (i.e. ARIMA (p,d,q)), respectively.

C. ARIMA Transfer Function ETHER models

In this section, we will take a look at the model structure for each model to better understand the rationale for the Day and Night models. Eight Evaluation of Thermal Response (ETHER) models are initially developed using observational data to predict interior conditions of the residence over a 6 hour forecast horizon. The ETHER models are needed to accurately represent the operational behavior of the residence for research and development purposes to include analysis and forecasting capabilities.

The dataset (i.e., one year's worth of data) used for this research was separated or partitioned to accommodate for solar loading and non-solar loading conditions. As a result, there are two sets of models formed: Day Models and Night Models. The Day models are developed based on the patterns of the observational time series data used to represent the actual performance of the residence conducted during the solar loading conditions of the day. During the daytime, the need and most importantly the use of artificial lighting is lessened. In addition, solar loading during the daytime hours can be viewed as a heating source for the interior envelope of the residence. The Night models are developed based on the same premise as the Day Models; however, targeting non-solar loading conditions. During the nighttime hours, artificial lighting becomes a necessity for a standard of living within the residence. Consequently, the developed models are reflective of the environmental conditions to provide an accurate evaluation of the interior conditions. In addition, the dataset used to develop each set of models factors in the seasonal influence for the entire year.

C.1. Rationale for Day & Night Model Techniques

There are four techniques used to separate the datasets into Day Models and Night Models based on the time of day conditions. All four of these techniques have been examined in detail and will be evaluated later in the dissertation.

C.1.1. Daylight Savings Time Technique

The first technique is known as the Daylight Savings Time technique. The baseline ETHER model development (i.e., eight models) and analysis were conducted using this technique. Figure VI-1 shows the conceptual orientation of the ETHER models for the Daylight Savings Time technique.

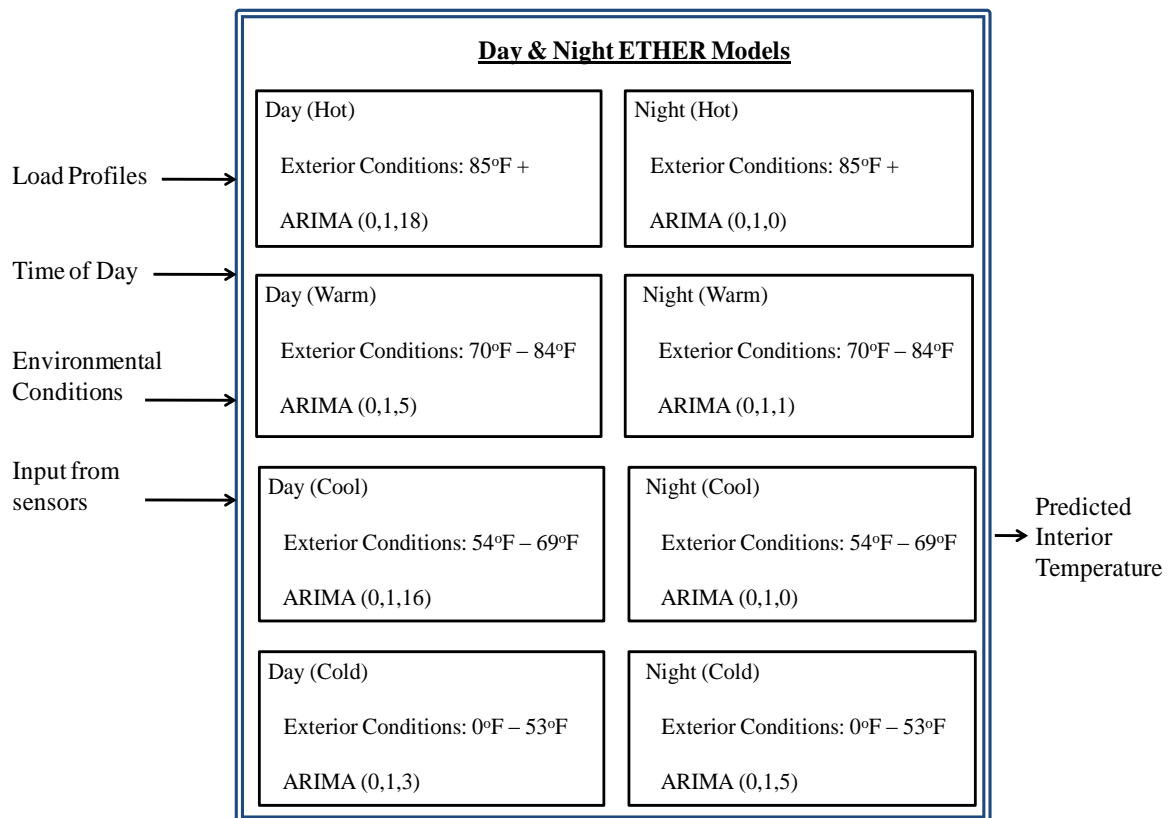


Figure VI-1: Conceptual Orientation of the Daylight Savings Time Technique

Daylight Savings Time is the practice of advancing the clocks forward one hour near the start of the spring season (e.g., March 13th) until the fall season (e.g., November 6th). The premise of Daylight Savings Time is that the evenings have more daylight and mornings have less. Daylight Savings Time allows the consumer to reduce evening usage of electricity and alter space heating and space cooling patterns.

For the Daylight Savings Time technique, an average sunrise and sunset time was estimated from the beginning until the end of the daylight savings timeframe. These times (i.e., Day (6:00 AM – 7:59 PM) and Night (8:00 PM – 5:59 AM)) were used to frame the time intervals for the Day Models and Night Models. Since Daylight Savings Time accounts for approximately nine months out of the year, an assumption was made to allow the selected time bounds of the Day and Night Models to be applied to the remaining three months (i.e., December, January and February). At the end of Daylight Savings Time, the period reverts back to standard time.

C.1.2. Equinox to Solstice Technique

The second technique is known as the Equinox to Solstice technique. The equinox is defined as the time when the sun crosses the celestial equator and the length of the daylight hours and nighttime hours are approximately the same [49]. The equinoxes happen twice a year; there is a spring or vernal equinox that occurs around March 20th and a fall or autumnal equinox that occurs around September 22nd. The astronomical definition of the solstice is when the sun has no apparent northward or southward motion, meaning that the sun stands still [49]. In other words, the solstice marks the conclusion of the increase or decrease in daylight hours depending on the time of the year. There is a summer solstice that occurs on June 21st and a winter solstice that occurs on December 21st.

The premise of the Equinox to Solstice technique is to split the year in half based on the changes in season and the variation in the length of the day. An average sunrise and sunset time were

estimated from the Spring Equinox to the Fall Equinox to account for the summer solstice (i.e., Day (6:15AM – 8:15PM) and Night (8:16PM - 6:14AM)) and from the Fall Equinox to the Spring Equinox to account for the winter solstice (i.e., Day (7:00AM – 5:59PM) and Night (6:00PM - 6:59AM)). These estimated times were used to frame the time intervals for the Day Models and Night Models.

A total of sixteen ETHER models are based on this technique. Figure VI-2 shows the conceptual orientation of the ETHER models for the equinox to solstice technique. As observed, an additional input (i.e. the date) is needed for the Day & Night ETHER models; otherwise, the model selection process as described earlier is the same. The date serves as the primary indicator for selection of which set of models should be used for forecasting and analysis purposes. For this research, the date range for the Spring Equinox to Fall Equinox models is March 20th through September 21st and the Fall Equinox to Spring Equinox models is September 22nd through March 19th.

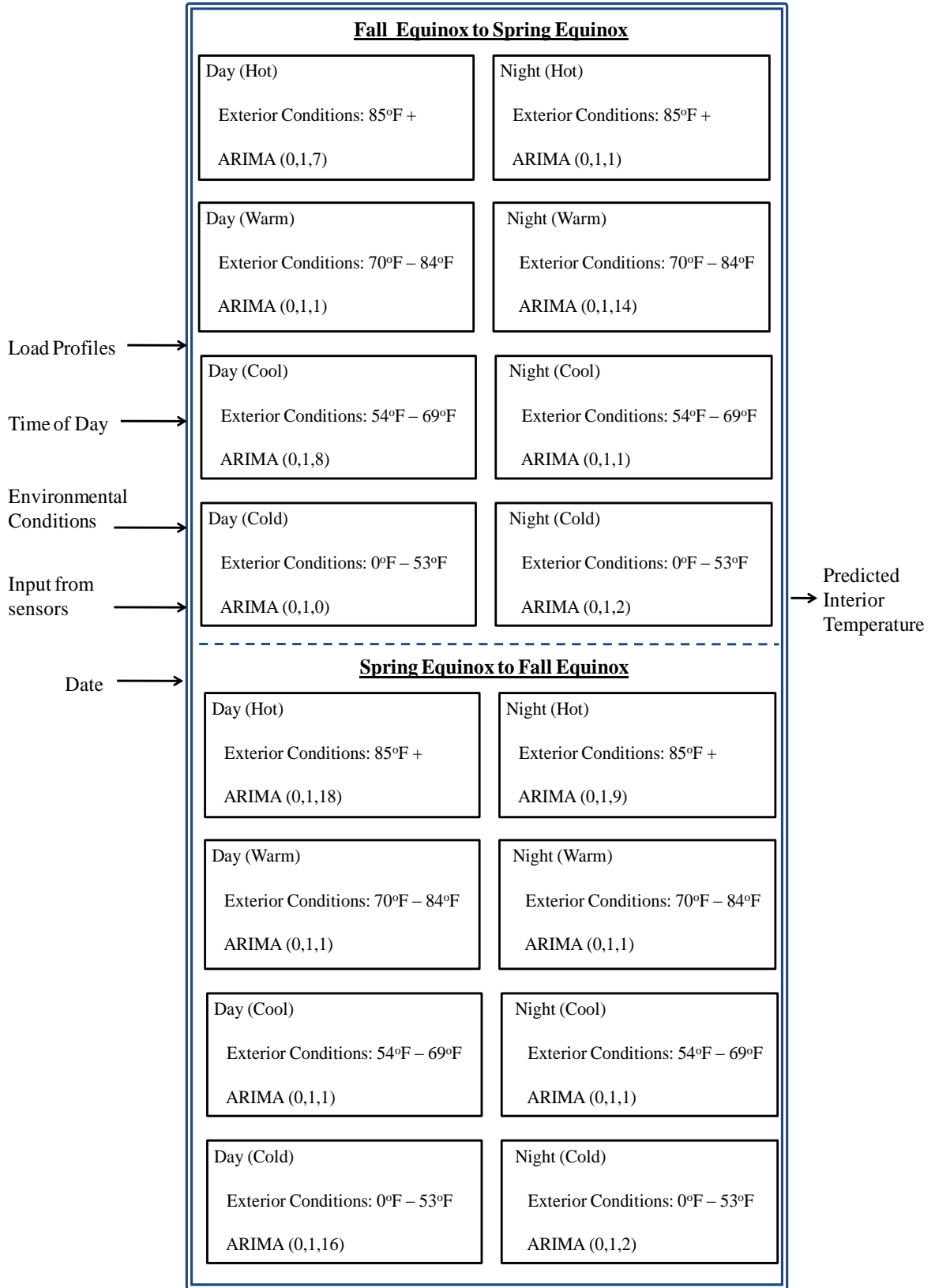


Figure VI-2: Conceptual Orientation of the Equinox to Solstice ETHER Model Technique

C.1.3. Seasonal Technique

The third technique is known as the Seasonal technique. This technique splits the year into quarters to accommodate for the four seasons (i.e., winter, spring, summer and fall). An average sunrise and sunset time were estimated for each season. For example, the Winter season (i.e., Day (7:00AM – 6:00PM) and Night (6:01PM - 6:59AM)) is from December 21st through March 19th, the Spring season (i.e., Day (6:00AM – 8:30PM) and Night (8:31PM - 5:59AM)) is from March 20th through June 20th, the Summer season (i.e., Day (6:10AM – 8:30PM) and Night (8:31PM - 6:09AM)) is from June 21st through September 21st and the Fall season (i.e., Day (7:00AM – 5:59PM) and Night (6:00PM - 6:59AM)) is from September 22nd through December 20th. These estimated times were used to frame the time intervals for the Day Models and Night Models.

There are a total of thirty-two ETHER models that are developed with this technique. Similar to the Equinox to Solstice technique, the Seasonal technique needs the date as an input into the system for the model selection process for the Day and Night ETHER models. Figure VI-3 highlights the conceptual orientation of the ETHER models for the Seasonal technique.

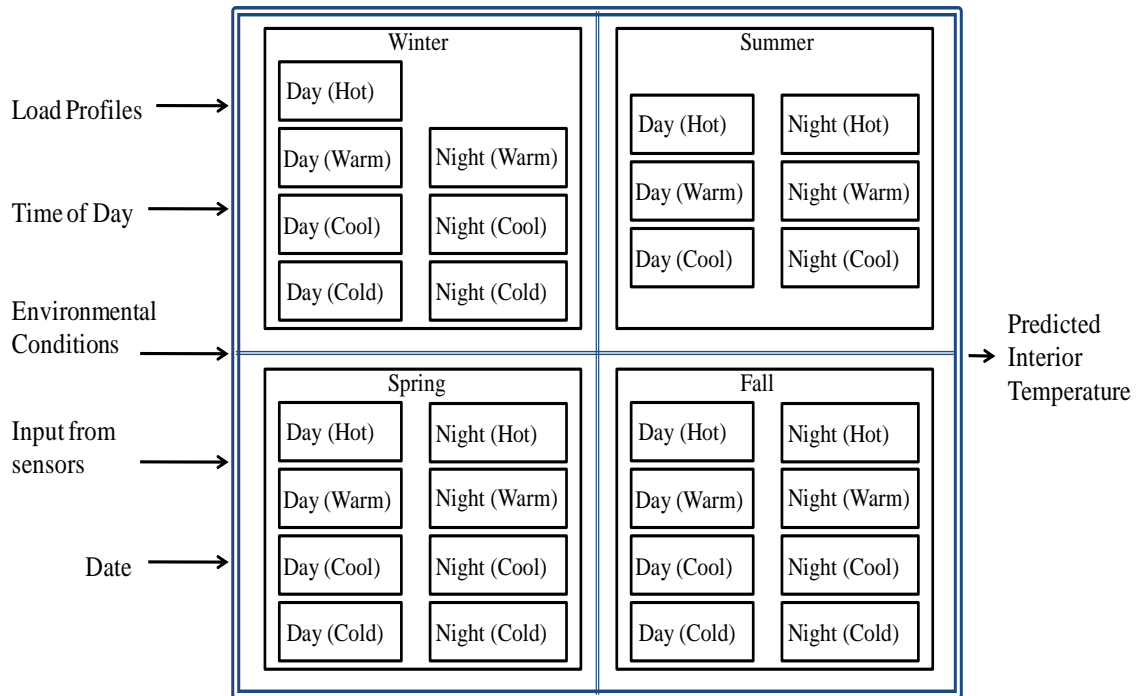


Figure VI-3: Conceptual Orientation of the Seasonal ETHER Model Technique

C.1.4. Entire Month Technique

The fourth technique is the Entire Month technique. This technique uses data from the previous month to predict the interior temperature of the residence for the subsequent month. Figure VI-4 provides a conceptual view for the Month to Month technique.

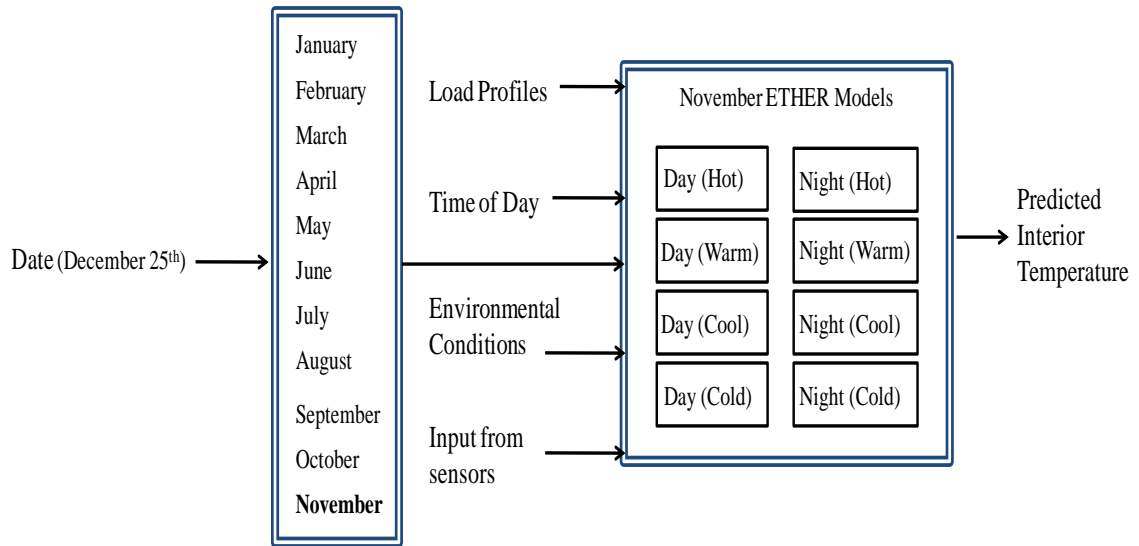


Figure VI-4: Conceptual Orientation of the Entire Month ETHER Technique

The average sunrise and sunset time were estimated for each month to separate the data into the appropriate time intervals to form the Day and Night ETHER models. Thirty days worth of data is being used for model development which significantly varies from the previously mentioned techniques. With the Entire Month technique, there are a total of seventy-eight ETHER models that were developed. An immediate challenge was identified with this technique which stems from the separation of the models based on outside temperature. In some cases, the previous month data does not contain the same exterior temperature ranges as the subsequent month which is needed for prediction. A potential solution would be to use the developed ETHER model with the closest outside temperature range. For evaluation purposes which will be shown later in the dissertation, the Entire Month technique using the outside temperature and time of day (OTTD) methodology will be compared to a Entire Month technique that was not separated based on the outside temperature and time of day conditions. In the next section, the rationale for the outside temperature conditions is discussed.

C.2. Rationale for Outside Temperature Conditions

As shown in Figure VI-2 above, each model is separated according to daytime and nighttime and outside temperature conditions. Environmental conditions (i.e., outside temperature) play a key role in understanding the operational behavior of certain equipment and appliances with the residence as well as understanding how the consumer responds to accommodate for their standard of living. The Energy Information Administration [14] reported the outside temperature condition is one of the many factors that affect residential energy consumption. For this reason, in addition to separating the dataset based on time of day conditions, the datasets were also separated based on outside temperature conditions. These outside temperature conditions have been categorized into four groups: Hot ($85^{\circ}\text{F} +$), Warm ($70^{\circ}\text{F} - 84^{\circ}\text{F}$), Cool ($54^{\circ}\text{F} - 69^{\circ}\text{F}$) and Cold ($0^{\circ}\text{F} - 53^{\circ}\text{F}$). The exterior temperature ranges were selected by observing the similarities of the operational behavior in the residence over an extended duration of time.

Our research methodology is centered on separating the ARIMA model structures based on time of day and exterior temperature conditions in an effort to identify controllable and uncontrollable variables that impact the interior conditions. By utilizing power consumption loads and other controllable parameters that consume a significant amount of energy, the multivariate time series prediction ETHER models can become a basis for predictive control aimed at optimizing load profiles to reduce and/or manage energy consumption during demand response periods. Generally speaking, the ARIMA Transfer Function ETHER models used in our research can be viewed as an effective technique to explain and clarify the mapping of the relationships between the interior temperature and its influence factors.

D. Case Study - Evaluation of Day & Night ETHER Model

Techniques

In this case study, an evaluation of the Day and Night ETHER Model techniques is investigated. The selected test cases are uniform for the purpose of testing each technique used for model development and forecasting. To maintain uniformity amongst the test cases, the history weather website for the Charlottesville, VA area was utilized to extract similar climatic conditions. Eight test cases have been selected from each individual month of the year. The purpose of the eight test cases is to assess the prediction accuracy of the developed models under varying outside temperature and time of day conditions (i.e., clear conditions and precipitation or unfair conditions; all within and outside the allotted six hour forecast horizon) as shown in Table VI-2.

Table VI-2: Eight test cases designed to evaluate model prediction accuracy

ETHER Models	Outside Temperature and Time of Day (OTTD) Conditions
Day	Clear (Within the 6 hour forecast horizon)
	Inclement (Within the 6 hour forecast horizon)
	Clear (Outside the 6 hour forecast horizon)
	Inclement (Outside the 6 hour forecast horizon)
Night	Clear (Within the 6 hour forecast horizon)
	Inclement (Within the 6 hour forecast horizon)
	Clear (Outside the 6 hour forecast horizon)
	Inclement (Outside the 6 hour forecast horizon)

Figure VI-5 and Figure VI-6 show the results for Day and Night ETHER models for the month of September, based on the description provided for each technique. The results for the remaining months can be found in Appendix XVIII. The Day (Warm) ETHER models and the Night (Hot)

ETHER models produced lower success ratio results compared to the other conditions. From analysis of the data, it was determined that this could possibly be due to the amount of missing data that was extracted from the dataset; approximately 10% of the data is missing. As mentioned earlier, the missing data or blanks found within the overall dataset were not replaced by methods of linear interpolation or mean of the series to fill in the missing sections. Our goal was to achieve a finite number of low-fidelity models to predict the interior temperature conditions to enhance decision making opportunities to achieve demand response.

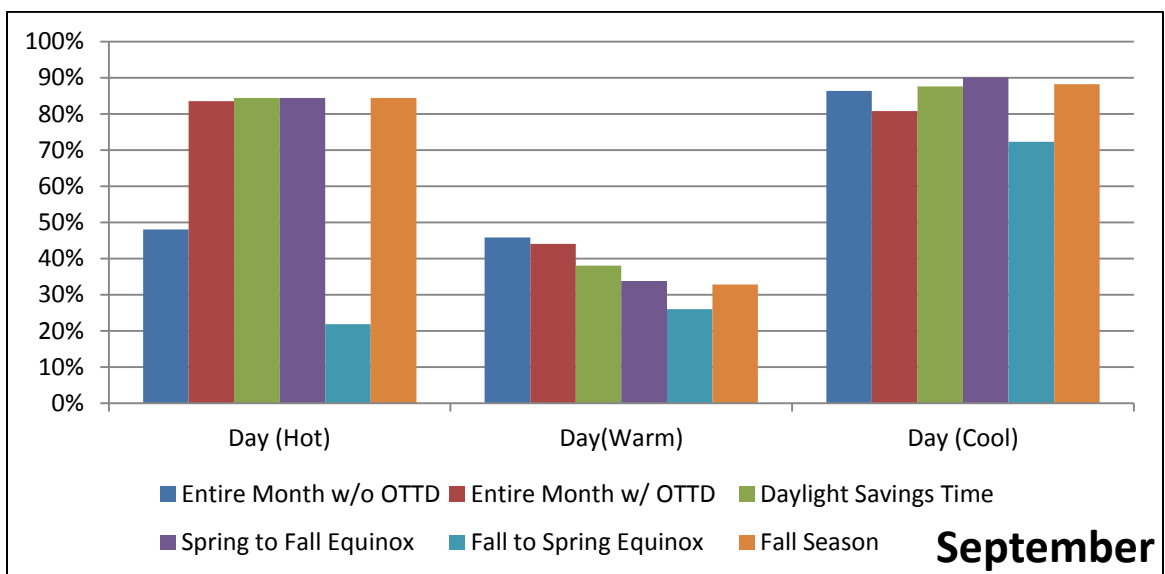


Figure VI-5: Evaluation of the Research Techniques for the Day ETHER Models

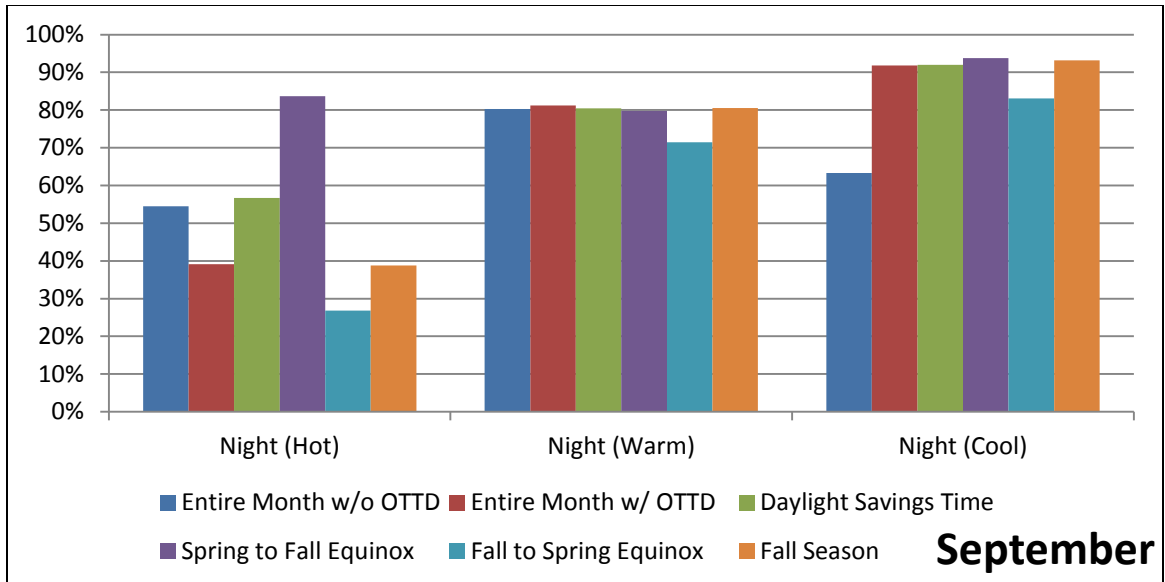


Figure VI-6: Evaluation of the Research Techniques for the Night ETHER Model

With the Day and Night models, the Entire Month w/o OTTD technique was outperformed in most cases by the Entire Month w/ OTTD technique. Thus, the Entire Month w/o OTTD technique was eliminated for further evaluation purposes. The Daylight Saving Time technique outperformed the Entire Month w/ OTTD technique. Therefore, the Entire Month w/ OTTD technique was eliminated for further evaluation. The Equinox to Solstice technique (i.e., Fall to Spring Equinox and Spring to Fall Equinox) performs quite well; however, this technique does not display consistency in performance with the models being separate to accommodate for the year being split in half. As a result, the Equinox to Solstice technique was eliminated for further evaluation. As depicted in the figures, the Seasonal technique (i.e., Fall Season) model shows comparable results to the other model techniques. The shortfall of this technique is that there are a total of seventy eight models that are developed that require tedious manipulation of the data during the data cleaning and pre-processing phase to separate the data into the appropriate outside temperature and time of day conditions. Consequently, the Seasonal technique was eliminated for further evaluation. The Daylight Savings Time technique, on the other hand, developed fewer models (i.e., eight) and performs on the same scale for both the Day and Night ETHER models.

In the next chapter, the ARIMA-based ETHER models will be investigated based on goodness of fit measures to observe the model accuracy of the developed models in this research.

VII. Model Validation

The Evaluation of Thermal Response (ETHER) models are trained and validated by using observational data in a historical period to test the accuracy of how well the ETHER models align with actual data collected from the residence. In the training and validation phase, enough data has to be taken into account to emulate the interior behavior of the residence based on the specified conditions (i.e., optimization of load profiles, etc.). In our research, enough data is defined as 70% of the observations in each dataset. In an effort to better understand the characteristics of the ETHER models and minimize the effects of data discrepancies, most of the observation data is used for training and the remaining data is used for testing [52]. Kusiak et al used a similar approach to build models based on observations for wind farms [3]. Table VII-1 highlights the validation information (i.e., fit statistics) presented for one of the ETHER models (i.e., Night (Cold)) using 70% of the observation data to train the model. The fit statistic metrics indicate how well the developed model fit the data.

Table VII-1: Example of Fit Statistics Metrics for the Trained Night (Cold) ETHER Model

Fit Statistic	Mean
R-squared	0.998
RMSE	0.082
MAPE	0.044
MaxAPE	7.446
MAE	0.032
MaxAE	5.586

The R-squared value measures how well the ETHER model predicts the actual interior conditions of the residence. Statistically speaking, an r-squared value of 1.0 generally indicates a good fit. For that reason, the mean value presented in Table VII-1 indicates a near perfect fit of the data to develop the fitted ETHER model. The root mean squared error (RMSE) is the square root of the average squared errors. The rule of thumb is that the smaller the statistical value, the better the model [53]. The mean absolute percentage error (MAPE) provides a statistical measure of how

large the error is, given that the value is always positive. The maximum absolute percentage error (MaxAPE) provides the single value of the maximum percentage error value that resulted during the development process for the model. The mean absolute error (MAE) is a statistical measure that averages the absolute value of the errors and the maximum absolute error (MaxAE) provides the single value for the maximum absolute error that occurred during model development. As a result of the values presented for the fit statistics, the fitted Night (Cold) ETHER model is acceptable for forecasting due to its prediction accuracy.

The remaining 30% of the data in each dataset is used for forecasting purposes. The data used during the forecasting period is not included in the model-building process; therefore, there is no redundant use of data. The data in each dataset is sectioned off into 6 hour time intervals for forecasting. In most cases, the forecasting time intervals overlap into the different time of day periods as well as outside temperature ranges. To that end, based on the outside temperature and time of day conditions, a model is selected that best fits the specified conditions to forecast up to 6 hours in advance to test the accuracy of the model. Model accuracy justifies how well the models predict within their own bounds. In addition, it gives a measure of how well the models can predict outside of their bounds as well.

A. Goodness of Fit Measures

There are six measures used in our research to measure the prediction accuracy of the ETHER models: success ratio, measure of effectiveness, predictive error (PE), root mean squared error (RMSE), mean average percentage error (MAPE), and mean absolute error (MAE). The success ratio, defined by (10), is a percentage measure of the accuracy of prediction to remain within the bounds of +/- 1 degree F.

$$\text{Success Ratio} = \frac{\# \text{ of Predictions} - \# \text{ Outside of Bounds}}{\# \text{ of Predictions}} \times 100 \quad (10)$$

These bounds were set to equate to the setpoint swing of the digital thermostat located within the residence. For example, if the setpoint temperature is 74 degrees F, the HVAC system will warm up conditions to shut off at 75 degrees F or cool down conditions to shut off at 73 degrees F. The validation of the forecast accuracy of the ETHER models is conducted by selecting eight test cases (i.e., eight random 6 hour intervals) per ETHER model to test the accuracy of the forecasting ETHER models based on different times of the day, varying environmental conditions and seasonal variation influence. Figure VII-1 shows the success ratio results for the Day and Night Models within 1 degree F of measurements.

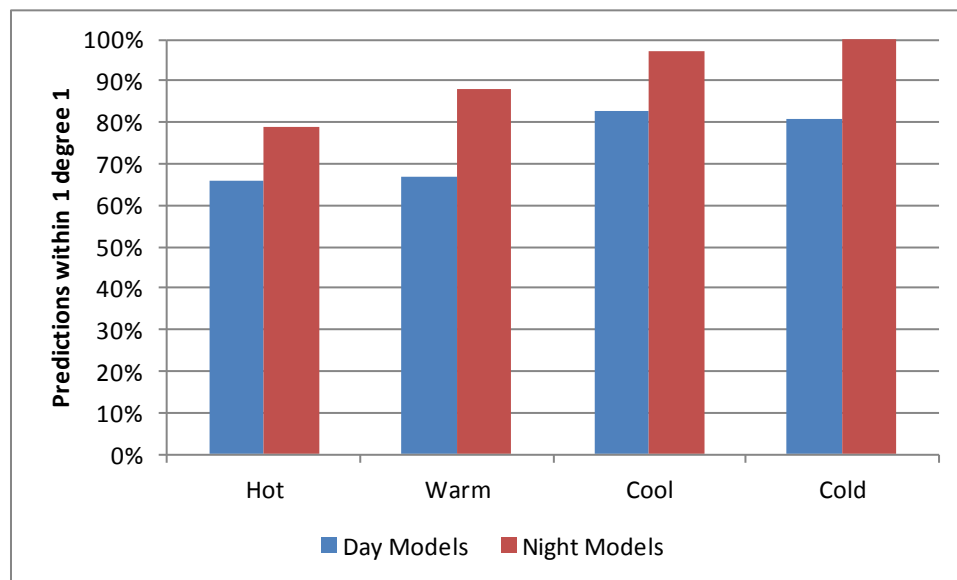


Figure VII-1: Success Ratio Results within 1 degree F for Day & Night ETHER Models

The structure of the Day Models are ARIMA (0,1,18), (0,1,5), (0,1,0), (2,1,2) and the structure of the Night Models are ARIMA (0,1,0), (0,1,1), (0,1,0), (1,1,5), respectively. The Day Models show a success ratio of 66% or higher within the 1 degree F comfort band and the Night Models with a success ratio of 82% or higher within the 1 degree F comfort band. The Day Models produce the lowest prediction accuracy; however, this may be due to other factors (i.e., magnitude of missing data within the dataset, opening windows and/or doors for natural ventilation, electric

fans used for cooling, electric heaters used for space heating, etc.). Experiments were conducted with the additional interior temperature sensor (i.e., Sensor T11) that was highly impacted by solar loading conditions to represent the direct solar radiation into the building structure; however, the results of prediction accuracy were not improved.

In literature, many researchers are concerned with forecasting accuracy within 2 degree C and 3 degree C of measurements. Measurements recorded in literature are reported with units of Celsius and the measurements recorded in our work are reported with units Fahrenheit. Evaluation tests have been conducted to compare the forecasting accuracy of our models to be within 2 degree F and 3 degree F as well. However, for the purposes of comparison, the magnitude of the error predictions is examined.

Mathews et al. performed hourly comparisons between predicted and measured interior temperature using a RC thermal network design model [28]. The results showed that more than 80% of predictions are within 2 degree C of measurements and more than 95% of predictions are within 3 degree C of measurements [28]. Figure VII-2 shows the results for the Day and Night Models within 2 degree F of measurements.

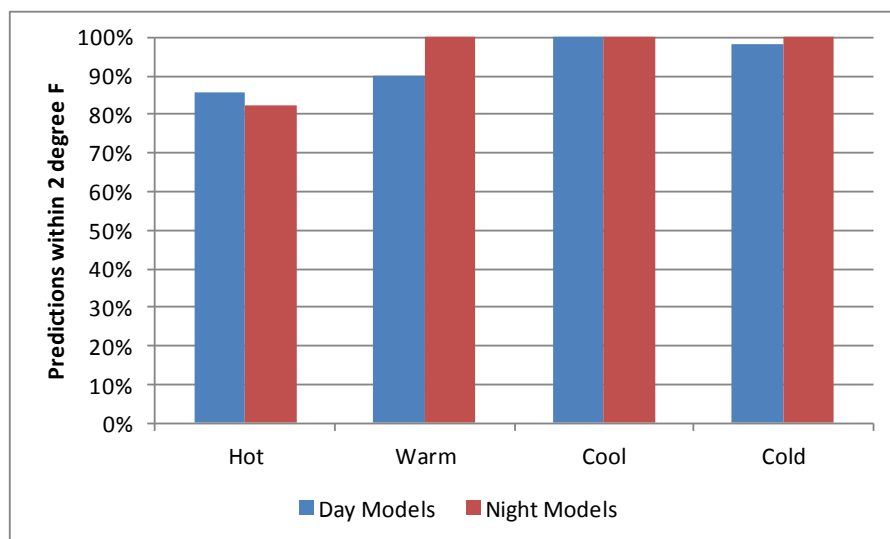


Figure VII-2: Success Ratio Results within 2 degree F for Day & Night ETHER Models

The forecast accuracy for the Day Models is more than 85% within 2 degrees F of measurements and the forecast accuracy for the Night Models is more than 86% within 2 degree F of measurements. Figure VII-3 shows the results for the Day & Night ETHER within 3 degree F of measurements.

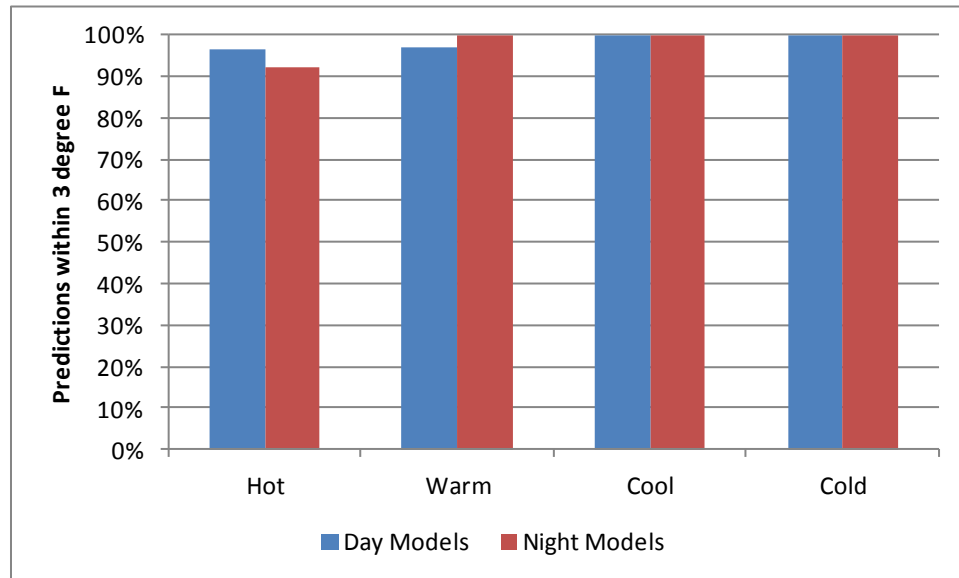


Figure VII-3: Success Ratio Results within 3 degree F for Day & Night ETHER Models

The forecast accuracy for the Day Models is more than 97% within 3 degrees F of measurements and the forecast accuracy for the Night Models is more than 98% within 3 degrees F of measurements. From the results shown, the Day and Night ETHER models produce highly comparable results with those presented by Mathews et al.

The measure of effectiveness, defined by (11), serves as the tradeoff metric to identify the percentage measure of accuracy to stay within the 1 degree F temperature bounds (i.e., success ratio) divided by the number of variables used for model development and forecasting.

$$\text{Measure of Effectiveness} = \frac{\text{Success Ratio}}{\# \text{ Variables}} \quad (11)$$

Figure VII-4 illustrates the measure of effectiveness results for the Day and Night ETHER models.

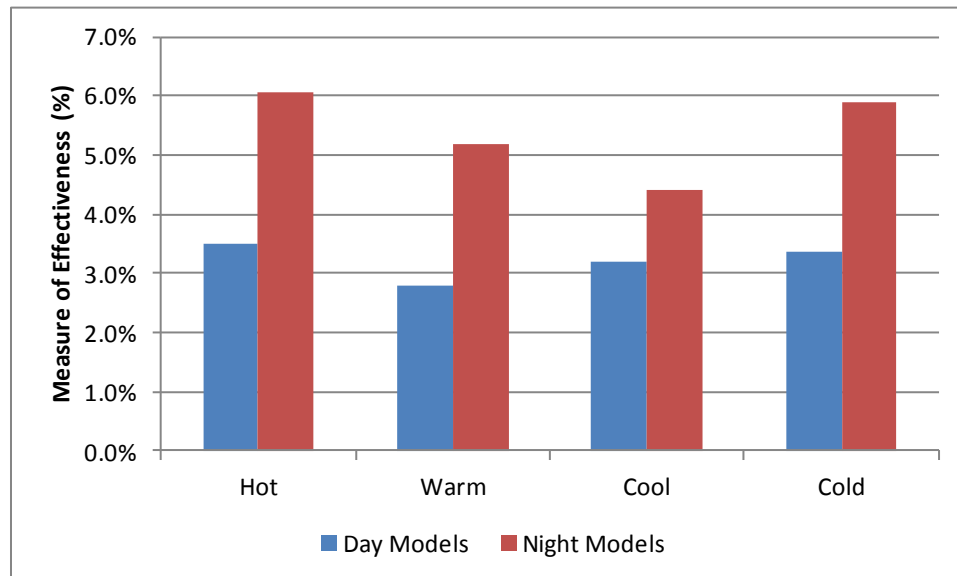


Figure VII-4: Measure of Effectiveness Results for the Day & Night ETHER Models

The rule of thumb for the measure of effectiveness metric is the higher the percentage, the better. There are a total of thirty-one independent variables used as model parameters to develop the ETHER models. In our research, we want to develop the best fit models with the fewest number of parameters to minimize data discrepancies, reduce the magnitude of data, and increase model accuracy by decreasing the correlation amongst the variables.

The prediction error (PE), defined by (12), is the difference between the actual time series data collected from the residence and the forecasts generated by the multivariate time series ETHER models.

$$PE = \text{Actual} - \text{Forecast} \quad (12)$$

The prediction error is a statistical indicator that provides the size of the obtained error, positive or negative in value. This error is used as the basis to derive the additional time series forecasting

error metrics. The subsequent equations were discussed previously in the above section; however, these equations are highlighted here for understanding its mathematical relevance to our work.

The RMSE, defined by (13), is the square root of the sum of the squares of the prediction errors.

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^N (\text{PE}(i))^2}{N}} \quad (13)$$

The MAPE, defined by (14), is the average resultant percentage error used to compare different predictions.

$$\text{MAPE} = \sum_{i=1}^N \left| \frac{\text{PE}(i)}{\text{Actual}(i)} \right| \times 100 \quad (14)$$

The MAE, defined by (15), is the average absolute difference between the actual time series data and the forecasts.

$$\text{MAE} = \frac{\sum_{i=1}^N \text{PE}(i)}{N} \quad (15)$$

All six of the goodness of fit metrics help to analyze and determine the accuracy of the developed models and its forecast capabilities. In addition to the success ratio and measure of effectiveness metrics, an evaluation of the accuracy of the ETHER models were also tested using the PE, RMSE, MAPE and MAE metrics. Table VII-2 shows the prediction accuracy of the Day and Night Models specifically highlighting the following goodness of fit metrics: PE, MAPE, RMSE, and MAE.

Table VII-2: Prediction Accuracy Results for Day & Night ETHER Models

Models	MAPE (%)	RMSE (°F)	MAE (°F)	PEMaximum (°F)	PEMinimum (°F)
Day (Hot)	1.2	1.3	0.9	3.6	-2.3
Day (Warm)	1.0	1.2	0.8	3.5	-1.6
Day (Cool)	0.7	0.7	0.5	1.7	-2.0
Day (Cold)	0.7	0.9	0.5	2.0	-2.4
Night(Hot)	1.3	1.4	1.0	1.0	-3.9
Night (Warm)	0.6	0.6	0.5	1.7	-1.5
Night (Cool)	0.3	0.4	0.3	1.4	-0.8
Night (Cold)	0.5	0.4	0.3	0.8	-0.7

The largest prediction error above the measured interior temperature is 3.6 degree F in the Day (Hot) Model and the largest prediction error below the measured interior temperature is -3.9 degree F in the Night (Hot) Model. Explicitly comparing the prediction performance results of the Day Models versus the Night Models in all six of the goodness of fit metrics, the Night Models perform quite well compared to the Day Models. The rationale behind the accuracy of prediction for the Night Models could be due to the uniformity of activity patterns within the residence at night during the off-peak hours. In addition, overall, the average interior temperature error in the Day and Night Models is well below 2 degree F, which is the typical interior temperature variation limit in the literature.

B. Assumptions

Since demand parameters vary from one residence to the next, using observational data for model development is very useful to factor in seasonal variation, deterioration of structural materials and the fluctuation of heating and cooling requirements [54]. Using 70% of the observation data to develop the ETHER models presents a feasible way to predict interior conditions over a short duration of time. In fact, the ETHER models demonstrate the ability to predict interior conditions while maintaining a desired comfort level for the occupants with 1 degree F of measurements.

The ETHER models can also assist in studying how changes in energy consumption can impact thermal comfort while trying to achieve demand response.

The presented validation results showed that the Day and Night ETHER models are able to maintain a good level of prediction accuracy of the interior conditions within the residence. There are, however, four assumptions that were made for model validation purposes of our research.

- ❖ 70% of the data used for training is enough data to recognize activity recognition patterns that deal with the optimization of loads and varying environmental conditions.
- ❖ The predictive error bound should be equivalent to the setpoint swing (1 degree F) of the digital thermostat to maintain the interior temperature at a thermal comfort level within the residence.
- ❖ The selected forecast period (i.e., up to 6 hours) is an acceptable duration of time to allow for modifications of load appliances to reduce energy consumption.
- ❖ The forty-five sensors used to collect measurements were strategically placed to capture the essential influential behavior of the environment within the residence.

Thirty-one out of the forty-five sensors are being used as model parameters for development of the ETHER models. Even though the goodness of fit metrics results in errors that are within our bounds, we would like to maintain the same level of accuracy or better while reducing the number of parameters and as a result, the magnitude of data. Accuracy of model predictions can provide the consumer with reliable information to make decisions in regards to energy usage and how best to achieve cost savings. In the next chapter, we will take a look at the evaluation of different model approaches to increase the accuracy of model predictions while minimizing the number of model parameters.

VIII. Evaluation Approaches

The evaluation approaches used throughout this research are known as model reduction approaches. These approaches are designed to increase the accuracy of the model predictions to remain within the evaluation bounds (i.e., 1 degree F of measurements) by decreasing the number of variables and reducing the magnitude of data used for model development. The ability to make predictions successfully is important to providing reliable information to an intelligent framework to enhance decision making opportunities on behalf of the consumer. A set of eight forecasting Evaluation of Thermal Response (ETHER) models have been developed, with each model reduction approach using observational data from a residence to predict interior conditions over a 6 hour forecast horizon. The advantage of being able to forecast conditions for the consumer will allow the opportunity for energy management practices and implementation of control strategies to reduce energy consumption and achieve cost savings.

As noted earlier, the observation data is recorded on a minute to minute basis by forty-five sensors within the residence. This presents a significant challenge in that the quantity of data within the dataset as well as the collection of sensors used to obtain the data can be redundant or highly correlated in nature. If this challenge is ignored 1) over-monitoring due to an abundance of sensors will continue, 2) energy usage will continue to increase from lack of insight and 3) demand response may not be achieved. In an effort to obviate the impact of correlation amongst the variables, there are four model reduction approaches that are examined to evaluate the tradeoff between prediction accuracy and the reduction of information.

The three model reduction approaches that will be discussed in detail in this chapter are: 1) the Statistical Package for Social Science (SPSS) Approach, 2) the Reduced Variable Approach and 3) the Select (Reduced) Variable Approach. During the data preparation and processing phase, thirty-one sensors (which are referred to as variables in this research), were identified to be used as independent variables for model development. The remaining fourteen sensors excluding one

(i.e., Sensor T11) were averaged into a composite function to serve as the dependent variable (i.e., interior temperature with no solar loading). All of the model reduction approaches were examined in the SPSS software application tool, resulting in both controllable and uncontrollable demand parameters used for model development, forecasting and analysis purposes.

A. Baseline Approach

The SPSS Approach is the baseline approach used in this research. This approach is divided into three phases: 1) Import Phase, 2) Model Development Phase and 3) Forecasting Phase. All the variables shown in the Chapter IV as input sensors are used for model development through the use of built-in applications in the SPSS tool; the best fitting models are produced and used for forecasting.

A.1. Import Phase

In the import phase, the data is imported into the SPSS 16.0 for Windows® tool as a comma delimited file (i.e., CSV) as shown in Figure VIII-1. The utilization of the CSV file format makes for a smooth data transfer into the SPSS tool. Prior to the transfer of data, the composite function for the dependent variable has previously been calculated in an Excel spreadsheet.

	DateT	Time	F1	F2	F3	F4	CT1	CT2	CT3	CT4
1	09-Mar-09	11:42	0	0	0.00000	0.000000	23.5357	60.0000	52.4643	4.37500
2	09-Mar-09	11:43	0	0	0.00000	0.000000	23.5273	60.0545	52.3636	4.32727
3	09-Mar-09	11:44	0	0	0.00000	0.000000	23.5536	60.1964	52.2500	4.28571
4	09-Mar-09	11:45	0	0	0.00000	0.000000	23.4909	60.3455	52.2545	4.29091
5	09-Mar-09	11:46	0	0	0.00000	0.000000	23.5179	62.2321	52.4286	4.28571
6	09-Mar-09	11:47	0	0	0.00000	0.000000	23.3455	62.4364	52.1818	4.23636
7	09-Mar-09	11:48	0	0	0.00000	0.000000	23.3455	59.6000	51.9455	4.21818

Figure VIII-1: CSV File Format into SPSS

A.2. Model Development Phase

In the model development phase, the SPSS tool has a built-in procedure known as the Time Series Modeler, shown in Figure VIII-2.

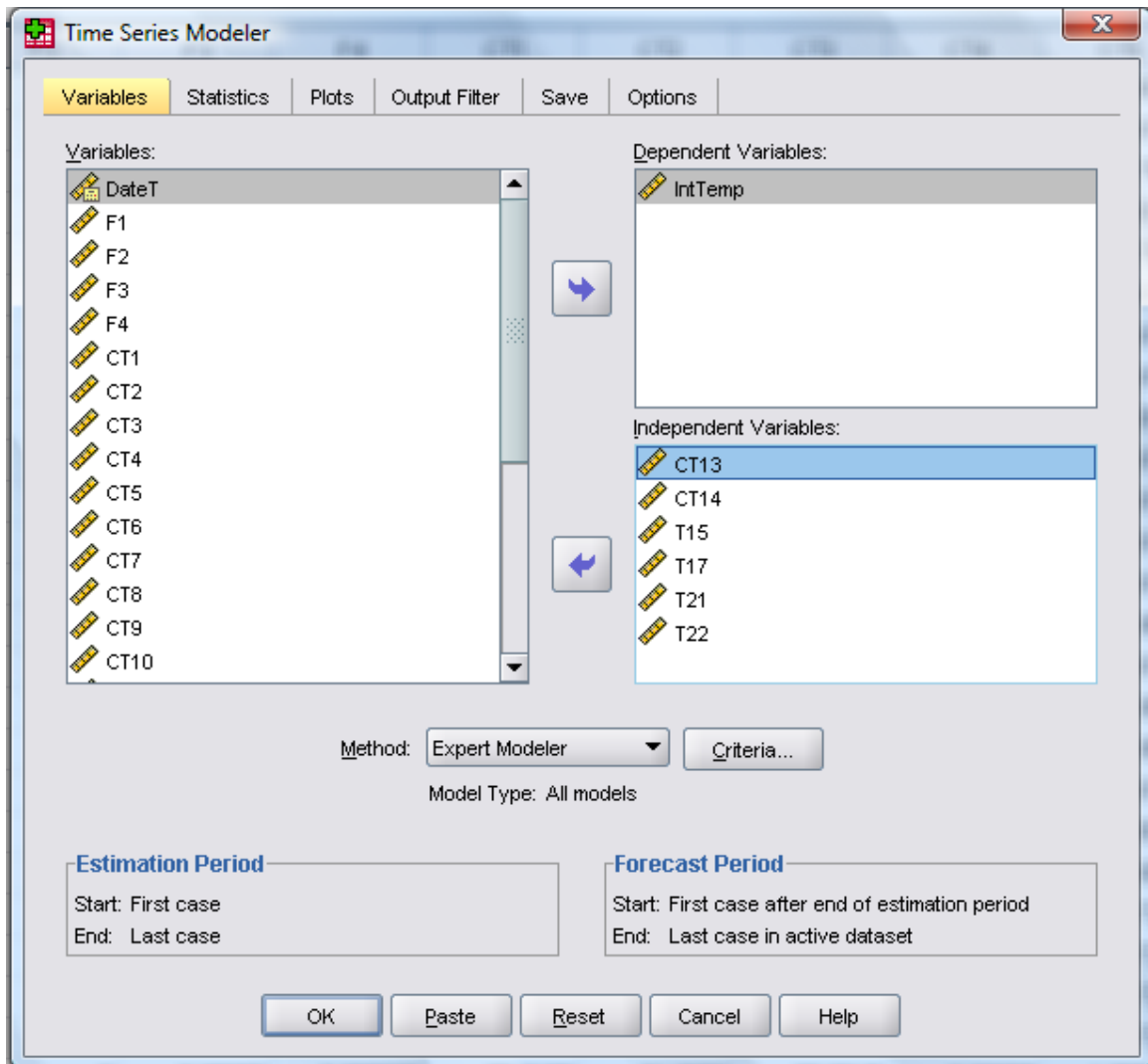


Figure VIII-2: Time Series Modeler with Dependent and Independent Variables

The Time Series Modeler is used to create models and produce forecasts [60]. All the variables that are present in the CSV file are shown within the Time Series Modeler for selection to facilitate the process of model development. This procedure includes an operating method, known as the Expert Modeler, which automatically determines the best fitting model based on the identified data series (i.e., dependent and independent variables). One of the advantages of using the Expert Modeler is that it eliminates the need for an analyst to use trial and error methods to identify an appropriate and acceptable model. Based on the indicated independent variables (i.e., thirty-one variables), the Expert Modeler selects the independent variables that have a statistically

significant relationship with the dependent variable series (i.e., interior temperature) based on outside temperature and time of day conditions [60]. The SPSS Approach allows the Expert Modeler to select and in most cases reduce the number of independent variables that are used for model development and forecasting purposes.

A.3. Forecasting Phase

Once the fitted model structure has been identified with the Expert Modeler, SPSS uses this model structure for forecasting up to 6 hours in advance. The developed model is exported and saved as an extensible markup language file (i.e., XML file) that stores all of the selected independent variables that make up the constructed ETHER model based on varying specified conditions. For the forecasting phase, the SPSS tool has a built-in procedure known as Apply Times Series Models, shown in Figure VIII-3. This procedure allows the developed ETHER model to be applied to predict interior temperature conditions over a 6 hour forecast horizon on more current data.

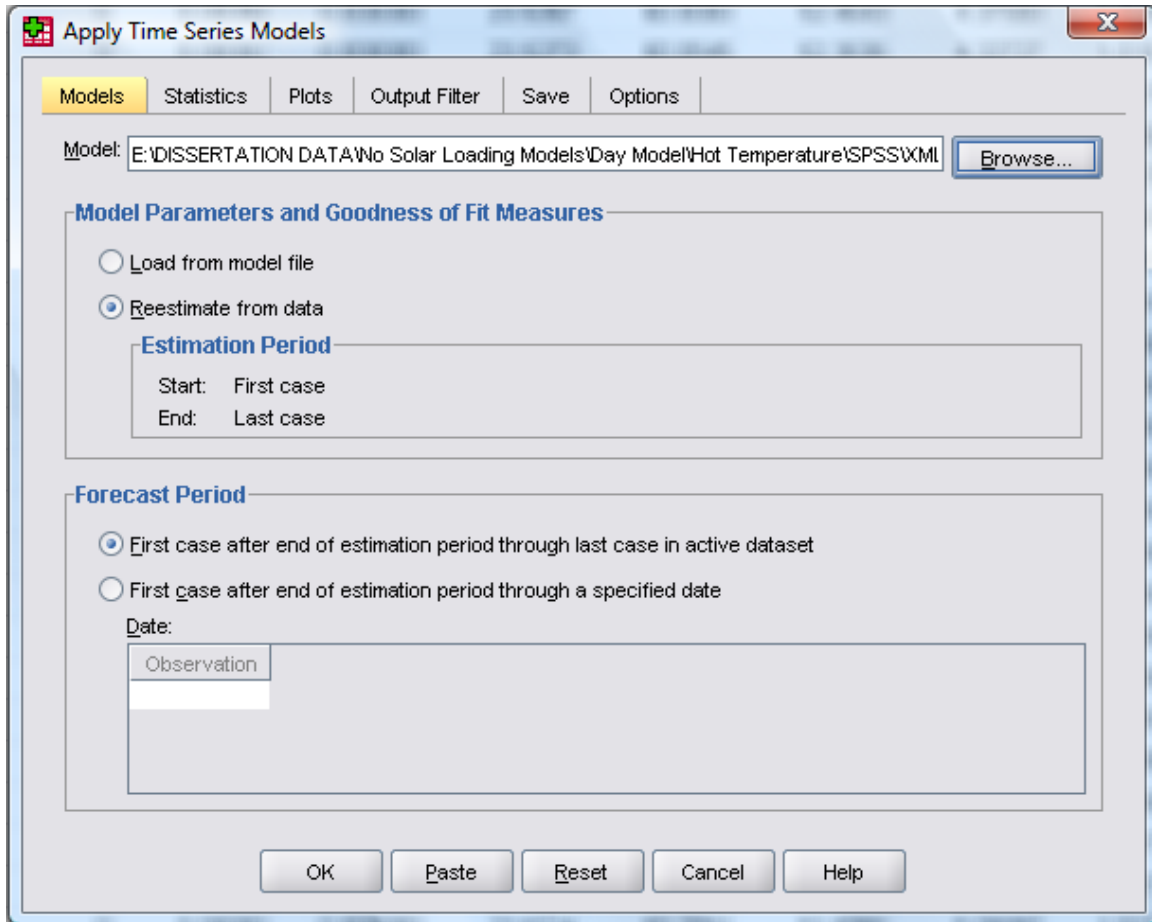


Figure VIII-3: Apply Time Series Models

With the Apply Time Series Model procedure, the structure and the independent variable predictors (i.e., model parameters) for the developed ETHER model remains the same; the only variation is the change in the coefficients of each model parameter. Therefore, the models are not rebuilt. This further enables the Apply Time Series Model to save processing time to obtain prediction results. With a new active dataset or more current data imported into the SPSS tool, the “Reestimate from data” option is always selected to incorporate the changes in the load consumptions profiles, temperature and humidity values, etc. An assumption is made that the new dataset has the same variable names as those used in the model development phase. If not, a new model will need to be developed to incorporate for the changes in the variable name.

A.4. Model Development & Forecasting Automation

One of the stated goals of this research is to automate the model development and forecasting process to produce self-constructing and self-refining models to predict interior temperature conditions within the residence to maintain comfort levels. The development process of the predictive ETHER models is fully automated through the use of SPSS. The automation of the model development process alleviates the manual entry of data into SPSS, the selected processes for model development as well as the updating of forecasts when actual data becomes available.

SPSS has its own command language that can be used to automate the process for model development and forecasting. A syntax file has been scripted which extracts data (i.e., temperature, humidity, carbon dioxide measurements, etc.) in a CSV file format and imports the data into SPSS where the data is processed in order to develop a time series model that represents the conditions of the residence. The command language syntax code has been written and generated for all three phases (i.e. data import, model development and forecasting). The syntax code for data import, model development and forecasting can be found in Appendix B through Appendix D, respectively.

The automation process of the data import into SPSS is shown in Figure VIII-4. The flowchart depicts the process used to read data into SPSS from the CSV file and format the values and variables accordingly for readability. Under the Identification of Variables header, there are three items listed to identify the data that is being read into SPSS. The variable type corresponds to a numeric, decimal, string, date, scientific notation, etc; decimals represent the number of significant digits after the decimal for the variable value and measure labels if the variable is a scalar, ordinal or nominal value. An assumption is made that all the dependent and independent variable values are scalar and numeric. The data import process has been validated based on the results obtained from manual data entry and pre-processing procedures.

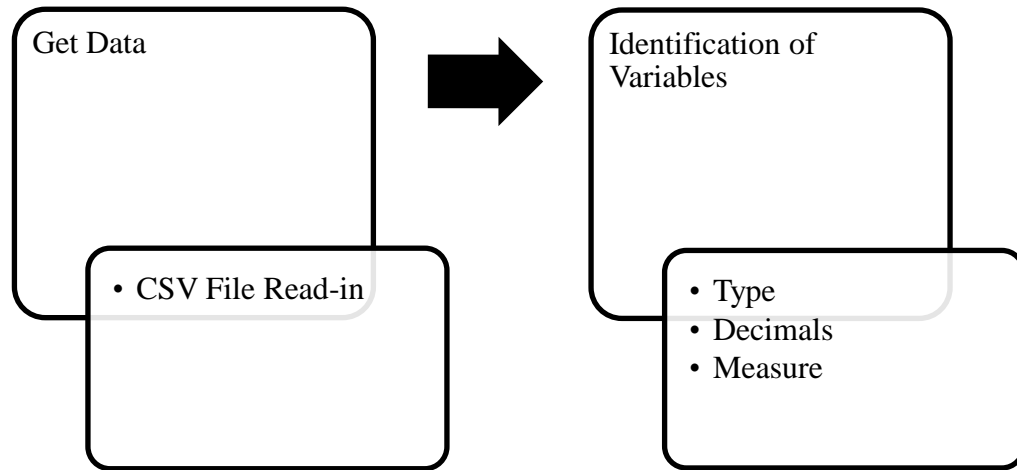


Figure VIII-4: Data Import Procedure into SPSS

The flowchart, shown in Figure VIII-5, shows the steps of creating the time series model structure in SPSS. The dependent and independent variables are selected based on the desired output for the residence. In our research, the dependent variable is the interior temperature and the independent variables are the remaining controllable and uncontrollable variables. Once the variables are identified for model development and the process has been executed, the fitted model structure is saved for forecasting purposes.

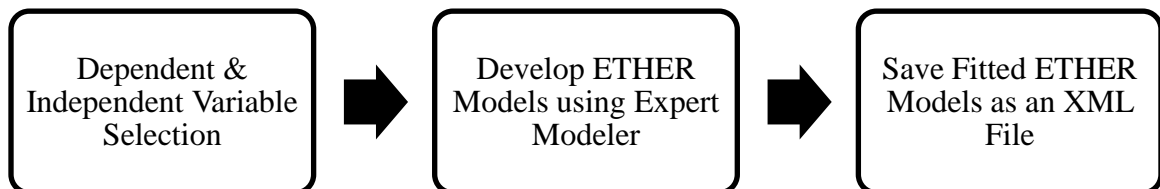


Figure VIII-5: Model Development Steps in SPSS

The ETHER forecasting process is a cyclic process which highlights the sequence of steps that are taken to constantly refine parameters as actual data becomes available, shown in Figure VIII-6. We envision real time data being collected from the input sensors to supply information to the set of ETHER models related to the independent and dependent variable series. The

parameters can then be re-estimated based on the actual data received prior to forecasting. As an input into the ETHER model, load profiles will be imported into the models for forecasting evaluation of the interior conditions in order to manage energy consumption to achieve demand response. Therefore, when actual data becomes available after a forecasting period, the data can be read into SPSS and the parameter coefficients and parameter values can be updated. This allows for constant refinement of the parameters.

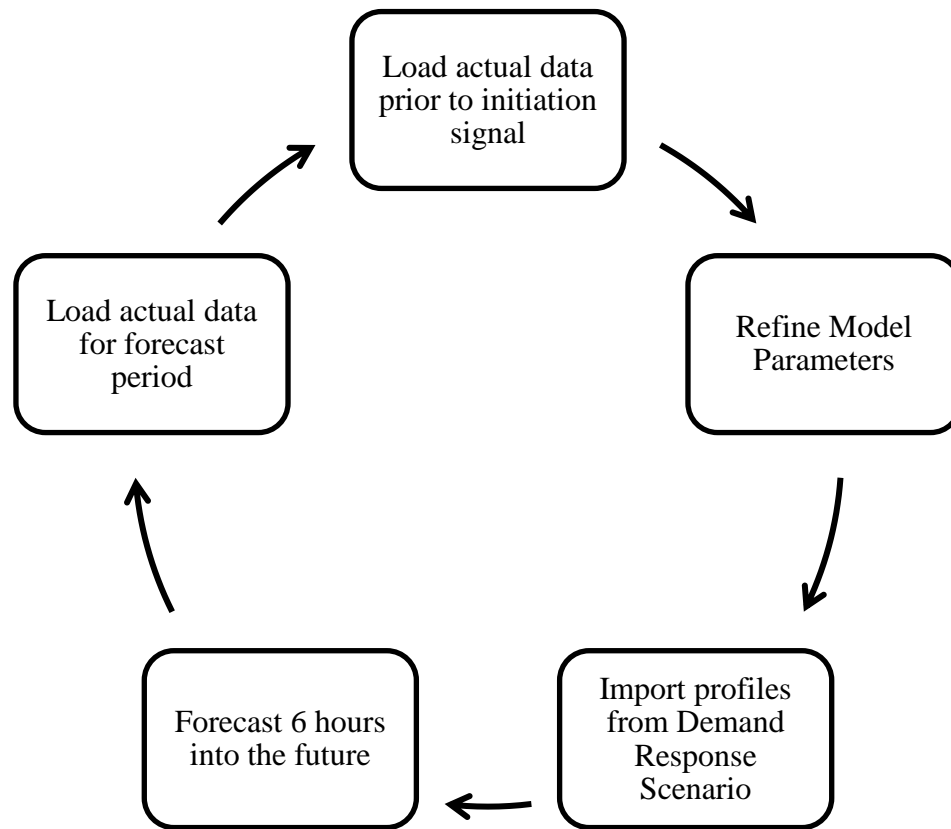


Figure VIII-6: ETHER Model Forecasting Process

Automation of the three phases highlighted within the SPSS Approach is essential for the remaining three model reduction approaches. The challenge with this approach is that SPSS does not automatically check for variables that are highly correlated, which is also referred to as multicollinearity. Multicollinearity means there is a strong linear relationship amongst the independent variables. As a result, it identifies some variables as statistically insignificant when

they should be significant [64]. Multicollinearity will not impact the model's reliability; however, it will impact the selected predictors (i.e., independent variables) to be used for model development and forecasting. In addition, multicollinearity makes it difficult to assess the importance in determining the dependent variable series (i.e. interior temperature) [65]. In an effort to assist in the mitigation of the multicollinearity risk, a reduced variable approach has been developed to check for correlation amongst the variables during a preliminary evaluation and observe the anticipated increase in the model prediction accuracy.

B. Reduced Variable Approach

The reduced variable approach was designed to identify and obviate multicollinearity amongst the independent variables in an effort to reduce the number of variables that will be used to develop the ARIMA time series ETHER models. The reduced variable reduction approach has been divided into a total of three reduction methods: 1) Data Source Correlation, 2) Regression Collinearity Diagnostics - Variable Inflation Factor (VIF) and 3) Significance (p-value) indicator. The methodology of the reduced variable approach is shown in Figure VIII-7. Each method is conducted within SPSS prior to model development.

Initially, the reduction process was conducted manually to be later automated to reduce the tedious tasks of evaluating each model parameter for reduction. We found that the reduction methods are acceptable for identifying the main influential variables of the dependent variable series as well as the reduction of variables. In addition, the variable reduction process benefited the research by providing insight on the identifying the top five variables (i.e., exterior temperature, interior humidity, exterior humidity, carbon dioxide and HVAC) that impact the interior temperature conditions. Due to these findings, the beginning phases of the variable reduction process have been automated using SPSS syntax code located in Appendix D. The variable reduction procedure was not automated in entirety due to the fact that the process did not eliminate the same sensors from all the developed ETHER models; however, the approach was

beneficial in identifying the main model parameters (i.e., top 5 variables) within each ETHER model.

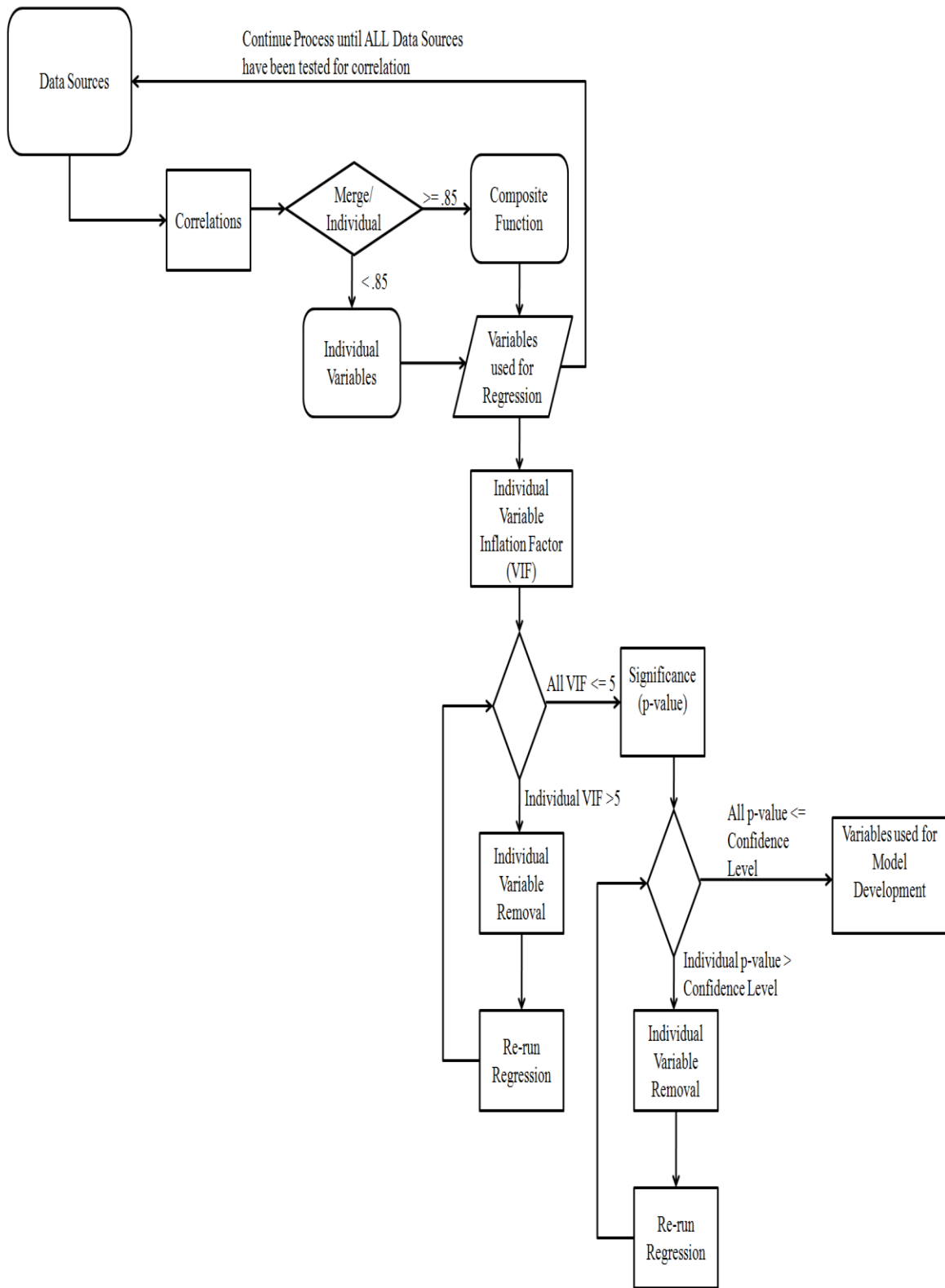


Figure VIII-7: Variable Reduction Flowchart Methodology

B.1. Data Source Correlation Method

The data source correlation method groups each data source into the same construct. There are a total of six data sources (i.e., humidity, power consumption, exterior temperature, solar hot water temperature, carbon dioxide measurements and water flow) detailed in Chapter IV. The purpose of data source correlation method is to identify the variables that are of the same type or same scale to reduce the redundancy of information up front. A bivariate correlation test is performed in SPSS which identifies a strong (i.e., $\geq .85$) or weak (i.e., $< .85$) correlation amongst the independent variables of the same construct. Lani [66] states the strength of a correlation ranges between the absolute value from 0 to 1; the closer the correlation is to 1, the stronger the relationship and the closer the correlation is to 0, the weaker the relationship. For purposes of the ETHER models, .85 has been selected as the cutoff correlation measure. If there is a weak correlation amongst all the variables, the independent variables are stored to be used in the next reduction method. However, if there is a strong correlation, the variables are averaged together to create a composite function. This composite function is in turn used in the next reduction method.

An example of the SPSS output [60] for the bivariate correlation reduction method is shown in Table VIII-1.

Table VIII-1: Bivariate Correlation Results with the CO₂ variables

		Correlations	
		T22	T23
T22	Pearson Correlation	1	.922**
	Sig. (2-tailed)		.000
	N	66071	66071
T23	Pearson Correlation	.922**	1
	Sig. (2-tailed)	.000	
	N	66071	66071

** . Correlation is significant at the 0.01 level (2-tailed).

The Pearson coefficient gives the bivariate correlation measure which quantifies the strength between the variables. The sig (2-tail) value represents the significance factor of the correlation.

The N represents the total number of observations for each variable used in the correlation reduction method. From the correlation matrix results in Table VIII-1, it can be observed that the first floor CO₂ sensor and the second floor CO₂ sensor are highly correlated with one another. As a result, these two variables would be averaged together to form a composite function. This method helps reduce the number of variables used for model development and in turn minimize the amount of data.

B.2. Variance Inflation Factor (VIF) Method

The second method used to help assess multicollinearity is the Regression Collinearity Diagnostics - Variance Inflation Factor (VIF). The VIF measurement is obtained by running the linear regression collinearity diagnostics option in SPSS. The VIF measures multicollinearity in the model which indicates that the variable is almost a perfect linear combination of the independent variables already in the equation and that it should not be added to the regression equation [67]. The purpose of the VIF method is for further reduction of the variables to reduce redundancy amongst the independent variables. This method compares each independent variable to the remaining variables to test for correlation; however, in the data source correlation method, each variable is compared through a pair-wise (e.g. one-on-one) process. Therefore, each independent variable has a VIF value. VIF is defined by the following equation:

$$VIF = \frac{1}{1-r^2} \quad (16)$$

where r indicates the correlation coefficient which is obtained by regressing each variable on the remaining independent variables or predictors [68].

There is no formal standard or consensus amongst statisticians for determining the magnitude of the VIF value that identifies the presence of multicollinearity; however, a crude rule of thumb, according to [69], [70], [71], suspects that a VIF greater than 10 represents severe multicollinearity. This rule has been noted in scholarly articles and statistical textbooks [72].

Studies have also shown that a VIF value around or greater than 5 also signifies evidence of multicollinearity [64]. For our model reduction approaches, if the VIF value is greater than 5.0, the independent variable is removed and the method is repeatedly executed until all VIF values are less than 5.0. Therefore, a VIF value from 1.0 to 5.0 indicates a statistical significance to the acceptable model. The remaining variables are then used in the next reduction method.

B.2.1 Case Study: Modification of VIF Cutoff

Dr. Xitao Fan, currently the Chair Professor and Dean at the University of Macau in China and a former professor here at the University of Virginia, suggested that a case study be performed on the dataset to identify the VIF cutoff due to the lack of formal standards used to indicate multicollinearity. The model reduction approaches used in this research were discussed with Dr. Fan and validated per the following response, “Your approach of combining variables with strong correlations is correct, assuming that the variables being combined measure the same or similar construct. In addition, VIF should give you very good indicate on about which variables should be considered as candidates for further removal.” Due to Dr. Fan’s statistical expertise, he suggested that a VIF value of 2.0 be considered “to make the decision for removal of correlated predictors.” The VIF value gives a factor of the how much the standard error will need to be inflated by to determine the regression coefficient (i.e., if the VIF value is a factor of 5.0, the standard error will be 5.0 times larger compared with the situation when the predictors are not correlated).

A case study was performed to evaluate the results with a VIF value of less than or equal to 2.0 and a VIF value of less than or equal to 5.0 (i.e., $VIF \leq 2.0$ vs. $VIF \leq 5.0$). With a VIF value of less than or equal to 2.0, more variables will be removed which means there will be a significant reduction in information used for model development and forecasting. This case study helps to identify how sensitive the model prediction results are to the reduction of information. In Figure VIII-8, the success ratio results are shown using the Daylight Savings technique for the

Daytime models after reducing the variables using the Variable Reduction Methods. As observed, the VIF value of less than or equal to 5.0 resulted in marginally better predictions to stay within the 1 degree F of measurements bound out of the eight test cases tested for both procedures. Therefore, a VIF of less than or equal to 5.0 is used as the factor of choice for the additional models and test and evaluation procedures.

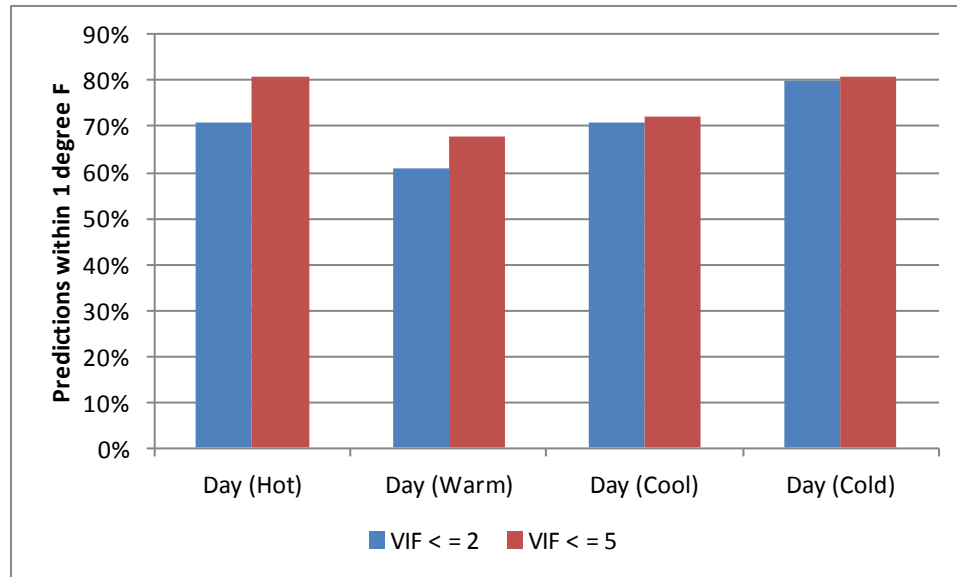


Figure VIII-8: VIF Cutoff Evaluation Results

B.3. Significance (p-value) Method

The significance (p-value) indication method is the third method for variable reduction. This method is obtained by running a linear regression test on the remaining variables to observe their influence of the dependent variable (i.e. interior temperature). The purpose of the significance (p-value) method gives an indication of the impact the each independent variable has on the dependent variable. The significance factor in SPSS has been set to .05 which means there is a 95% confidence that the independent variable has a significant impact on the dependent variable. If the significance factor is very small (i.e., less than .05) then the independent variable has a significant influence on the dependent variable [73]. However, if the significance factor is large

(i.e., greater than .05) then the independent variable has no significant influence on the dependent variable. Therefore, if the significant factor value is greater than .05, the independent variable can be removed from any further evaluation. This method is repeatedly executed until all significance values are less than or equal to the confidence level.

C. Select (Reduced) Variable Approach

The select (reduced) variable approach is an extension of the reduced variable approach. However, in this approach, the water flow rate variables and the solar hot water temperature variables are removed during a pre-analysis procedure (i.e., intuitive engineering decision) to identify which variables provide minimal influence to the dependent variable series. Taking a holistic view of the residence, the water flow rate variables are binary in nature and the solar hot water is stored in insulated tanks; neither of which produce any meaningful information nor contributes to the dependent variable series.

D. Case Study – Comparison of Evaluation Approaches

To test the accuracy of the ETHER models, the three model reduction approaches were evaluated based on outside temperature and time of day conditions. Approach 1 is referred to as the “SPSS Approach”; this is the baseline approach. In this approach, all the data for thirty-one independent variables are imported into SPSS and used to create the best fit model for each dataset. Approach 2 is referred to as the “Variable Reduction Approach”. In this approach, the thirty-one independent variables are reduced based on the variable reduction procedure above in Figure VIII-7. After the variable reduction procedure has been executed, the remaining variables are imported into SPSS and then used to create the best fit model for each dataset. Approach 3 is referred to as the “Select (Reduced) Variable Approach”. In this approach, the water flow rate variables and the solar hot water temperature variables are removed during a pre-analysis procedure (i.e., intuitive engineering decision) to identify which variables provide minimal to no

influence on the dependent variables series. The remaining independent variables (i.e., power consumption loads, exterior temperature, humidity and carbon dioxide measurements) are then reduced based on the variable reduction procedure and imported into SPSS to create the best fit model for each dataset.

One of the research goals in this dissertation is to reduce the number of variables used in the model development process; thus, a closer look at all three model reduction approaches for the Day and Night ETHER Models are required to determine which approach produces the best model prediction results. The goodness of fitness measures that were explained in Chapter VII were also used as evaluation metrics.

Table VIII-2 shows the prediction accuracy given by the goodness of fit measures (i.e., MAPE, RMSE, MAE, Maximum PE and Minimum PE) for each Day Model using the Daylight Savings technique. The goodness of fit measures represents the size of error that is obtained using the different outside temperature and time of day ETHER models. From the values observed, the Select (Reduced) Variable Approach produces the best model prediction accuracy for all of the Day Models.

Table VIII-2: Model Reduction Approaches – Goodness of Fit Measures for Day Models

	Day Models Approaches	MAPE (%)	RMSE (°F)	MAE (°F)	PE Max (°F)	PE Min (°F)
Hot	SPSS	1.2	1.3	0.9	3.6	-2.3
	Reduced Variable	0.9	1.0	0.7	3.0	-1.1
	Select (Reduced)	0.8	0.9	0.7	2.3	-1.3
Warm	SPSS	1.0	1.2	0.8	3.5	-1.6
	Reduced Variable	1.3	1.3	1.0	3.2	-1.1
	Select (Reduced)	0.9	1.1	0.7	3.2	-1.2
Cool	SPSS	0.7	0.7	0.5	1.7	-2.0
	Reduced Variable	0.9	0.9	0.6	1.9	-2.8
	Select (Reduced)	0.7	0.6	0.5	1.9	-1.5
Cold	SPSS	0.7	0.8	0.5	2.0	-2.4
	Reduced Variable	0.8	0.9	0.6	2.4	-2.5
	Select (Reduced)	0.8	0.9	0.5	2.3	-2.6

Figure VIII-9 shows the success ratio comparisons between the three model approaches in the Day Models. The success ratio indicates the performance accuracy of the model reduction approaches within each ETHER model. All the models have a success ratio of more than 66% within 1 degree F of measurements.

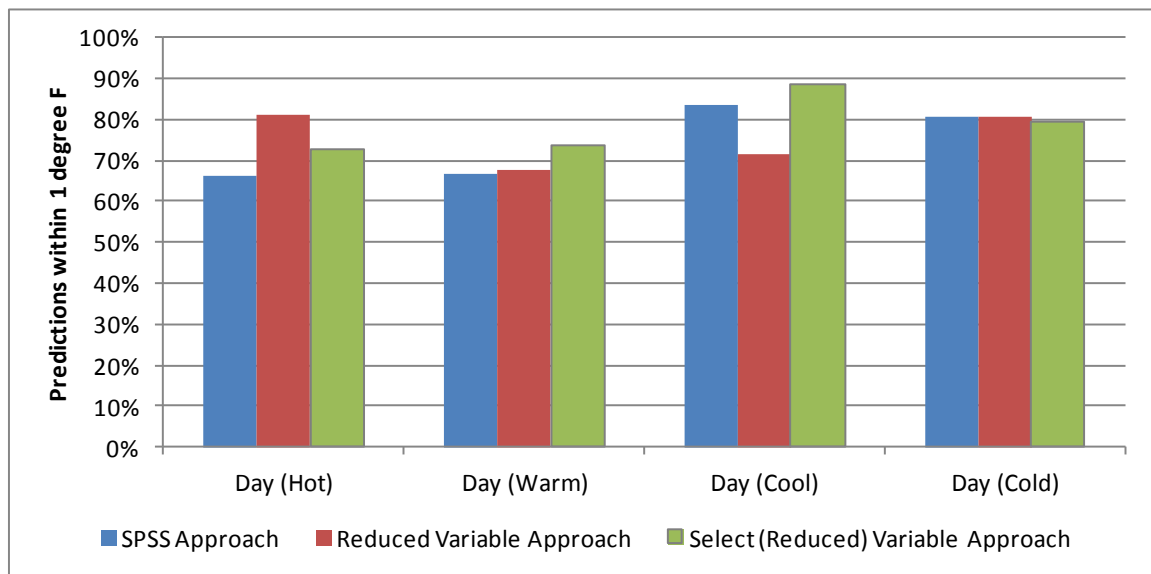


Figure VIII-9: Model Reduction Approaches – Success Ratio for Day Models

Figure VIII-10 looks specifically at the measure of effectiveness amongst the three approaches. This measure shows the comparison of the quality of the models to predict with a varying number of predictors (i.e. independent variables) used for model development and forecasting. With a reduction in variables, the magnitude of data is decreased and in addition, it helps the analyst determine the minimal number of variables that are required to obtain accurate model predictions. This will decrease over-monitoring of the residence and processing time for development and analysis. As observed, amongst the three approaches, the Selected (Reduced) Variable Approach results in the lowest number of predictors (i.e., independent variables) to be used for model development and forecasting. It can be seen from Table VIII-2 and Figure VIII-9 that the Select

(Reduced) Variable Approach produces good results overall; producing accurate predictions with the fewest number of variables.

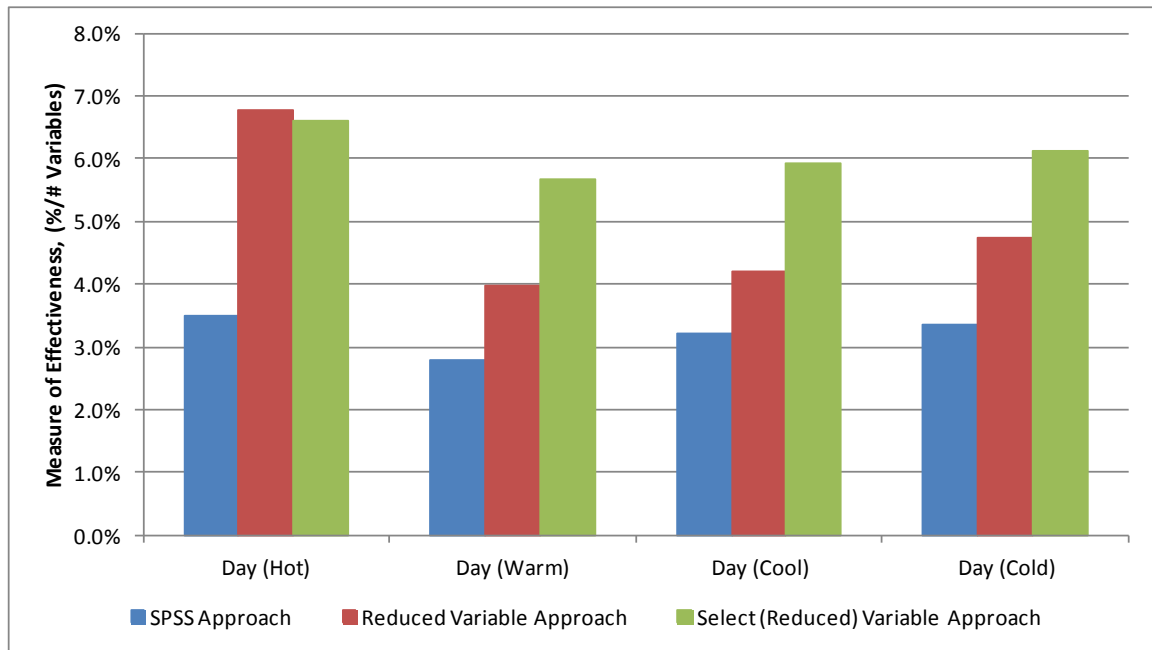


Figure VIII-10: Model Reduction Approaches - Measure of Effectiveness for Day Models

Table VIII-3 shows the prediction accuracy using the three model reduction approaches for the Night Models.

Table VIII-3: Model Reduction Approaches – Goodness of Fit Measures for Night Models

	Night Models Approaches	MAPE (%)	RMSE (°F)	MAE (°F)	PE Max (°F)	PE Min (°F)
Hot	SPSS	1.3	1.4	1.0	1.0	-3.9
	Reduced Variable	1.0	1.0	0.8	1.4	-3.1
	Select (Reduced)	1.0	1.1	0.8	1.0	-3.5
Warm	SPSS	0.6	0.6	0.5	1.7	-1.5
	Reduced Variable	0.7	0.7	0.5	1.7	-1.5
	Select (Reduced)	0.7	0.7	0.6	1.8	-1.5
Cool	SPSS	0.3	0.4	0.3	1.4	-0.8
	Reduced Variable	0.4	0.4	0.3	1.5	-0.8
	Select (Reduced)	0.4	0.4	0.3	1.7	-0.7
Cold	SPSS	0.5	0.4	0.3	0.8	-0.7
	Reduced Variable	0.5	0.4	0.3	0.9	-0.7
	Select (Reduced)	0.5	0.4	0.4	1.0	-0.7

It can be seen in Table VIII-3 that even with a significant reduction in information, both the Reduced Variable and Select (Reduced) Variable Approaches resulted in comparable and minimal prediction errors compared to the baseline SPSS Approach.

Figure VIII-11 and Figure VIII-12 show the success ratio and the measure of effectiveness comparisons between the three model approaches in the Night Models, respectively. The success ratio and the measure of effectiveness results are shown to indicate the performance accuracy of the model reduction approaches within each ETHER model. All the models have a success ratio of more than 75% within 1 degree F of measurements. The success ratio for the Night Models are highly comparable with one another; however, from observing the measure of effectiveness graph, amongst all three approaches, the Selected (Reduced) Variable Approach results in the lowest number of predictors (i.e., independent variables) to be used for model development and forecasting. Even though the success ratio results are similar, the Select (Reduced) Variable Approach produces accurate predictions with the fewest number of variables.

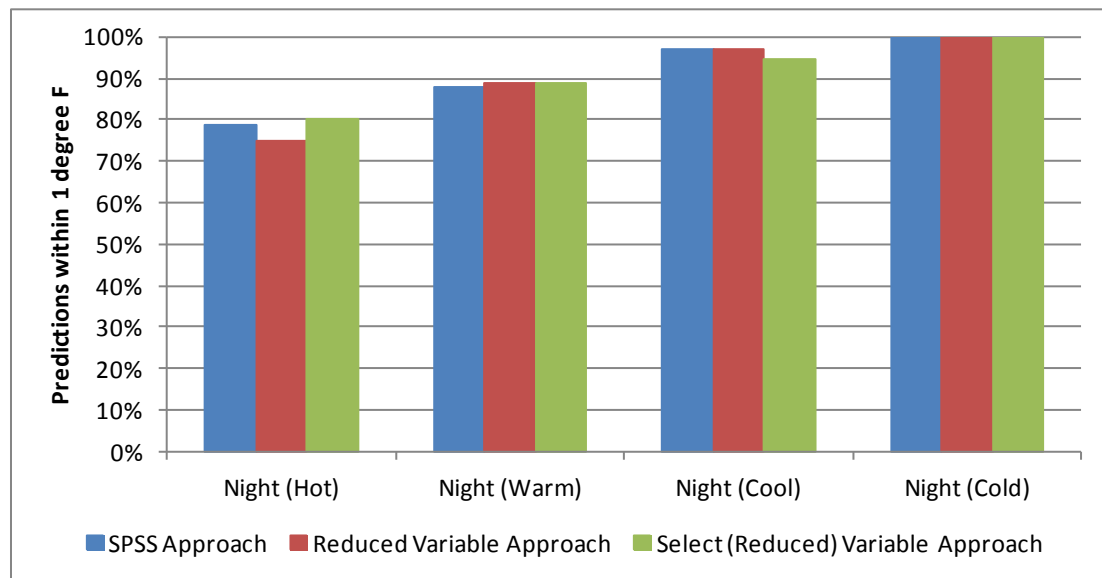


Figure VIII-11: Model Reduction Approaches – Success Ratio for Night Models

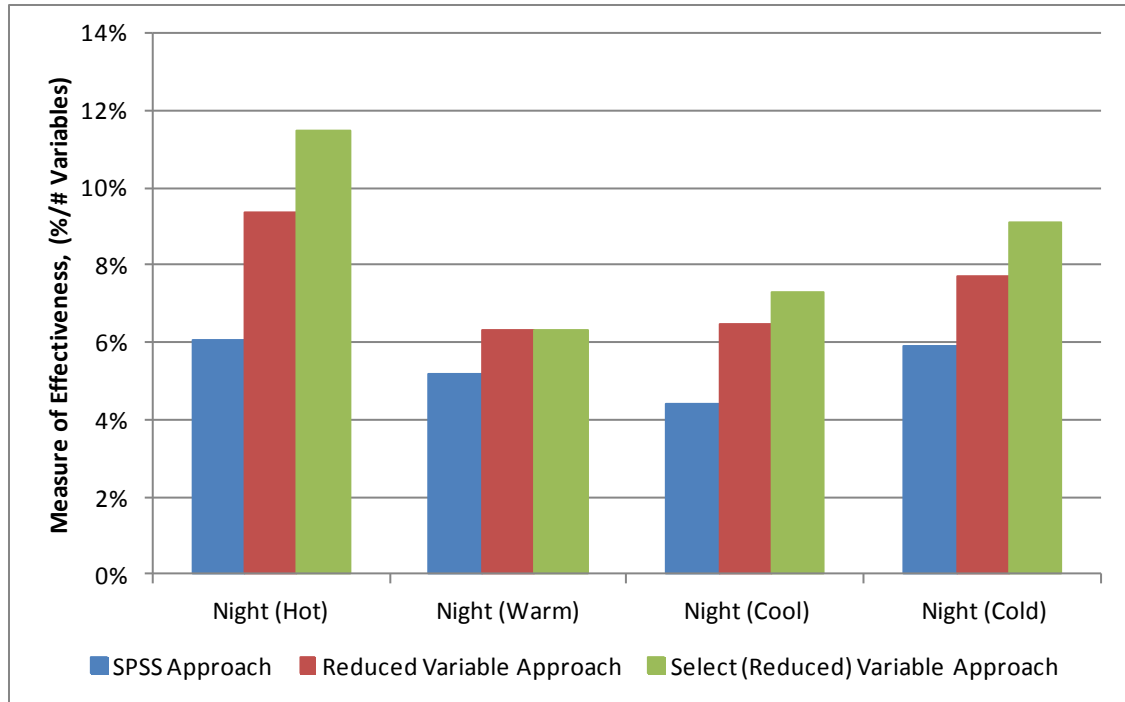


Figure VIII-12: Model Reduction Approaches - Measure of Effectiveness for Night Models

Aforementioned, the objective of the developed ARIMA time series ETHER models is to provide prediction accuracy for the interior temperature conditions to be within 1 degree F of measurements, more than 65% of the time, over a 6 hour forecast horizon using the model reduction approaches. By obtaining accurate predictions, the ETHER models can be used to provide insightful information to the consumers or as part of an intelligent framework to make timely decisions in an effort to achieve residential demand response.

Through the evaluation of the model reduction approaches, it became apparent that the Select (Reduced) Variable Approach produced the best model prediction results and resulted in containing the lowest number of predictors for model development. In the next section, the ETHER models developed from the Select (Reduced) Variable Approach is compared to the model development of using the top five variables that were common amongst all the developed ETHER.

E. Case Study – Select (Reduced) Variable Approach vs. Top 5

Variables

The top five variables (i.e. exterior temperature, interior humidity, exterior humidity, carbon dioxide and the HVAC components) were observed across all the Day and Night ETHER models that were developed during the variable reduction process. These variables were noted as the top five due to their continuous presence within each model as well as their statistical significance factors. The advantages of using the top five variables are the reduction of sensors located in the residence which eliminates the challenge of over-monitoring, the decrease in the processing time for model development and forecasting and the elimination of the same variables across all the developed ETHER models. These advantages satisfy all the specified points for the variable reduction goal.

The objective of this case study was to evaluate the accuracy of the model predictions by using the Select (Reduced) Variable Approach rather than the top five variables that provide a significant influence to the interior temperature conditions within the residence. Interior temperature conditions controlled by a heating, ventilation, and air-conditioning (HVAC) system within a building structure allows occupants to maintain their desired thermal comfort zone [80]. Thermal comfort can be defined as “a condition of mind that expresses satisfaction with the thermal environment” [81]. It has been observed that thermal comfort depends on a range of environmental, personal and behavioral factors. For example, clothing, age, air temperature, humidity, air movement, activity levels (e.g., working out, cleaning, and arguments, etc.) can be contributing factors of thermal comfort. Since the interior temperature within a building structure is highly influenced by variations of outside conditions as well as interior factors [82], the top 5 variables, also referred to as the five thermal comfort variables, were identified as model development parameters to test for model prediction accuracy using the ARIMA process. In Figure VIII-13 and Figure VIII-14, the success ratios for the Day ETHER Models and Night

ETHER Models are evaluated by using the Select (Reduced) Variable Approach and the top five variables, respectively.

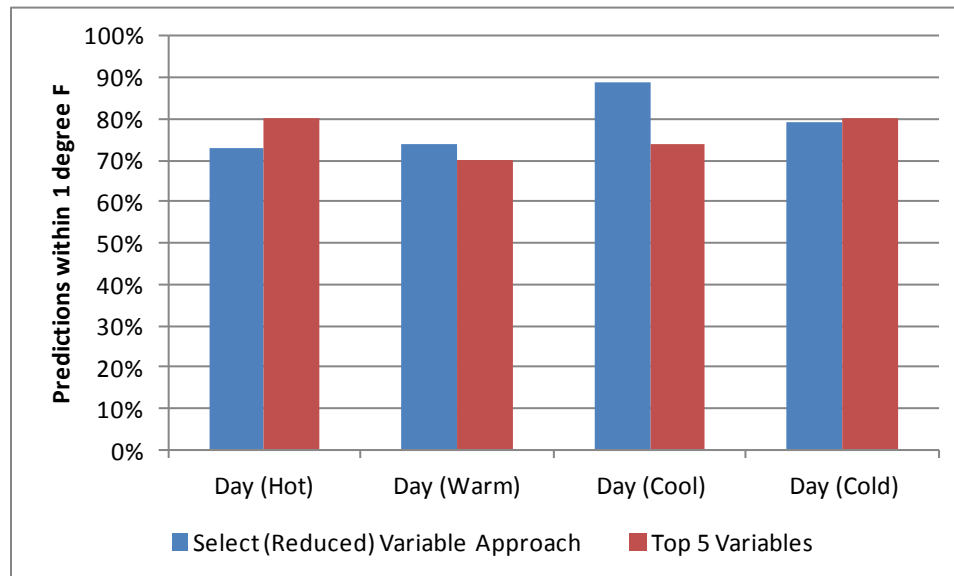


Figure VIII-13: Select (Reduced) Variable Approach vs. Top 5 Variables – Day Models

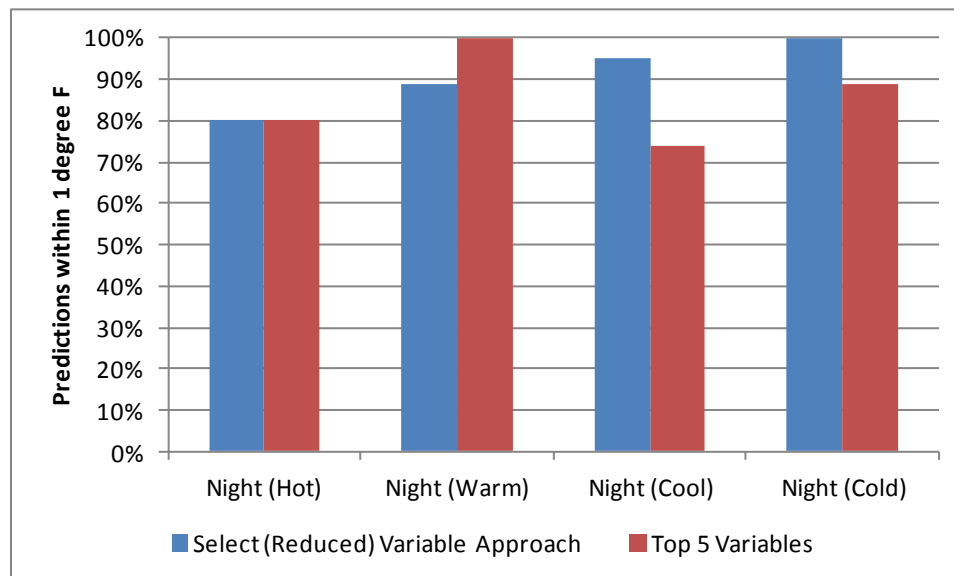


Figure VIII-14: Select (Reduced) Variable Approach vs. Top 5 Variables – Night Models

Both approaches present comparable results for model prediction accuracy with more than 70% within 1 degree F of measurements for the Day and Night models. As observed, there is a tradeoff

between the prediction accuracy and the number of variables used for model development. In some instances, the top five variables produce slightly better results. But, in most cases, the Select (Reduced) Variable Approach outperformed the top five variables; however, it does not align with all the specified points of achievement established for the variable reduction procedure. Yet, the top five variables approach uses the same variables to develop each ETHER model as well as it produces results that are more than 65% within 1 degree F of measurements, which is the baseline prediction accuracy chosen for this research. For further analysis of the Top 5 Variables Approach, in the next chapter, an evaluation of additional datasets will be examined to validate the Top 5 Variables Approach as well as the overall methodology and techniques used to construct the ETHER models.

IX. Evaluation of Additional Datasets

In this section, our methodology is tested on data obtained from the Pavilion II residence and the University of Virginia Art Museum (i.e., Bayly Museum) with the Top 5 Variables Approach. The data was supplied by Scott Martin, a facilities management representative. A year of data (i.e., CY2011) is used to create each model using the Daylight Savings Time technique for each building. The data is collected on a fifteen minute basis rather than the minute to minute basis that was used for our baseline residence. Seventy percent of the data is used to train the models, while the remaining thirty percent is used for testing. None of the data used during the training phase is used in the test and evaluation phase of the models. Validation of the methodology and newly developed evaluation of thermal response (ETHER) models has been carried out by comparing the output data with the actual data collected for the interior temperature conditions.

Both buildings are located on the grounds of the University of Virginia in Charlottesville, VA. The buildings vary in square footage, the number of rooms, and the construction materials in comparison to our baseline residence. Our baseline residence consists of two floors and a basement which includes three bedrooms, two bathrooms with a combination living/dining and kitchen area. The finished area of the home (excluding the basement) is approximately 1,244 square feet.

As previously mentioned, the top five variables are: exterior temperature, exterior humidity, interior humidity, carbon dioxide measurements, and the HVAC component. The buildings are equipped with sensors that align with the top five variables extracted from our baseline dataset, excluding the carbon dioxide measurement sensors. As a result, there are a total of four variables that is used for model development. The exterior temperature and exterior humidity data was collected by sensors at the Leake Building (i.e., the facilities building) that is also located on campus grounds.

A. Pavilion II

The Pavilion II building was built in 1826 and has a gross square footage of 7,610. The building is equipped with four zone temperature sensors and four humidity sensors. The four zone temperature sensors are averaged together to create the interior temperature variable. The four humidity sensors are also averaged together to create the humidity variable for the entire space. The chilled water and the medium temperature hot water components are viewed as the source of cooling and heating for the building, respectively. The composite function of the chilled water and medium temperature hot water components is equivalent to the HVAC component for our baseline residence.

During the data cleaning and pre-processing phases, it was observed that approximately 1% of the data was missing from the dataset. This could have been due to sensor malfunctions or recording errors. To align with the techniques used in our baseline dataset, the missing data or blanks found in the dataset were omitted.

Eight ETHER models were developed based on the Daylight Savings Time technique to predict interior temperature conditions over a 6 hour forecast horizon. Figure IX-1 shows the results of succession to remain within the 1 degree F of measurements. As observed, the Day and Night ETHER models perform quite well, resulting in predictions of more than 77%.

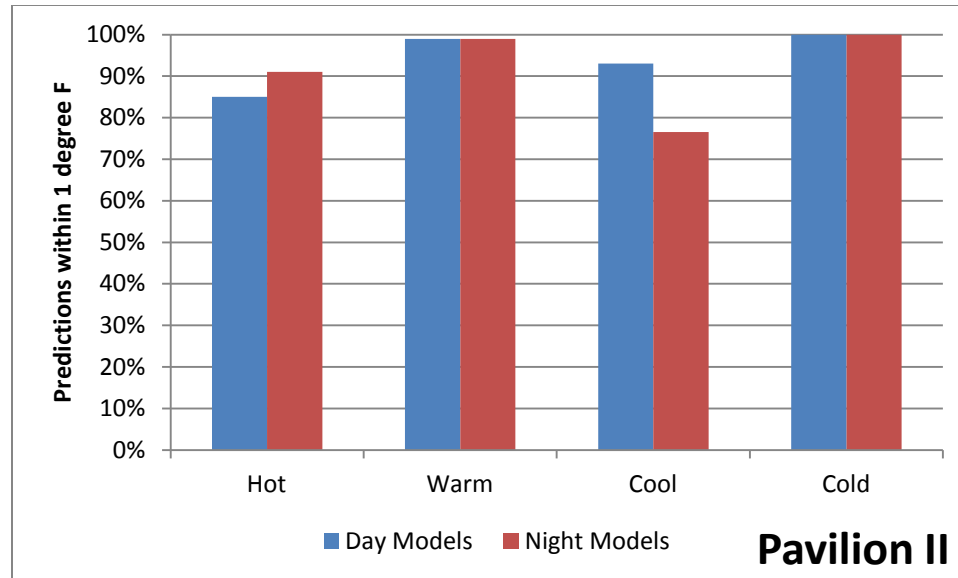


Figure IX-1: Prediction Accuracy Results for Pavilion II using ETHER Models

B. Bayly Museum

The Bayly Museum was built in 1934 and has a gross square footage of 13,469. The building is equipped with twelve zone temperature sensors and ten humidity sensors. Similar to the process described for the Pavilion II, the twelve zone temperature sensors are averaged together to create interior temperature variable and the ten humidity sensors are averaged together to create the interior humidity variable for the entire space. The chilled water and the medium temperature hot water components are also viewed as the source of cooling and heating for the building, respectively.

During the data cleaning and pre-processing phases, it was observed that less than .5% of the data was missing from the dataset. Compared to the other two datasets evaluated in this dissertation (i.e., baseline (10% of missing data) and Pavilion II (1% of missing data)), this dataset has the lowest percentage of missing data. As a result, it is anticipated that there is a notable increase in prediction accuracy using our methodology as compared with the other results.

Figure IX-2 shows the results for the Day and Night ETHER models that were developed using the data from the Bayly Museum. The eight ETHER models provide a prediction accuracy of more than 88% to remain within the 1 degree F of measurements.

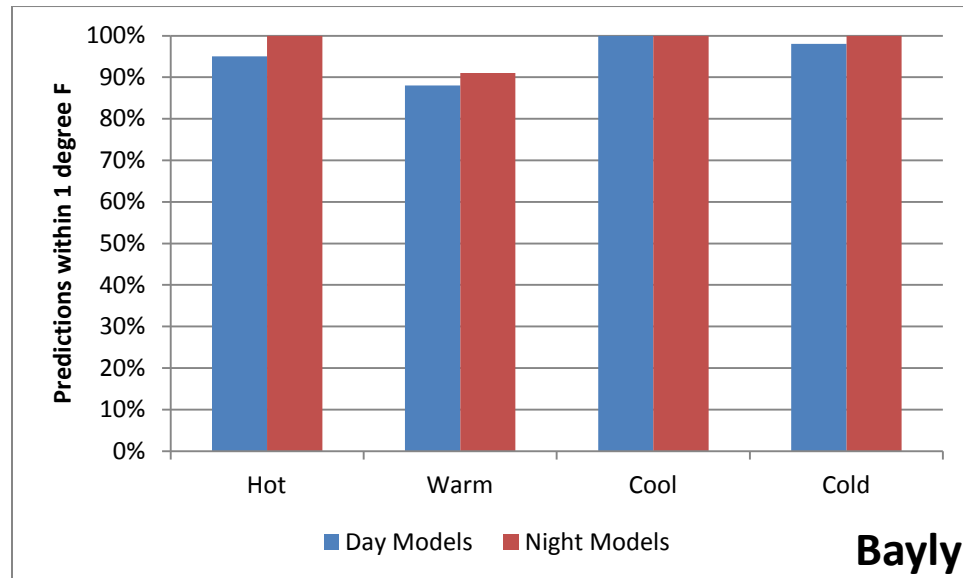


Figure IX-2: Prediction Accuracy Results for the Bayly Museum using ETHER Models

The evaluation of the additional datasets using our outside temperature and time of day conditional methodology helped to validate the originality and significance of our approach. The three buildings that were evaluated varied in size, year of development as well as structural materials. As mentioned earlier, it is difficult and quite expensive to build a “real” model of the system (i.e., thermal network design) to include thermodynamic properties of the materials, etc. Second, such a “real” model will degrade over time as the real system operation diverges from the modeled system. Therefore, our approach of using the observational data to automatically build and continually update our ETHER models, allows for addressing both of the limitations noted for a “real” modeled system.

X. Conclusions

In this dissertation, we developed a methodology and a set of forecasting evaluation of thermal response (ETHER) models (based on an Autoregressive Integrated Moving Average (ARIMA) method) to predict interior temperature conditions within a residence. We envision the ETHER models to serve as an automated agent of an intelligent framework aimed at enhancing decision making opportunities with regards to residential demand response.

Eight forecasting ETHER models have been constructed based on outside temperature and time of day conditions to predict interior conditions up to 6 hours in advance. For purposes of this research, time of day refers to day time and night time (i.e., Day (6:00 AM – 7:59 PM) and Night (8:00 PM – 5:59 AM)). The outside conditions are based on exterior temperature conditions (i.e., Hot (85°F +), Warm (70°F – 84°F), Cool (54°F – 69°F) and Cold (0°F – 53°F)). The time of day and exterior temperature ranges were selected by utilizing the history weather website for the Charlottesville area [4].

In previous work, thermal response models have been represented as thermal RC networks to predict and analyze conditions within a building structure for optimization through simulating processes of conduction, convection and radiation in the building under specific environmental conditions to achieve energy reduction [1]. Given the increasing demand of energy, price scheduling and reduction demand signals are being sent from the utility to the consumer to explore optimal energy reduction schemes that can prove to be beneficial to both parties. As a result, our anticipatory controlled thermal response model provides an opportunity to reduce electricity expenditures and maintain a level of comfort for the occupants by responding to pricing and demand reduction signals that vary at different times during the day [2].

With the notion of the smart grid, there is a need for an approach on the consumer-side which adds further intelligence in order to achieve demand response. Since demand parameters vary

from one residence to the next, using observational data for model development is very useful to factor in seasonal variation, deterioration of structural materials and the fluctuation of heating and cooling requirements [28]. The methodology and forecasting ETHER models present a feasible way to predict interior conditions over a short time span and assist in studying the changes in energy consumption to achieve demand response. The model structure of the ETHER models identifies the controllable and uncontrollable variables that impact the interior conditions during the outside temperature and time of day conditions. Ultimately, the multivariate time series prediction models can become a basis for predictive control aimed at optimizing load profiles to reduce and/or manage energy consumption during demand response periods.

The research shows that the ARIMA time series method can be used to obtain short-term thermal response forecasts to provide insight on comfort levels as well as assisting to identify the contributing controllable loads that consume energy and influence interior temperature conditions. The advantage of this research is the ability to forecast interior temperature conditions and identify contributing loads based on consumption recognition conducted in the residence during varying outside temperature and time of day conditions. Validation of the forecasting models has been carried out by comparing the models' output with the actual data collected for the interior temperature conditions to be within 1 degree F of measurements.

The evaluation approaches and techniques chosen to showcase the accuracy of the interior temperature conditions provide a successful outlook into which methods produced the best results on a consistent basis. Throughout the dissertation, case studies were shown to compare techniques and evaluation approaches. The techniques highlighted in this work compared the manipulation of the datasets with an increased number of developed models. Out of the five model development techniques that were mentioned, the Daylight Savings Time technique resulted with the best results. The model reduction approaches presented in this work detailed a variable reduction procedure to preclude highly correlated variables and minimize the quantity of

data and variables used for model development. Out of the three model reduction approaches, the Select (Reduced) Variable Approach provided the best means of prediction accuracy for the interior temperature conditions to be within 1 degree F, more than 66% of the time for the Day Models and more than 79% for the Night Models, over a 6 hour forecast horizon. In addition, the Select (Reduced) Variable Approach resulted in the fewest number of predictors to be used for model development without sacrificing prediction accuracy.

Due to the accurate predictions obtained from the combination of the Daylight Savings Time technique and the Select (Reduced) Variable Approach, an investigation was conducted to see how comparable the results would be with a more increased reduction of information. The Top 5 Variables Approach was then investigated for further analysis purposes on additional datasets. The Top 5 Variables Approach aligned with one of the goals of our research which was to minimize the number of variables and reduce the magnitude of data used for model development and forecasting. As shown in Chapter X, the top five variables using the Daylight Savings Time technique produced interior temperature conditions to be within 1 degree F, more than 85% of the time for the Day Models and more than 77% for the Night Models for Pavilion II. In addition, the top five variables using the Daylight Savings Time technique produced interior temperature conditions to be within 1 degree F, more than 88% of the time for the Day Models and more than 91% for the Night Models. After further analysis of the top five variables within each dataset, it was observed that as the number of missing data decreased within the dataset, the prediction accuracy for the Day and Night ETHER models increased.

Our results indicate that the outside temperature and time of day conditional methodology using the Daylight Savings Time technique provides a means for prediction accuracy that is acceptable for model development and forecasting. The main practical conclusion is that developing the ETHER models from observational data assists to continuously adjust demand parameters to align with conditions within the residence to predict the interior temperature.

As a product of this research, we were able to submit two manuscripts for journal publication

which are currently under review. These manuscripts highlighted the research methodology and model reduction approaches used to develop the ETHER models that predict the interior temperature conditions of the residence based on observational data. The titles of the two manuscripts are:

- ❖ Submitted/Under review: Craft, S., Williams, R., Marshall, P., Dugan J.B., “Development of Forecasting Time Series ARIMA Thermal Response Models to achieve Residential Demand Response”
- ❖ Submitted/Under review: Craft, S., Williams, R., Dugan J.B., “Model Reduction Approaches to Achieve Residential Demand Response”

A. Contributions

Based on the results achieved in this dissertation, the model development and forecasting methods present a compelling case for the applicability of this work thus far. The contributions that arise from this work include:

- ❖ The developed modeling methodology identifies the exact load appliances that impact the thermal response of the residence. The identification of the controlled independent variables is essential for testing the optimal settings which are required for achieving demand response. This is different from the approaches proposed in the literature which focus mainly on pre-programmed strategies and customer preferences for operating appliances.
- ❖ The automation of the modeling process allows for self-constructing and self-parameterizing models to be created based on data collected from the residence. This eliminates the human interference factor and captures the essential data automatically for constant refinement. Consequently, based on outside temperature and time of day conditions, a model can be selected to give accurate predictions for interior temperature conditions.

- ❖ The variable reduction procedure reduces the number of sensors placed in the home to monitor activity within the structure. This procedure contributes to the end-users' knowledge regarding strategic placement of sensors to obtain useful information on the conditions of the residence.
- ❖ The selection of the top five variables allows the consumer to reduce over-monitoring of his/her residence, decreases the cost of purchasing sensors for installation and presents an analyst with a more straight-forward approach to identify which variables are highly significant.

B. Future Work

In this section, potential ideas for an expansion of our work will be provided. For future work, our methodology and modeling approach techniques can be expanded by:

- ❖ Coding the ARIMA modeling process in an embedded computing device to predict interior temperature conditions within a residence based on observational data collected from sensors. In Appendix A, information on the statistical algorithms is provided at a high level to serve as a roadmap of the process that needs to be executed within the embedded environment. This embedded computing device can be programmed to develop the ETHER models to serve as a prototype. One area of exploration could possibly be to investigate the processing speed and storage space required for continuous refinement of the models.
- ❖ Writing a publication of a proposal to utilize the Intelligent Demand Energy Adaptive System (IDEAS) framework to bridge together all the demand response models that are currently present in literature. At present, researchers have proposed various demand response models to reduce energy consumption and achieve cost savings for the consumer. The IDEAS framework represents an intelligent agent that can act autonomously on behalf of the consumer to make decisions regarding demand reduction

to achieve demand response. Bridging together all the literary findings will also serve as a survey for the demand response field of study.

- ❖ Performing an in-depth investigation on how humidity impacts the thermal comfort of occupants. Researchers can investigate psychometric charts and the effective temperature factors to obtain a basic understanding in this area. ETHER models were developed with both the interior humidity and interior temperature serving as dependent variables; however, in our body of work, these models were tested to observe the impact on model prediction results with having the interior humidity variable as a dependent variable rather than an independent variable.
- ❖ Performing a sensitivity analysis on the variables used for model development. There are total of thirty-one sensors used to collect data. Out of the thirty-one variables/sensors, five were selected as the top variables that influence the interior temperature conditions. A sensitivity analysis can be conducted to increase the understanding of why these five variables were selected and understand their relationship or influential importance to the output variable (i.e. interior temperature).

C. Summary

The scope of our work was focused on developing a finite set of low fidelity models using observational data to predict interior temperature conditions within the residence based on the variations of controllable and uncontrollable factors to aid the consumer in planning for energy consumption. The smart grid technologies enable utilities as well as consumers to have an open channel of communication in an effort to achieve demand response, especially targeting the residential sector. In the traditional utility environment, there are no mechanisms in place to help the consumer alter his/her consumption patterns. Literature has proven that if consumers are provided with mechanisms to help them observe their consumption, they are more acceptable to altering their energy usage. With the adoption of the Smart Grid, mechanisms are provided to

help the consumer alter their consumption patterns to achieve demand response. The ETHER models are viewed as mechanisms to enhance the smart grid paradigm on the consumption side. With a better understanding of consumption patterns, utilities are able to design demand response programs to benefit the residential end-user and take advantage of the best practices within the smart grid paradigm. In addition, consumers are more equipped to make decisions regarding the reduction of energy usage of specific appliances and devices when mechanisms are provided to observe their consumption patterns.

The ETHER models rely on the transmission of information through the advanced metering infrastructure to add value to the consumer. The external signals which are transmitted from the utility are received by an in-home controller or smart meter which sends alerts to the ETHER models to predict interior temperatures based on changes made to contributing loads within the building. The interior temperature is considered as the primary indicator for the achievement of thermal comfort. Thermal comfort is determined by a number of factors (i.e., gender, health status, humidity, etc.). Due to this fact, estimating thermal comfort levels is rather complex because of the subjective nature of the influential variables and the response of the individuals. Thus, using an outside temperature and time of day methodology coupled with the development of ETHER models using the ARIMA process presents a viable case for utilization of interior temperature prediction.

The ETHER models are designed to add intelligence to the consumer-side to aid in the management of energy usage. Based on the variations of energy consumption of the contributing loads, the ETHER models are designed to predict interior temperature conditions so that real-time decisions can be made. Automation of the reduction of energy consumption can be achieved in an intelligent framework similar to the framework suggested in this dissertation (e.g. IDEAS). Having the ETHER models act as automated agents eliminates the human factor and increases decision making opportunities in an intelligent framework.

For residential consumers, anticipatory models are needed to maximize demand-side management to get the greatest value out of the smart grid. The ETHER models are a set of forecasting models using an outside temperature and time of day condition methodology to help make decisions on minimizing consumption, achieving thermal comfort levels and achieving cost savings on the consumer-side. Being able to test the methodology and ETHER models using data from additional buildings helped to validate the adaptability and practicality of our work.

XI. Bibliography

1. Reuters, T.: Global energy demand seen up 44 percent by 2030: EIA. (2009)
2. O. Ameri Sianaki, O.: Intelligent Decision Support System for Including Consumers' Preferences in Residential Energy Consumption in Smart Grid. In : Second International Conference on Computational Intelligence, Modelling and Simulation (2010)
3. Farhangi, H.: The Path of the Smart Grid. IEEE Power and Energy Magazine, 18-28 (2010)
4. Ipakchi, F.: Demand Response as a Market Response Under the Smart Grid Paradigm. IEEE Transactions on Smart Grid 1(1), 82-88 (2010)
5. Mouftah, M.: The Impact of Smart Grid Residential Energy Management Schemes on the Carbon Footprint of the Household Electricity Consumption. In : IEEE Electric Power and Energy Conf. (2010)
6. Q. Binh Dam, S.: Intelligent Demand Response Scheme for Customer Side Load Management. In : IEEE Energy 2030 Conf., Atlanta (2008)
7. A. Iwayemi, P.: Knowing When to Act: An Optimal Stopping Method for Smart Grid Demand Response. IEEE Network 25(5), 44-49 (2011)
8. Lee, S.: A Residential Consumer-Centered Load Control Strategy in Real-Time Electricity Pricing Environment. In : 39th North American Power Symposium (2007)
9. Y. Guo, R.: A Simulator for Self-Adaptive Energy Demand Management. In : Second IEEE International Self-Adaptive and Self-Organizing Systems Conf (2008)
10. Mathews, C.: Efficient, Steady State Solution of a Time Variable RC Network, for Building Thermal Analysis. Building and Environment 27(3), 279-287 (1992)
11. Mathews, E. H., Richards, P. G., Lombard, C.: A first-order thermal model for building design. Energy and Buildings 21(2), 133-145 (1994)
12. Mathews, E. H., Rousseau, P. G., Richards, P. G., Lombard, C.: A Procedure to Estimate the Effective Heat Storage Capacity of a Building., 179-188 (1991)
13. Administration, E.: U.S. DoE. In: U.S. Household Electricity Report. (Accessed July 14, 2005) Available at: www.eia.doe.gov
14. Energy, U. S.: Energy Efficiency Trends in Residential and Commercial Buildings. (Accessed October 2008) Available at: http://apps1.eere.energy.gov/buildings/publications/pdfs/corporate/bt_stateindustry.pdf

15. Energy, U. S.: The Smart Grid: An Introduction. (Accessed 2008) Available at:
http://www.oe.energy.gov/DocumentsandMedia/DOE_SG_Book_Single_Pages_percent281_percent29.pdf
16. Vojdani, A.: Smart Integration. (2008)
17. Hayes, D.: Hypothesis Testing. In : Code Of Best Practice Experimentation. CCRP Publication Series (2003) 22-23
18. Hayes, D.: Hypothesis Testing. In : Campaigns of Experimentation: Pathways to Innovation and Transformation. CCRP Publication Series (2005) 74
19. J. Contreras, R.: ARIMA Models to Predict Next-Day Electricity Prices. IEEE Trans. Power Systems 18(3), 1014-1020 (Aug. 2003)
20. I. Razak, M.: Short Term Load Forecasting Using Data Mining Technique. In : IEEE 2nd International Power and Energy Conference (2008)
21. Energy Monitoring System. Available at:
<https://ecomod.seas.virginia.edu/ecomod/int2/home.php>
22. Kidd, B.: ecoMOD | Data Monitoring System., Charlottesville (2007)
23. Nugent, X.: Partitioning Time Series Sensor Data for Activity Recognition. In : International Information Technology and Applications in Biomedicine Conference (2009)
24. Vuppala, S. K., Padmanabh, K., Bose, S. K., Paul, S.: Incorporating fairness within Demand response programs in smart grid. IEEE PES Innovative Smart Grid Technologies (ISGT), 1-9 (January 2011)
25. Rahimi, F., Ipakchi, A.: Demand Response as a Market Resource Under the Smart Grid Paradigm., 82-88 (2010)
26. Sianaki, O. A., Hussain, O., Dillon, T., Tabesh, A. R.: Intelligent Decision Support System for Including Consumers' Preferences in Residential Energy Consumption in Smart Grid. In : Second International Conference on Computational Intelligence, Modelling and Simulation, pp.154-159 (2010)
27. Ipakchi, A.: Grid of the Future. IEEE Transactions on Smart Grid 7(2) (2009)
28. Wang, J., Biviji, M., Wang, W. M.: Case studies of smart grid demand response programs in North America. In : IEEE PES Innovative Smart Grid Technologies (ISGT), pp.1-5 (2011)
29. Conejo, A. J., Morales, J. M., Baringo, L.: Real-Time Demand Response Model., 236-242 (2010)

30. Hobby, J.: Constructing Demand Response Models for Electric Power Consumption. In : First IEEE International Smart Grid Communications Conf. (2010)
31. O'Neill, D., Levorato, M., Goldsmith, A., Mitra, U.: Residential Demand Response Using Reinforcement Learning. In : 2010 First IEEE International Smart Grid Communications (SmartGridComm) Conference, pp.409-414 (2010)
32. Sutton, R., Barto, A.: Reinforcement Learning. (1998)
33. Shao, S., Pipattanasomporn, M., S., R.: Development of Physical-Based Demand Response-enabled Residential Load Models. IEEE PES Trans. Power Systems (2011)
34. Iwayemi, A., Yi, P., Dong, X., Zhou, C.: Knowing When to Act: An Optimal Stopping Method for Smart Grid Demand Response. IEEE Network 25(5), 44-49 (2011)
35. Ferguson, T. S.: Optimal Stopping and Applications. Available at: <http://www.math.ucla.edu/~tom/Stopping/Contents.html>
36. Roe, C., Meliopoulos, S., Entriken, R., Chhaya, S.: Simulated demand response of a residential energy management system. In : 2011 IEEE Energytech, pp.1-6 (2011)
37. Han, J., Piette, M. A.: Solutions for Summer Electric Power Shortages: Demand Response and its Applications in Air Conditions and Refrigerating Systems., 1-4 (2008)
38. Albadi, M. H., El-Saadany, E. F.: Demand Response in Electricity Markets: An Overview. In : IEEE Power Engineering Society General Meeting (2007)
39. Dam, Q. B., Mohagheghi, S., Stoupis, J.: Intelligent Demand Response Scheme for Customer Side Load Management. In : IEEE Energy 2030 Conference, Atlanta, pp.1-7 (2008)
40. Schneider, K. P., Fuller, J. C., Chassin, D.: Analysis of Residential Demand Response and double-auction markets. (2011)
41. Richards, E.: A Tool for Predicting Hourly Air Temperatures and Sensible Energy Loads in Buildings at Sketch Design Stage. (1989)
42. Braun, J., Chaturvedi, N.: An Inverse Gray-Box Model for Transient Building Load Prediction., 73-99 (2002)
43. Zheng, A.: Data Mining for Prediction of Wind Farm Power Ramp Rates. (2008)
44. Contreras, J., Espinola, R., Nogales, F., Conejo, A.: ARIMA Models to Predict Next-Day Electricity Prices. IEEE Trans. Power Systems 18(3), 1014-1020 (Aug. 2003)

45. Razak, I. A. B. W. A., bin Majid, S., Rahman, H. A., Hassan, M. Y.: Short Term Load Forecasting Using Data Mining Technique. In : IEEE 2nd International Power and Energy Conference, pp.139-142 (2008)
46. Hubert, T., Grijalva, S.: Realizing smart grid benefits requires energy optimization algorithms at residential level. In : IEEE PES Innovative Smart Grid Technologies (ISGT), pp.1-8 (2011)
47. Dupont, B., De Jonghe, C., Kessels, K., Belmans, R.: Short-term consumer benefits of dynamic pricing. In : 2011 8th International European Conference, pp.216-221 (2011)
48. Widergren, S., Subbarao, K., Chassin, D., Fuller, J., Pratt, R.: Residential real-time price response simulation. In : IEEE Power and Energy Society General Meeting, pp.1-5 (2011)
49. Study, P.: Bringing the Smart Grid to the Smart Home - It's Not Only About the Meter. In: SmartGridNews. Available at:
http://www.smartgridnews.com/artman/publish/Technologies_Metering_News/Bringing-the-Smart-Grid-to-the-Smart-Home-It-s-Not-Only-About-the-Meter-1711.html
50. Guo, Y., Li, R., Poulton, G., Zeman, A.: A Simulator for Self-Adaptive Energy Demand Management. In : Second IEEE International Self-Adaptive and Self-Organizing Systems Conf, pp.64-73 (2008)
51. Saele, H., Grande, O. S.: Demand Response From Household Customers: Experiences From a Pilot Study in Norway., 102-109 (2011)
52. Venkatesan, N., Solanki, J., Solanki, S. K.: Market Optimization for Microgrid with Demand Response Model. In : North American Power Symposium (NAPS), pp.1-6 (2011)
53. Webexhibits: Daylight Savings Time - When we change our clocks. In: Daylight Savings Time. Available at: <http://www.webexhibits.org/daylightsaving/b2.html>
54. Weather Underground. (Accessed 2011) Available at: www.wunderground.com
55. Design Guide - Thermal Comfort. In: Engineering Guide - Thermal Comfort. Available at: www.price-hvac.com/Catalog/Section_I/html/DG_ThermalComfort.htm
56. Zhang, X., Fields, R. L., Abreu, K.: Financial Benefits of Implementing Demand Response in CAISO Market. In : Power and Energy Society General Meeting, 2010 IEEE, pp.1-3 (2010)
57. Chu, C.-M., Jong, T.-L.: Enthalpy estimation for thermal comfort and energy saving in air conditioning system., 1620-1628 (2008)
58. Wan, J. W., Yang, K., Zhang, W. J., Zhang, J. L.: A new method of determination of indoor

- temperature and relative humidity with consideration of human thermal comfort., 411-417 (2009)
59. Architecture, S.: Evaluating ecoMOD., University of Virginia, Charlottesville (2005)
 60. University, C.: Ambient Environment: Thermal Comfort. In: Cornell University Ergonomics Web. Available at:
<http://ergo.human.cornell.edu/studentdownloads/DEA3500notes/Thermal/thcomnotes1.html>
 61. Khasnabis, S., Lyoo, S. H.: Use of Time Series Analysis to Forecast Truck Accidents. Transportation research record (1989)
 62. Kumar, K., Yadav, A. K., Singh, M. P., Hassan, H., Jain, V. K.: Forecasting Daily Maximum Surface Ozone Concentrations in Brunei Darussalam – An ARIMA Modeling Approach. Journal of the Air and Waste Management Association 54(5), 809-814 (July 2004)
 63. Jenkins, J. B.: Time Series Analysis Forecasting and Control. Holden-Day, San Francisco (1976)
 64. Wang, W., Niu, Z.: Time Series Analysis of NASDAQ Composite Based on Seasonal ARIMA Model. In : International Management and Service Science Conference, pp.1-4 (2009)
 65. Cho, M. Y., Hwang, J. C., Chen, C. S.: Customer Short Term Load Forecasting by using ARIMA Transfer Function Model. In : International Energy Management and Power Delivery Conference, vol. 1, pp.317-322 (1995)
 66. SPSS Statistics 16.0. (2011)
 67. Morris, W., ed.: The American Heritage Dictionary of the English Language. Houghton Mifflin Company, Boston (1979)
 68. Microsoft: Partitioning Data into Training and Testing Sets (Analysis Services - Data Mining). In: Microsoft Technet Website. Available at: <http://technet.microsoft.com/en-us/library/bb895173.aspx>
 69. Kusiak, A., Zheng, H.: Data Mining for Prediction of Wind Farm Power Ramp Rates. In : IEEE International Sustainable Energy Technologies Conference, pp.1099-1103 (2008)
 70. NIST: Motivation: How do we Construct a Goodness-of-fit Metric for a Model? In: Engineering Statistics Handbook. Available at:
<http://www.itl.nist.gov/div898/handbook/pri/section5/pri5992.htm>
 71. Multicollinearity. In:

- <http://www.chsbs.cmich.edu/fattah/courses/empirical/multicollinearity.html>.
72. Blalock, H. M.: Correlated Independent Variables: The Problem of Multicollinearity., 233-237 (1963)
 73. Lani, J.: Dissertation Statistics Help. Available at:
<http://statisticssolutions.blogspot.com/2008/06/what-does-bivariate-correlation.html>
 74. Consultation, R.: Identifying Multicollinearity in Multiple Regression Statistics Help for Dissertation Students & Researchers. Available at:
<http://www.researchconsultation.com/multicollinearity-regression-spss-collinearity-diagnostics-vif.asp>
 75. Sun, P., Tan, J., Wang, H. L., Fang, Y. S., Xu, S., Yan, W. N., Liu, H.: QSAR Study on Acute Toxicity to Photobacterium Phosphoreum of Fluorobenzene Derivatives. In : 2011 5th International Conference on Bioinformatics and Biomedical Engineering, (iCBBE) , pp.1-4 (2011)
 76. Yakubu, A.: Fixing Collinearity Instability in the Estimation of Body Weight from Morpho-Biometrical Traits of West African Dwarf Goats., 61-66 (2009)
 77. Rook, A. J., Dhanoa, M. S., Gill, M.: Prediction of the voluntary intake of grass silages by beef cattle. 2. Principal component and ridge analysis., 439-454 (1990)
 78. Gill, J. L.: Outliers and influence in multiple regression., 161-175 (1986)
 79. O'Brien, R.: A Caution Regarding Rules of Thumb for Variance Inflation Factors., 673-690 (2007)
 80. University, T.: Interpreting Regression Output. Available at:
http://dss.princeton.edu/online_help/analysis/interpreting_regression.htm
 81. Alahmer, A., Omar, M. A., Mayyas, A., Dongri, S.: Effect of relative humidity and temperature control on in-cabin thermal comfort state: Thermodynamic and psychometric analysis., 2636-2644 (2011)
 82. 55-2004, A.: Thermal Environmental Conditions for Human Occupancy. (2004)
 83. Gerlich, V.: Validation heat transfer model by experimental measurement. In : 2011 12th International Carpathian Control Conference (ICCC) , pp.131-134 (2011)
 84. Kontoleon, K. J., Bikas, D. K.: The influence of the zone's indoor temperature settings on the cooling/heating loads for fixed and controlled ventilation., 75-86 (2006)
 85. Leon-Garcia, A.-H.: Optimal Residential Load Control With Price Prediction in Real-Time

- Electricity Pricing Environments. (2010)
86. Perdomo, R., Banguero, E., Gordillo, G.: Statistical modeling for global solar radiation forecasting in Bogotá. In : 2010 35th IEEE Photovoltaic Specialists Conference (PVSC), pp.2374-2379 (2010)
 87. Sfetos, A.: A novel approach for the forecasting of the mean hourly wind speed time series. (2002)
 88. Identifying Multicollinearity in Multiple Regression. In:
<http://www.researchconsultation.com/multicollinearity-regression-spss-collinearity-diagnostics-vif.asp>. (Accessed 2007)
 89. Richards, E. H.: A Tool for Predicting Hourly Air Temperatures and Sensible Energy Loads in Buildings at Sketch Design Stage. *Energy and Buildings* 14(1), 61-80 (1989)
 90. McGee, R. Y.: *Introduction to Time Series Analysis and Forecasting with Applications of SAS and SPSS*. Academic Press, Inc., New York (2000)
 91. Group, T.: *Quantifying Demand Response Benefits in PJM*. Cambridge (2007)
 92. Torres, J. L., Garcia, A., De Blas, M., De Francisco, A.: Forecast of hourly average wind speed with ARMA models in Spain., 65-77 (2005)
 93. Perfumo, C., Ward, J., Braslavsky, J.: Reducing energy use and operational cost of air conditioning systems with multi-objective evolutionary algorithms. In : *IEEE Congress on Evolutionary Computation*, vol. 22, p.57 (2010)
 94. McArthur, S. D. J., Davidson, E. M.: Concepts and Approaches in Multi-agent Systems for Power Applications. In : *13th International Intelligent Systems Application to Power System Conference*, p.5 (2005)
 95. Hindi, H., Greene, D., Laventall, C.: Coordinating regulation and demand response in electric power grids using multirate model predictive control., 1-8 (2011)
 96. Wang, X., Jayantilal, A.: Dispatch wind generation and demand response., 1-4 (2011)
 97. Vos, A.: Effective business models for demand response under the Smart Grid paradigm. In : *Power Systems Conference and Exposition, 2009. PSCE '09. IEEE/PES*, p.1 (2009)
 98. Tiptipakorn, S., Lee, W.-J.: A Residential Consumer-Centered Load Control Strategy in Real-Time Electricity Pricing Environment. In : *39th North American Power Symposium, 2007. NAPS '07* , pp.505-510 (2007)
 99. Torriti, J., Hassan, M. G., Leach, M.: Demand response experience in Europe: Policies,

- programmes and implementation. *Energy* 35(4), 1575-1583 (2010)
100. Sane, H., Guay, M.: Minmax dynamic optimization over a finite-time horizon for building demand control. In : American Control Conference, pp.1469-1474 (2008)
 101. Mohsenian-Rad, A.-H., Leon-Garcia, A.: Optimal Residential Load Control With Price Prediction in Real-Time Electricity Pricing Environments., 120-133 (2010)
 102. Mathieu, J. L., Price, P. N., Kiliccote, S., Piette, M. A.: Quantifying Changes in Building Electricity Use, With Application to Demand Response., 507-518 (2011)
 103. Katipamula, S., Lu, N.: Evaluation of residential HVAC control strategies for demand response programs., 535-546 (2006)
 104. Erol-Kantarci, M., Mouftah, H. T.: The impact of smart grid residential energy management schemes on the carbon footprint of the household electricity consumption. In : 2010 IEEE Electric Power and Energy Conference (EPEC), pp.1-6 (2010)
 105. Chassin, D. P., Du, P., Fuller, J. C.: The potential and limits of residential demand response control strategies. In : 2011 IEEE Power and Energy Society General Meeting, pp.1-6 (2011)
 106. Abreu, J., Camara Pereira, F., Vasconcelos, J., Ferrao, P.: An approach to discover the potential for demand response in the domestic sector. In : Innovative Technologies for an Efficient and Reliable Electricity Supply (CITRES), pp.240-245 (2010)
 107. Fernandes, F., Sousa, T., Silva, M., Morais, H., Vale, Z., Faria, P.: Genetic algorithm methodology applied to intelligent house control. In : 2011 IEEE Computational Intelligence Applications In Smart Grid (CIASG) Symposium, pp.1-8 (2011)
 108. Ota, N., Ahrens, S., Redfern, A., Wright, P., Yang, X.: An Application-Driven Architecture for Residential Energy Management with Wireless Sensor Networks. In : IEEE International Mobile Adhoc and Sensor Systems (MASS) Conference, pp.639-644 (2006)
 109. Houwing, M., Negenborn, R. R., De Schutter, B.: Economic advantages of applying model predictive control to distributed energy resources: The case of micro-CHP systems. In : 16th Mediterranean Control and Automation Conference, pp.1550-1555 (2008)
 110. Shahgoshtasbi, D., Jamshidi, M.: Energy efficiency in a smart house with an intelligent neuro-fuzzy lookup table. In : 6th International System of Systems Engineering (SoSE) Conference, pp.288-292 (2011)
 111. Ma, Y., Anderson, G., Borrelli, F.: A distributed predictive control approach to building temperature regulation. In : American Control Conference (ACC), pp.2089-2094 (2011)

112. Cherkassky, V., Chowdhury, S. R., Landenberger, V., Tewari, S., Bursch, P.: Prediction of electric power consumption for commercial buildings. In : The 2011 International Joint Neural Networks (IJCNN) Conference, pp.666-672 (2011)
113. Olofsson, T., Andersson, S.: Long-term energy demand predictions based on short-term measured data. *Energy and Buildings* 33(2), 85-91 (2001)
114. Nogales, F. J., Contreras, J., Conejo, A. J., Espinola, R.: Forecasting next-day electricity prices by time series models., 342-348 (2002)
115. Zheng, H., Xie, L., Zhang, L.: Intelligence Model for the Sensitivity Analysis of Financial Time Series. In : Third International Conference on Natural Computation, 2007. ICNC 2007. , vol. 2, pp.533-537 (2007)
116. Liao, T. W.: Clustering of time series data - a survey., 1857-1874 (2005)
117. Ding, J., Chen, X., Ding, F.: Time series AR model parameter estimation with missing observation data. In : 7th World Congress on Intelligent Control and Automation, pp.5632-5636 (2008)
118. Brown, B. G., Katz, R. W., Murphy, A. H.: Time series models to simulate and forecast wind speed and wind power., 1184-1195 (1984)
119. Lee, J.-Y., Wu, T.-N.: Discovery of Temporal Variation of Arsenic in a Historical Blackfoot Disease Territory by Time Series Analysis. In : Sixth International Conference on Fuzzy Systems and Knowledge Discovery , vol. 2, pp.542-546 (2009)
120. Staackmann, M., Bor, Y., Yun, D. Y. Y.: Dynamic driving cycle analyses using electric vehicle time-series data. In : 32nd Intersociety Energy Conversion Engineering Conference, vol. 3, pp.2014-2018 (1997)
121. Li, F., Luan, P.: ARMA model for predicting the number of new outbreaks of newcastle disease during the month. In : IEEE International Computer Science and Automation Engineering (CSAE) Conference, vol. 4, pp.660-663 (2011)
122. Yao, J. T.: Sensitivity analysis for data mining. In : 22nd International Conference of the North American Fuzzy Information Processing Society, pp.272- 277 (2003)
123. Guo, J., Wang, H., Gao, Y., Zhu, Z.: A New Data Mining Method of Iterative Dimensionality Reduction Derived from Partial Least-Squares Regression. In : Third International Intelligent Information Technology Application Symposium, vol. 2, pp.471-474 (2009)
124. Larose, D.: *Data Mining Methods and Models.*, 93-154 (2006)

125. El Hag, H. M. A., Sharif, S. M.: An adjusted ARIMA model for internet traffic. In : AFRICON, pp.1-6 (2007)
126. Li, J., Hu, C., Xu, D., Xiao, J., Wang, H.: Application of time-series autoregressive integrated moving average model in predicting the epidemic situation of newcastle disease. In : 2010 World Automation Congress (WAC) , pp.141-144 (2010)
127. Lihua, N., Xiaorong, C., Qian, H.: ARIMA model for traffic flow prediction based on wavelet analysis. In : 2010 2nd International Information Science and Engineering (ICISE) Conference, pp.1028-1031 (2010)
128. Meesad, P., Srikhacha, T.: Universal Data Forecasting with an Adaptive Approach and Seasonal Technique. In : International Computational Intelligence for Modelling, Control and Automation, 2006 and International Conference on Intelligent Agents, Web Technologies and Internet Commerce Conference, p.66 (2006)
129. Ji, W., Chan, C., Loh, J., Choo, F., Chen, L. H.: Solar radiation prediction using statistical approaches. In : 7th International Information, Communications and Signal Processing Conference , pp.1-5 (2009)
130. Piña-Monarez, M.: A NEW THEORY IN MULTIPLE LINEAR REGRESSION., 310-316 (2011)
131. Bin, X., Cang, Z.: The Application of Multiple Regression Analysis Forecast in Economical Forecast: The Demand Forecast of Our Country Industry Lavation Machinery in the Year of 2008 and 2009. In : Second International Workshop on Knowledge Discovery and Data Mining, 2009. WKDD 2009. , pp.405-408 (2009)
132. Lin, D., Foster, D. P.: VIF Regression: A Fast Regression Algorithm for Large Data. In : Ninth IEEE International Conference on Data Mining, 2009. ICDM '09. , pp.848-853 (2009)
133. White, J. A., Reichmuth, R.: Simplified method for predicting building energy consumption using average monthly temperatures. In : Energy Conversion Engineering Conference, 1996. IECEC 96., vol. 3, pp.1834-1839 (1996)
134. Lindberg, R., Binamu, A., Teikari, M.: Five-year data of measured weather, energy consumption, and time-dependent temperature variations within different exterior wall structures., 495-501 (2004)
135. E. Mathews, P. G.: A first-order thermal model for building design. (1994)
136. Chu, C., Jong, T.: Enthalpy estimation for thermal comfort and energy saving in air conditioning system., 1620-1628 (2008)

137. Han, J., Zhang, G., Zhang, Q., Liu, J., Tian, L., Zheng, C., Hao, J., Lin, J., Liu, Y., Moschandreas, D.: Field study on occupants' thermal comfort and residential thermal environment in a hot-humid climate of China., 4043-4050 (2007)
138. Karyono, T. H.: Report on thermal comfort and building energy studies in Jakarta - Indonesia., 77-90 (2000)
139. Peeters, L., Dear, R., Hensen, J., D'haeseleer, W.: Thermal comfort in residential buildings: Comfort values and scales for building energy simulation., 772-780 (2009)
140. Djongyang, N., Tchinda, R., Njomo, D.: Thermal comfort: A review paper., 2626-2640 (2010)
141. Hoffman, M., Feldman, M.: Calculation of the Thermal Response of Buildings by the Total Thermal Time Constant Method., 71-85 (1981)

XII. Appendix A - Embedded Environment

As described in the Research Methodology chapter, the Autoregressive Integrated Moving Average (ARIMA) models for this research are developed in a software application tool known as Statistical Package for Social Sciences (SPSS). SPSS has statistical algorithms programmed behind its user interface to determine the stationarity of the data, identify the appropriate model, determine the estimation of the parameters, evaluate the model accuracy and apply the developed model to forecast on additional datasets. For the continuation of our research for future work, we envision the fundamental procedure of the ARIMA model development and prediction to be conducted in an embedded environment without the utilization of the SPSS tool.

An embedded environment can be defined as a consortium environment that encompasses small computerized parts that are designed to perform specific tasks for an overall general purpose. For example, HVAC systems use thermostats that can adjust the thermal comfort levels for an occupant based on his/her standard of living specifications (i.e., environmental conditions, time of day, etc.). The general purpose of the embedded environment in the context of our research would be to predict the interior temperature conditions of the residence based on observational data collected from sensors using an ARIMA method. The functionality of the embedded device can be programmed with C++, Assembly Language or any other applicable programming platform.

A roadmap for building ARIMA-based evaluation of thermal response (ETHER) models at the function level within an embedded environment will be described in this appendix. The various mathematical theories that can be used or applied to form the ARIMA models will be mentioned; however, a deep dive of each theory will not be provided within this dissertation. A brief overview will be provided which describes each step required to develop the ARIMA models using the statistical processes. The rationale and procedures rendered in the selection of the data and processes for model development and forecasting can be found in Chapter VI and Chapter VIII, respectively. Figure VII-1 shows a flowchart of establishing the ARIMA model.

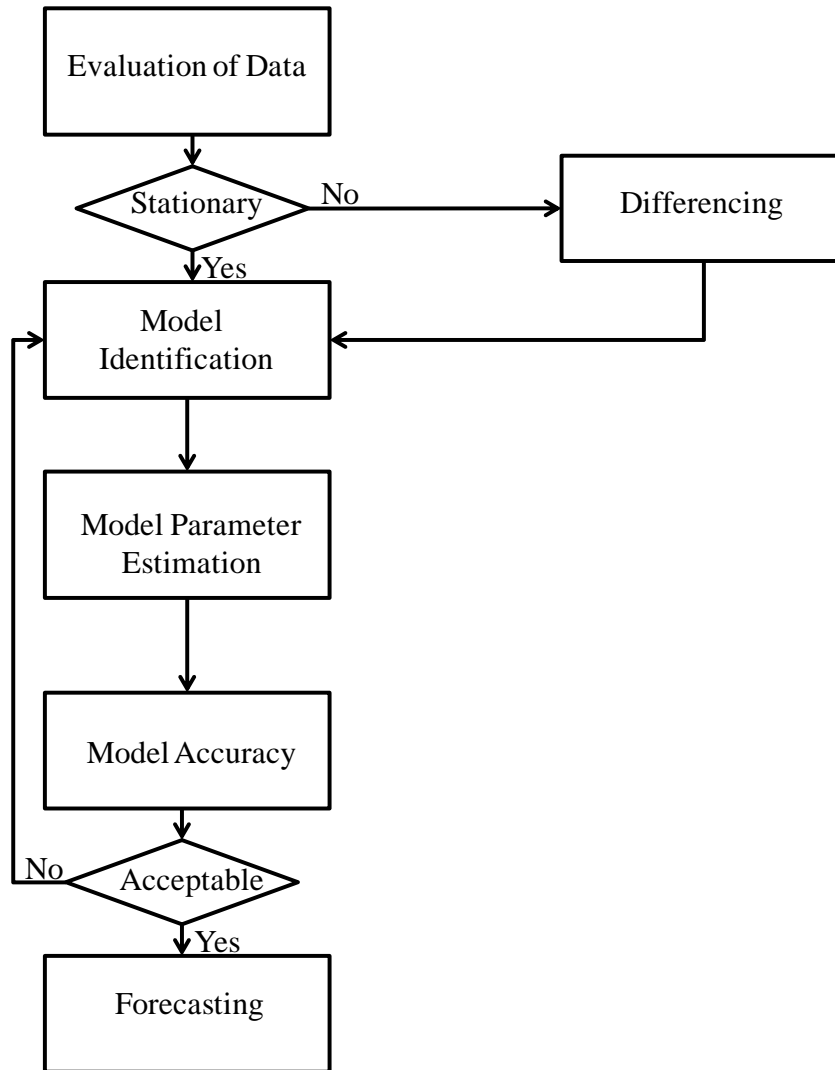


Figure XII-1: Flowchart of establishing the ARIMA model

A. Evaluation of Data

In the evaluation of data step, the data is tested for trends and nonstationary behavior. Typically, this test is performed by plotting the data as a scatter plot to observe its patterns or observing the variance and the mean. In practice, graphical representation of the data is a powerful method of evaluating and observing the nature of the data. The scatter plot uses pairs of the response variable series separated by a constant interval or an equivalent lag value [57]. If the data presents no evidence of a trend, the data is stationary in nature. If the data presents evidence of a trend, the

differencing technique is initiated to remove the trend. The process of differencing is iterative; that is, it is a process of trial and error. Commonly, the difference order is 0, 1, or at most 2, which represents the number of times it takes to differentiate the nonstationary series to achieve stationary behavior [62]. Once the data is stationary, the model identification process can start.

B. Model Identification

There are three functions, introduced by the Box and Jenkins approach, which are used to determine the form of the ARIMA Transfer Function ETHER models. These functions include: the autocorrelation function, the partial autocorrelation function and the cross-correlation function. The model identification step can be an iterative process that is conducted through trial and error methods. Based on the results obtained from the functions, a number of different models can be formed and tested for validation.

The autocorrelation function (ACF) and partial autocorrelation function (PACF) are both used to determine the form as well as determine the values for the ARIMA (p,d,q) model. The differencing order, d, has already been found in the previous step. The ACF can be represented as a bar chart plot of the coefficients of correlation between a time series and lags of itself. The ACF provides information on the moving average order (q) in the ARIMA model. The PACF, on the other hand, provides information to determine the p value in ARIMA model. If the ACF and PACF decay rapidly to zero with an increased lag, the data series is said to be stationary. The autoregressive (AR) and moving average (MA) terms are identified using the ACF and PACF functions to determine whether the AR and MA terms are needed to correct any autocorrelation that remains in the differenced series.

The cross-correlation function (CCF) between the input and output is used to identify the transfer function models. The cross-correlation provides information on the denominator and numerator operators in the transfer function model. The denominator operator represents the times series of

the dependent series while the numerator represents the time series of the input or independent variables.

C. Model Parameter Estimation

In the previous step, the orders of the autoregressive (AR) and the moving average (MA) terms are investigated to obviate repeating patterns amongst the data series, the orders of the transfer function and the order of the differencing (d) term were investigated in the evaluation of data step. Now that the general form of the ARIMA model has been estimated, the coefficients for the autoregressive, moving average, and transfer function parameters have to be estimated.

The coefficients of the autoregressive, moving average, and transfer function terms can be estimated using different methods: maximum likelihood method or least squares method. Simulations have been performed by many statisticians to empirically investigate and compare the performance between the two different methods. It has been found that the maximum likelihood method produces satisfactory parameter value approximations for small to moderate datasets and the least squares method is more commonly used for larger datasets [57]. Once the ARIMA model has been formed with the order of the autoregressive, differencing, moving average and transfer function terms and associated coefficients for the parameters, the model can be tested for accuracy.

D. Model Accuracy

The purpose of the model accuracy step is to assess the desired level of model adequacy. This is done by observing the difference between the model output and the actual data. Goodness of fit measures, which are discussed in the next chapter of this dissertation, are suggested for use to obtain a reasonable fit to the actual data. If the model accuracy is acceptable, the estimated parameters found in the model parameter estimation step are applied to the forecasting equation

which is modified by factoring in the lead time (i.e., $t+h$). If the model accuracy is not acceptable, the model identification step is re-examined.

The mathematical processes described in this section provide an introductory look at building ARIMA models outside of statistical application tools. For future work, each step can be expanded and elaborated in more detail. Understanding the core fundamentals of each step will help to foster implementation into an embedded environment.

XIII. Appendix B – Data Import Syntax Code

```
GET DATA
/TYPE=TXT
/FILE='C:\Documents and Settings\username\File Directory\CSV File'
/DELCASE=LINE
/DELIMITERS=","
/ARRANGEMENT=DELIMITED
/FIRSTCASE=2
/IMPORTCASE=ALL
/VARIABLES=
DateT DATE9
T A4
F1 F1.0
F2 F3.1
F3 F3.1
F4 F3.1
CT1 F5.1
CT2 F4.1
CT3 F5.1
CT4 F6.1
CT5 F3.1
CT6 F6.1
CT7 F4.1
CT8 F6.1
CT9 F4.1
CT10 F3.1
CT11 F5.1
CT12 F5.1
CT13 F6.1
CT14 F5.1
IntTemp F4.1
T15 F4.1
T16 F1.0
T17 F4.1
T18 F4.1
T19 F4.1
T20 F4.1
T21 F4.1
T22 F5.1
T23 F5.1
T24 F4.1
T25 F4.1
T26 F4.1
T27 F4.1.
CACHE.
VARIABLE LEVEL F1(SCALE).
VARIABLE LEVEL F2(SCALE).
VARIABLE LEVEL F3(SCALE).
VARIABLE LEVEL F4(SCALE).
```

```
VARIABLE LEVEL CT1(SCALE).  
VARIABLE LEVEL CT2(SCALE).  
VARIABLE LEVEL CT3(SCALE).  
VARIABLE LEVEL CT4(SCALE).  
VARIABLE LEVEL CT5(SCALE).  
VARIABLE LEVEL CT6(SCALE).  
VARIABLE LEVEL CT7(SCALE).  
VARIABLE LEVEL CT8(SCALE).  
VARIABLE LEVEL CT9(SCALE).  
VARIABLE LEVEL CT10(SCALE).  
VARIABLE LEVEL CT11(SCALE).  
VARIABLE LEVEL CT12(SCALE).  
VARIABLE LEVEL CT13(SCALE).  
VARIABLE LEVEL CT14(SCALE).  
VARIABLE LEVEL T15(SCALE).  
VARIABLE LEVEL T16(SCALE).  
VARIABLE LEVEL T17(SCALE).  
VARIABLE LEVEL T18(SCALE).  
VARIABLE LEVEL T19(SCALE).  
VARIABLE LEVEL T20(SCALE).  
VARIABLE LEVEL T21(SCALE).  
VARIABLE LEVEL T22(SCALE).  
VARIABLE LEVEL T23(SCALE).  
VARIABLE LEVEL T24(SCALE).  
VARIABLE LEVEL T25(SCALE).  
VARIABLE LEVEL T26(SCALE).  
VARIABLE LEVEL T27(SCALE).  
EXECUTE.  
DATASET NAME DataSet1 WINDOW=FRONT.  
  
SAVE OUTFILE= 'C:\Documents and Settings\username\File Directory\filename.sav'  
/COMPRESSED.
```

XIV. Appendix C – Model Development Syntax Code

PREDICT THRU END.

* Time Series Modeler.

TSMODEL

```

/MODELSUMMARY PRINT=[MODELFIT]
/MODELSTATISTICS DISPLAY=YES MODELFIT=[ SRSQUARE]
/MODELDETAILS PRINT=[ PARAMETERS]
/SERIESPLOT OBSERVED FORECAST
/OUTPUTFILTER DISPLAY=ALLMODELS
/AUXILIARY CILEVEL=95 MAXACFLAGS=24
/MISSING USERMISSING=EXCLUDE
/MODEL DEPENDENT=IntTemp INDEPENDENT=Desired Variables
  OUTFILE=' C:\Documents and Settings\username\File Directory\filename.xml'
  PREFIX='Model'
/EXPERTMODELER TYPE=[ARIMA EXSMOOTH] TRYSEASONAL=YES
/AUTOOUTLIER DETECT=OFF.

```

XV. Appendix D – Forecasting Syntax Code

GET DATA

```

/TYPE=TXT
/FILE='C:\Documents and Settings\username\File Directory\CSV File'
/DELCASE=LINE
/DELIMITERS=","
/ARRANGEMENT=DELIMITED
/FIRSTCASE=2
/IMPORTCASE=ALL
/VARIABLES=
DateT DATE9
T A4
F1 F1.0
F2 F3.1
F3 F3.1
F4 F3.1
CT1 F5.1
CT2 F4.1
CT3 F5.1
CT4 F6.1
CT5 F3.1
CT6 F6.1
CT7 F4.1
CT8 F6.1
CT9 F4.1
CT10 F3.1
CT11 F5.1
CT12 F5.1
CT13 F6.1
CT14 F5.1

```



```
IntTemp F4.1
T15 F4.1
T16 F1.0
T17 F4.1
T18 F4.1
T19 F4.1
T20 F4.1
T21 F4.1
T22 F5.1
T23 F5.1
T24 F4.1
T25 F4.1
T26 F4.1
T27 F4.1.
CACHE.
VARIABLE LEVEL F1(SCALE).
VARIABLE LEVEL F2(SCALE).
VARIABLE LEVEL F3(SCALE).
VARIABLE LEVEL F4(SCALE).
VARIABLE LEVEL CT1(SCALE).
VARIABLE LEVEL CT2(SCALE).
VARIABLE LEVEL CT3(SCALE).
VARIABLE LEVEL CT4(SCALE).
VARIABLE LEVEL CT5(SCALE).
VARIABLE LEVEL CT6(SCALE).
VARIABLE LEVEL CT7(SCALE).
VARIABLE LEVEL CT8(SCALE).
VARIABLE LEVEL CT9(SCALE).
VARIABLE LEVEL CT10(SCALE).
VARIABLE LEVEL CT11(SCALE).
VARIABLE LEVEL CT12(SCALE).
VARIABLE LEVEL CT13(SCALE).
VARIABLE LEVEL CT14(SCALE).
VARIABLE LEVEL T15(SCALE).
VARIABLE LEVEL T16(SCALE).
VARIABLE LEVEL T17(SCALE).
VARIABLE LEVEL T18(SCALE).
VARIABLE LEVEL T19(SCALE).
VARIABLE LEVEL T20(SCALE).
VARIABLE LEVEL T21(SCALE).
VARIABLE LEVEL T22(SCALE).
VARIABLE LEVEL T23(SCALE).
VARIABLE LEVEL T24(SCALE).
VARIABLE LEVEL T25(SCALE).
VARIABLE LEVEL T26(SCALE).
VARIABLE LEVEL T27(SCALE).
EXECUTE.
DATASET NAME DataSet1 WINDOW=FRONT.

SAVE OUTFILE='C:\Documents and Settings\username\File Directory\filename.sav'
/COMPRESSED.
```

GET

FILE='C:\Documents and Settings\username\File Directory\filename.sav '.

DATASET NAME DataSet0 WINDOW=FRONT.

PREDICT THRU END.

* Apply Time Series Models.

TSAPPLY

/MODELSUMMARY PRINT=[MODELFIT]

/MODELSTATISTICS DISPLAY=YES MODELFIT=[SRSQUARE]

/SERIESPLOT FORECAST

/OUTPUTFILTER DISPLAY=ALLMODELS

/AUXILIARY CILEVEL=95 REESTIMATE=YES

/MISSING USERMISSING=EXCLUDE

/MODEL FILE=' C:\Documents and Settings\username\File Directory\filename.xml'.

GET DATA

/TYPE=TXT

/FILE='C:\Documents and Settings\username\File Directory\CSV File'

/DELCASE=LINE

/DELIMITERS=","

/ARRANGEMENT=DELIMITED

/FIRSTCASE=2

/IMPORTCASE=FIRST 360

/VARIABLES=

DateT DATE9

T A4

F1 F1.0

F2 F3.1

F3 F3.1

F4 F3.1

CT1 F5.1

CT2 F4.1

CT3 F5.1

CT4 F6.1

CT5 F3.1

CT6 F6.1

CT7 F4.1

CT8 F6.1

CT9 F4.1

CT10 F3.1

CT11 F5.1

CT12 F5.1

CT13 F6.1

CT14 F5.1

IntTemp F4.1

T15 F4.1

T16 F1.0

T17 F4.1

T18 F4.1

T19 F4.1

T20 F4.1

T21 F4.1

```
T22 F5.1
T23 F5.1
T24 F4.1
T25 F4.1
T26 F4.1
T27 F4.1.
CACHE.
VARIABLE LEVEL F1(SCALE).
VARIABLE LEVEL F2(SCALE).
VARIABLE LEVEL F3(SCALE).
VARIABLE LEVEL F4(SCALE).
VARIABLE LEVEL CT1(SCALE).
VARIABLE LEVEL CT2(SCALE).
VARIABLE LEVEL CT3(SCALE).
VARIABLE LEVEL CT4(SCALE).
VARIABLE LEVEL CT5(SCALE).
VARIABLE LEVEL CT6(SCALE).
VARIABLE LEVEL CT7(SCALE).
VARIABLE LEVEL CT8(SCALE).
VARIABLE LEVEL CT9(SCALE).
VARIABLE LEVEL CT10(SCALE).
VARIABLE LEVEL CT11(SCALE).
VARIABLE LEVEL CT12(SCALE).
VARIABLE LEVEL CT13(SCALE).
VARIABLE LEVEL CT14(SCALE).
VARIABLE LEVEL T15(SCALE).
VARIABLE LEVEL T16(SCALE).
VARIABLE LEVEL T17(SCALE).
VARIABLE LEVEL T18(SCALE).
VARIABLE LEVEL T19(SCALE).
VARIABLE LEVEL T20(SCALE).
VARIABLE LEVEL T21(SCALE).
VARIABLE LEVEL T22(SCALE).
VARIABLE LEVEL T23(SCALE).
VARIABLE LEVEL T24(SCALE).
VARIABLE LEVEL T25(SCALE).
VARIABLE LEVEL T26(SCALE).
VARIABLE LEVEL T27(SCALE).
EXECUTE.
DATASET NAME DataSet2 WINDOW=FRONT.

SAVE OUTFILE='C:\Documents and Settings\username\File Directory\filename1.sav'
/COMPRESSED.
```

XVI. Appendix E – Re-estimation Syntax Code

PREDICT THRU END.

* Apply Time Series Models.

TSAPPLY

/MODELSUMMARY PRINT=[MODELFIT]

/MODELSTATISTICS DISPLAY=YES MODELFIT=[SRSQUARE]

/SERIESPLOT FORECAST

/OUTPUTFILTER DISPLAY=ALLMODELS

/AUXILIARY CILEVEL=95 REESTIMATE=YES

/MISSING USERMISSING=EXCLUDE

/MODEL FILE=' C:\Documents and Settings\username\File Directory\filename.xml'.

XVII. Appendix F - Variable Reduction Syntax Code

COMMENT - Correlation code for the data sources of the same construct

CORRELATIONS

/VARIABLES=F1 F2 F3 F4

/PRINT=TWOTAIL NOSIG

/MISSING=PAIRWISE.

CORRELATIONS

/VARIABLES=T15 T16 T17

/PRINT=TWOTAIL NOSIG

/MISSING=PAIRWISE.

CORRELATIONS

/VARIABLES=T18 T19 T20 T21

/PRINT=TWOTAIL NOSIG

/MISSING=PAIRWISE.

CORRELATIONS

/VARIABLES=T22 T23

/PRINT=TWOTAIL NOSIG

/MISSING=PAIRWISE.

CORRELATIONS

/VARIABLES=T24 T25 T26 T27

/PRINT=TWOTAIL NOSIG

/MISSING=PAIRWISE.

COMMENT – Code to the save output of the correlation matrix

SAVE OUTFILE='C:\Documents and Settings\username\MyDocuments\SPSS\filename.sav'.

SELECT IF ROWTYPE_="CORR".

```

SAVE OUTFILE='C:\Documents and Settings\username\MyDocuments\SPSS\filename.sav'.
SELECT IF ROWTYPE_="CORR".
SAVE OUTFILE='C:\Documents and Settings\username\MyDocuments\SPSS\filename.sav'.
SELECT IF ROWTYPE_="CORR".
SAVE OUTFILE='C:\Documents and Settings\username\MyDocuments\SPSS\filename.sav'.
SELECT IF ROWTYPE_="CORR".
SAVE OUTFILE='C:\Documents and Settings\username\MyDocuments\SPSS\filename.sav'.
SELECT IF ROWTYPE_="CORR".
/COMPRESSED.

```

COMMENT - Code to extract and save values from the pair-wise correlation

```

USE ALL.
VARSTOCASES
/ID=T
/MAKE CVAL FROM T18 T19 T20 T21
/INDEX=Index1(CVAL)
/KEEP=ROWTYPE_ VARNAME_
/NULL=KEEP.
SAVE OUTFILE='C:\Documents and Settings\username\My Documents\SPSS\filename.sav'.

```

COMMENT – Code to sort and save the values of the pair-wise correlation of the variables

```

SORT CASES BY CVAL(A).
SAVE OUTFILE='C:\Documents and Settings\username\My Documents\SPSS\filename.sav'.
/COMPRESSED.

```

COMMENT - Code to select and save the desired values within the correlation matrix.

```

USE ALL.
SELECT IF (CVAL<1).
DATASET COPY Filename.
DATASET ACTIVATE Filename.
FILTER OFF.
SAVE OUTFILE='C:\Documents and Settings\username\MyDocuments\SPSS\filename.sav'.

```

COMMENT – Code to identify duplicate cases and sort matching groups in the correlation matrix

* Identify Duplicate Cases.

```

USE ALL.
SORT CASES BY CVAL(A) VARNAME_(A) Index1(A).
MATCH FILES
/FILE=*
/BY CVAL
/FIRST=PrimaryFirst
/LAST=PrimaryLast.
DO IF (PrimaryFirst).
COMPUTE MatchSequence=1-PrimaryLast.
ELSE.
COMPUTE MatchSequence=MatchSequence+1.
END IF.

```

```

LEAVE MatchSequence.
FORMAT MatchSequence (f7).
COMPUTE InDupGrp=MatchSequence>0.
SORT CASES InDupGrp(D).
MATCH FILES
  /FILE=*
  /DROP=PrimaryLast InDupGrp MatchSequence.
VARIABLE LABELS PrimaryFirst 'Indicator of each first matching case as Primary'.
VALUE LABELS PrimaryFirst 0 'Duplicate Case' 1 'Primary Case'.
VARIABLE LEVEL PrimaryFirst (ORDINAL).
FREQUENCIES VARIABLES=PrimaryFirst.
SAVE OUTFILE='C:\Documents and Settings\username\My Documents\SPSS\filename.sav'.

```

COMMENT – Code to identify primary cases within the correlation matrix

```

USE ALL.
SELECT IF (PrimaryFirst=1).
DATASET COPY Primary_Cases.
DATASET ACTIVATE Primary_Cases.
FILTER OFF.
SAVE OUTFILE='C:\Documents and Settings\username\My Documents\SPSS\filename.sav'.
DATASET ACTIVATE DataSet1.
DATASET CLOSE Filename.
DATASET ACTIVATE Primary_Cases.
USE ALL.
DO IF ((VARNAME_="T18") & (Index1="T20") & (CVAL >= 0.85))
  DATASET ACTIVATE DataSet1
  COMPUTE Average = MEAN(T18,T20)
END IF.
UPDATE FILE= 'C:\Documents and Settings\username\My Documents\SPSS\filename.sav'.
SAVE OUTFILE='C:\Documents and Settings\username\My Documents\SPSS\filename.sav'.
/COMPRESSED.
/DROP = T18 T20.
& (Index1="T19")
ELSE IF (VARNAME_="T18") & (Index1="T20") & (CVAL >= 0.85)
  DATASET ACTIVATE DataSet1
  COMPUTE Average1 = MEAN(T18,T20)
USE ALL.
IF ( VARNAME_="T18" & Index1="T19" & CVAL >= .85).
DATASET ACTIVATE DataSet1.
COMPUTE Average = MEAN(T18,T19).
DELETE T18.
DELETE T19.
FILTER OFF.
DATASET ACTIVATE Primary_Cases.
USE ALL.
IF ( VARNAME_="T18" & Index1="T20" & CVAL >= .85).
DATASET ACTIVATE DataSet1
COMPUTE Average=MEAN(T18,T20)
DELETE T18
DELETE T20.

```

FILTER OFF.

COMMENT – Code to perform a regression test on the selected variables

REGRESSION

/MISSING LISTWISE

/STATISTICS TOL

/CRITERIA=PIN(.05) POUT(.10)

/NOORIGIN

/DEPENDENT IntTemp

/METHOD=ENTER F2 F3 F4 CT1 CT2 CT3 CT4 CT5 CT6 CT7 CT8 CT9 CT10 CT1
1 CT12 CT13 CT14 T15 T16 T17 T18 T19 T20 T21 T22 T23 T24 T25 T26 T27.

XVIII. Appendix G – Evaluation Technique Prediction Results

