A METHODOLOGY FOR FORECASTING IMPACT OF DEMAND RESPONSE ON CAPACITY EXPANSION PLANNING

A Dissertation

Presented to

The Academic Faculty

by

Ju Hyun Kim

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy in the

School of Aerospace Engineering

Georgia Institute of Technology

August 2018

A METHODOLOGY FOR FORECASTING IMPACT OF DEMAND RESPONSE ON CAPACITY EXPANSION PLANNING

Approved by:

Prof. Dimitri N. Mavris, Advisor
School of Aerospace Engineering
Georgia Institute of Technology

Dr. Scott J. Duncan
School of Aerospace Engineering
Georgia Institute of Technology

Prof. Daniel P. Schrage
School of Aerospace Engineering
Georgia Institute of Technology

Mr. Ken Caird
GE Energy

Prof. Jachiel Jogoda
School of Aerospace Engineering
Georgia Institute of Technology

Date Approved: May 11, 2018

To my family

ACKNOWLEDGEMENTS

It’s been a long journey, but I have reached the finish line and can move on to another chapter of my life with the help of others. I would like to thank to everyone who supported me throughout the years. First, I would like to thank my advisor, Dr. Dimitri Mavris, for giving me the opportunity to join ASDL and supporting many years of my study. Without his guidance and inspiration, this work would not have existed. I would like to thank my mentor, Dr. Scott Duncan, for his advice and support. Those regular or irregular meetings with him kept me motivated. I also would like to thank the other members of my thesis committee, Dr. Jeff Jagoda, Dr. Daniel Schrage, and Mr. Ken Caird, for reading my thesis and giving valuable feedback. Thanks to Smart Campus team, especially Dr. Woongje Sung and Dr. Michael Balchanos, who helped me frame research in modeling and simulations. Finally, I would like to thank my family in Korea for their love and endurance. I cannot thank my parents enough for that they have never doubted and always cheered me on in whatever I did.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ASHRAE	American Society of Heating, Refrigerating and Air-Conditioning Engineers
AHP	Analytical Hierarchy Process
ANN	Artificial Neural Network
BTO	Building Technologies Office
COP	Coefficient of Performance
CVRMSE	Cumulative Variation of Root Mean Squared Error
DOE	Department of Energy
DR	Demand Response
DSM	Demand Side Management
ELCC	Effective Load Carrying Capability Demand Side Management
ELECTRE	Elimination and Choice Translating Reality
EIA	Energy Information Administration
ENS	Energy Not Served
GPM	Gallon Per Minute
HVAC	Heating, Ventilation, and Air Conditioning
LBL	Lawrence Berkeley National Laboratory

LP	Linear Programming
LDC	Load Duration Curve
LLD	Loss of Load Duration
LOLE	Loss of Load Expectation
LOLP	Loss of Load Probability
MBE	Mean Bias Error
MTBF	Mean Time between Failure
MTTR	Mean Time to Repair
MILP	Mixed Integer Linear Programming
MINLP	Mixed Integer Non Linear Programming
MIP	Mixed Integer Programming
MCS	Monte Carlo Simulation
MADM	Multi Attribute Decision Making
MCDM	Multi Criteria Decision Making
MODM	Multi Objective Decision Making
NPV	Net Present Value
PROMETHEE	Preference Ranking Organization Method for Enrichment Evaluation

QOI	Quantity of Interest
TOPSIS	Technique for Order Preference by Similarity to Ideal Solution

SUMMARY

In parallel to tighter energy regulations and increasing demand for emissions reduction, the Department of Energy (DOE) has set a goal to reduce energy consumption in the building sector to 50% of 2010 levels by 2030. This encourages the use of advanced operational strategies and demand side management concepts to improve energy efficiency and reduce peak energy loads. In response to this societal need, many communities such as cities and university campuses are trying to transform their energy systems into smart and sustainable ones. At the same time, capital planners in these organizations are interested in deferring the need to expand energy supply capacity, e.g., new chillers for district cooling systems, to avoid incurring those costs until farther in the future.

A campus level district energy system is in the mixed position of an energy producer and consumer. It consumes the electrical energy to produce the thermal energy for multiple buildings within the community. A district energy plant tends to have the excessive capacity with a redundancy to ensure the system reliability. Advancement in building energy technologies and communication systems enables the concept of Demand Side Management (DSM) to become a reality. Among the DSM concepts, Demand Response (DR) is now the most established practice for the power system management. The established DR can be utilized by the owner of the district energy system to alleviate the oversizing problem and defer the capital investment on new chillers if it can replace a redundant chiller.

This thesis proposes a methodology to quantitatively evaluate the impact of DR on a district energy system's operation and planning. The proposed methodology utilizes measured data, converts them into actionable information by developing models to capture

the interaction between demand and supply sides, and provides an insight into the planning design space by connecting the planning optimization and the reliability analysis modules. The methodology consists of seven steps including preparatory steps. Each of the steps is developed from research questions seeking efficient modeling and analysis methods. A data driven cooling demand model discerns the amount of load shed by a specific DR method, and the physics based chiller plant model is approximated as a performance curve for fast assessment of DR. Capacity expansion planning is formulated as a Mixed Integer Linear Programming problem and follows a simulation-based reliability analysis.

The proposed methodology is applied to an example system, which is based on the real system data, to demonstrate capability. Results show that the proposed method can quantify and present the trade-off space between the capital cost and the reliability of the optimal system expansion plans and can reveal hidden trends in the capacity planning design space. Compared to the baseline N+1 system, it shows how long the plant expansion can be deferred and how much cost can be saved from the preferred plan.

CHAPTER 1 INTRODUCTION

1.1 Motivation

1.1.1 Building Energy

Buildings are significant energy consumers. According to the US Energy Information Administration (EIA), 41% of the U.S. energy consumption is used by residential and commercial buildings [1]. Climate change creates two conflicting trends in the building energy sector. Building occupants demand more energy for cooling throughout a year due to the higher temperatures and the longer summer than before. On the other hand, many governments and international organizations set goals to reduce energy consumptions and greenhouse gas (GHG) emission. The Energy Information Administration of the United States (EIA) projects that world energy consumption will increase 56% within 30 years since 2010 and the total consumption will be up to 820 quadrillions British Thermal Unit (Btu) by 2040 [2].

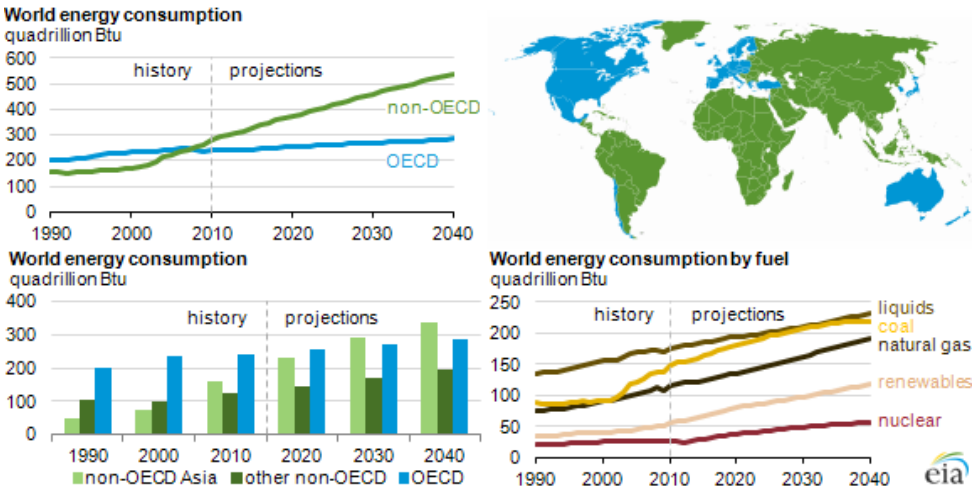


Figure 1 World Energy Consumption Trends [2]

Fossil fuels such as coal and oil are continuously large contributors to energy sources, even though renewable energy is growing as presented in Figure 1. This concern raises interest in innovating and deploying emerging technologies for buildings to improve efficiency and maintain a sustainable energy system.

The Building Technologies Office (BTO) of the Department of Energy (DOE) sets a goal to reduce building energy consumption by 50% by 2030 as displayed in Figure 2. This is a very aggressive goal because the energy demand increases naturally, especially for the cooling. To overcome this challenge, they identified technologies available in the near future and the next generation technologies for the building HVAC (Heating, Ventilation, and Air Conditioning) systems, which will be able to "leapfrog" the current technological barrier as described in Figure 3 [3].

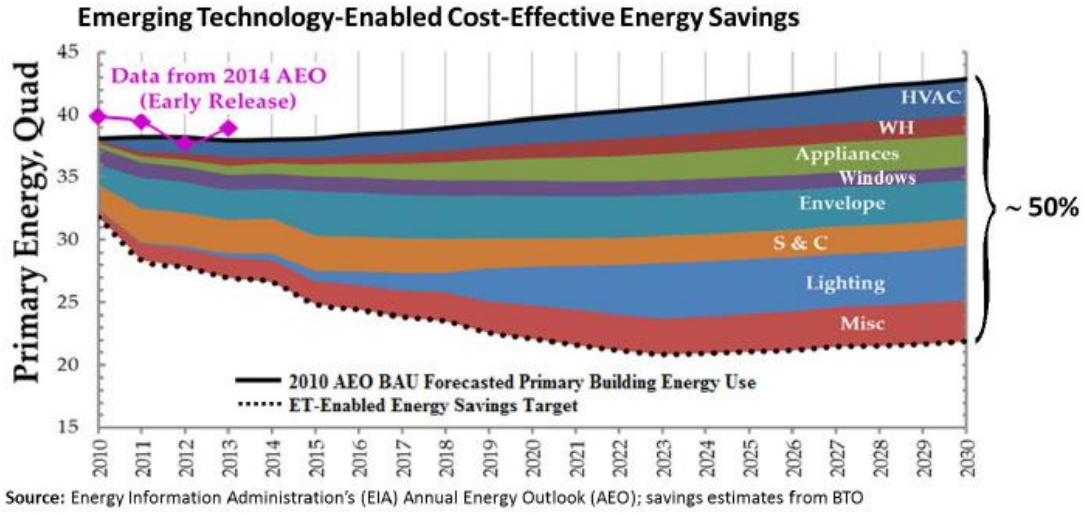


Figure 2 BTO Emerging-Technology-Enabled Energy Savings Targets [3]

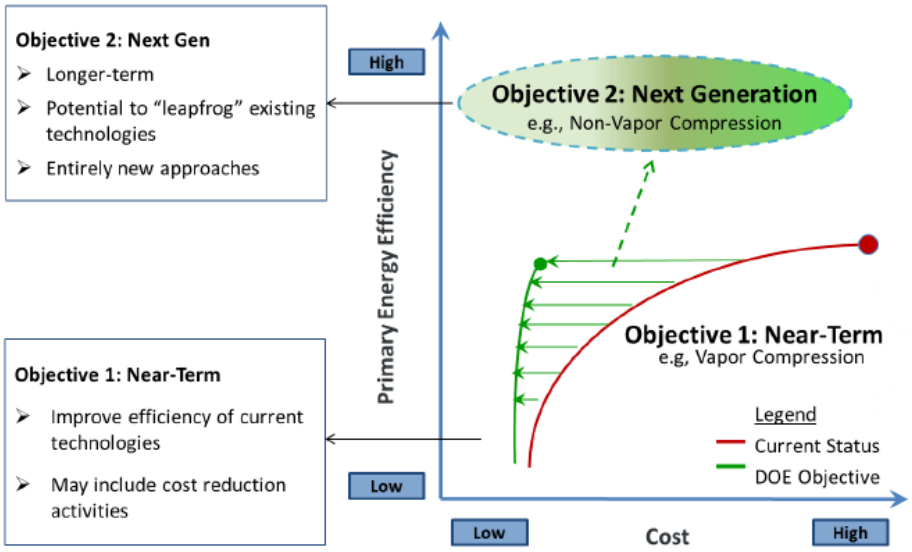


Figure 3 BTO Approach to HVAC R&D [3]

1.1.2 Demand Side Management and District Energy Systems

Advancement of building HVAC system technologies, especially the building automation and communication technologies, enables and expedites demand side management. Demand side management (DSM) was developed initially to overcome the energy crisis in the 1970s. It is an operational concept of the electrical grid to alter load shape by reducing peak load to mitigate the risk and increase the operating flexibility [4]. DSM concepts are categorized into six groups based on how they change the shape of the original load curve as shown in Figure 4 [5].

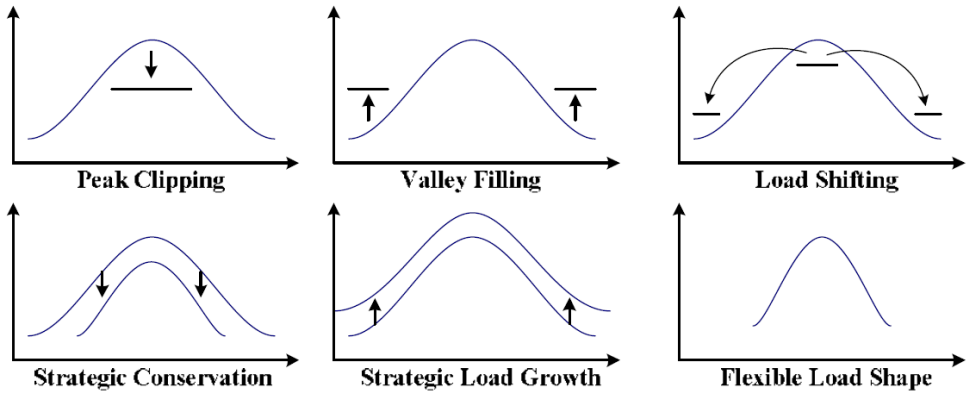


Figure 4 Load Shaping Concepts by DSM [5]

Palensky and Dietrich categorize DSM into four groups based on the timing and the impact on the process [6] as illustrated in Figure 5: Energy Efficiency, Time of Use, Demand Response, and Spinning Reserve. Demand Response (DR) is the end users' voluntary changes of energy usage in response to the utility company's request. Energy Efficiency programs change the load shape and energy use permanently by using building energy technologies such as the ones introduced earlier while others have the immediate impact on the load shape and demand pattern will return to the original shape without controls.

They distinguish market DR and physical DR based on the motivation of DR. The market DR is to use the real-time pricing or incentives to encourage customers' participation. On the other hand, the physical DR is considered as an emergency operation method when the load shedding is required due to system failures. Even though DR does not change the load shape permanently, it can affect the system capacity plan by reducing peak load, which happens only a few hours but the system should be built to provide that amount of power [7].

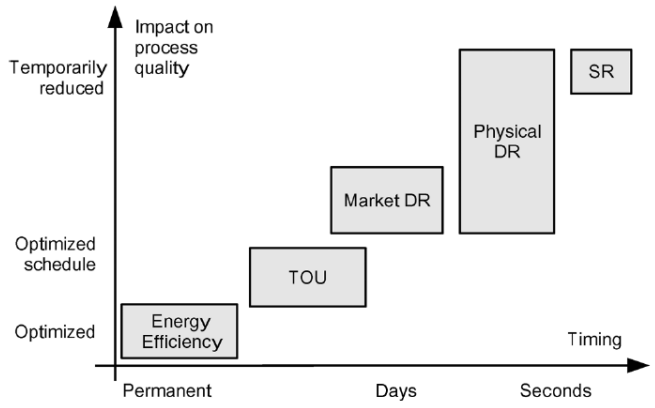


Figure 5 Categories of DSM [6]

An introductory report about DR strategies and techniques listed three reasons why HVAC systems are very attractive targets [8]. First, building HVAC systems consume a significant amount of electrical energy. As of 2010, energy consumption by HVAC systems is more than 40% of the entire U.S. buildings [3]. Second, because of the thermal mass of buildings, which causes the '*thermal flywheel'* effect, the HVAC load can be reduced temporarily without noticing by the building occupants. Lastly, the building automation systems allow implementation of pre-planned DR programs.

For heating and cooling for buildings, many cities or campuses adopt a district energy system. A district energy system consists of one or more central plants that produce cooling or heating energy, the distribution piping systems, and multiple buildings connected to the network. The produced thermal energy is distributed as the form of water or air through a dedicated underground piping network to multiple end users as illustrated in Figure 6. End users can be any types of buildings: residential buildings, offices, industrial buildings, or research labs. Each building is equipped with a heat exchanger to obtain heat from the network instead of individual buildings' own chillers or furnace. This improves overall efficiency and reliability of the entire system. For that reason, district

cooling systems are widely used in high density areas where cooling and dehumidification are required [9].

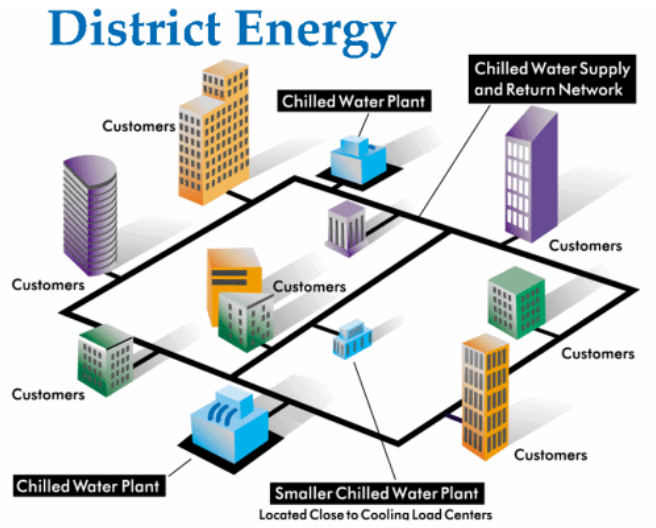


Figure 6 District Energy Chilled Water System [10]

In the current market DR practice managed by electrical utility companies, the buildings respond to the price signal from the grid side and change their set-point temperature of HVAC systems to reduce energy consumption instantly. While a building can get slight energy savings during DR, the chilled water plant can have much more savings and operational flexibility. However, the benefit of DR for a community with district energy plant is not well appreciated because the plant side benefit is overlooked.

The system operator of a district energy system can accomplish the same practice by themselves. There are few attempts to use the DR concept for the district energy systems. For example, Wernstedt et al. developed a multi-agent based control method for district heating system and performed field test [11]. They demonstrated that the total energy consumption is reduced by 4% by shifting load to the off-peak period when the production cost is cheaper.

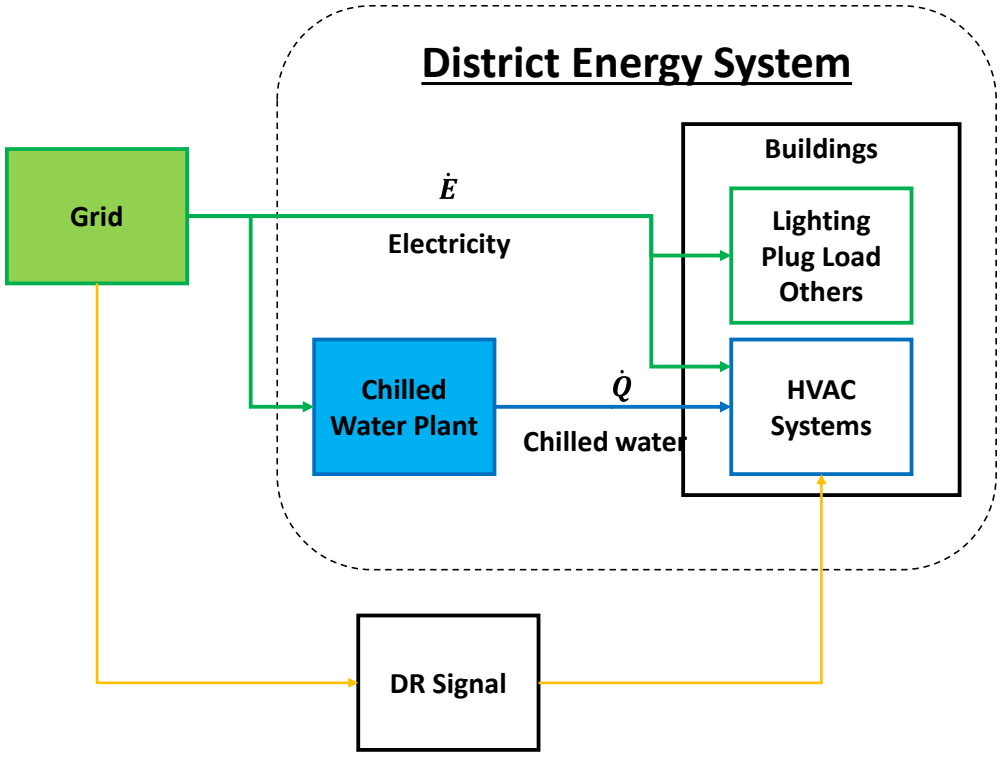


Figure 7 Energy Flow with Demand Response in District Energy System

If the DR program is implemented for wide area, the impact on the system will be significant. A study using smart meter data shows that the residential A/C in California has the potential peak demand reduction of 310MW to 3.5GW depending on the duration of load shed [12]. Even though the DR concept is devised to help the operation of the electrical grid, the same concept can be applied to district energy systems because both systems have the similar structure and functionality to balance energy supply and demand. Since the building HVAC system uses both types of energy – electricity and thermal energy, it is obvious that the chiller plant can see the reduced load during the DR events. The chiller is an energy conversion machine that transforms the electricity into the thermal energy. Therefore, if the required cooling load to be produced at the chiller plant is reduced, then the electricity consumption by the plant can be saved. The energy flow of the two types of energy from the sources to the end users in a district energy system is illustrated in Figure

7. Only a cooling system is represented here, but it can be extended with the heating plants with additional primary energy resources such as natural gas.

1.1.3 Capacity Expansion Planning

A school campus is not a static system. The landscape is changing with time. Buildings are continuously built and demolished as needed. The population within the community also keeps changing. As the demand is changing over time, the energy planner of the community should consider expanding or upgrading the energy plant at the right time to meet the increasing demand. Purchasing and installing major equipment usually takes more than a year [13]. Therefore, the decision making for capacity expansion should be made with consideration of the lead time.

A district cooling system is typically sized to have N+1 redundancy for reliability [14]. This requirement should be kept while expanding the plant. In industrial practice, the need for capacity expansion planning is established by a future demand projection based on the campus master plan. The future peak load projection is estimated by the building area (gross square feet) of the new construction projects specified by the master plan. Once the existing system's capacity is evaluated, it is determined when a new chiller or a plant is needed by the demand projection reaching the capacity limit. A notional capacity expansion plan is illustrated in Figure 8. For this notional system, three chillers currently provide cooling energy. Additional chillers will be needed in years of 3, 8, and 11. The firm capacity is the usable capacity when the largest chiller is unavailable. The plant should meet the peak demand projection with the firm capacity to maintain the N+1 redundancy requirement over the planning horizon. A few available options selected by subject matter experts are evaluated for their economic viability. Once specific chiller models for the

additional capacity are selected, the capital phasing plan is created with the preferred options. This process can be summarized as in Figure 9.

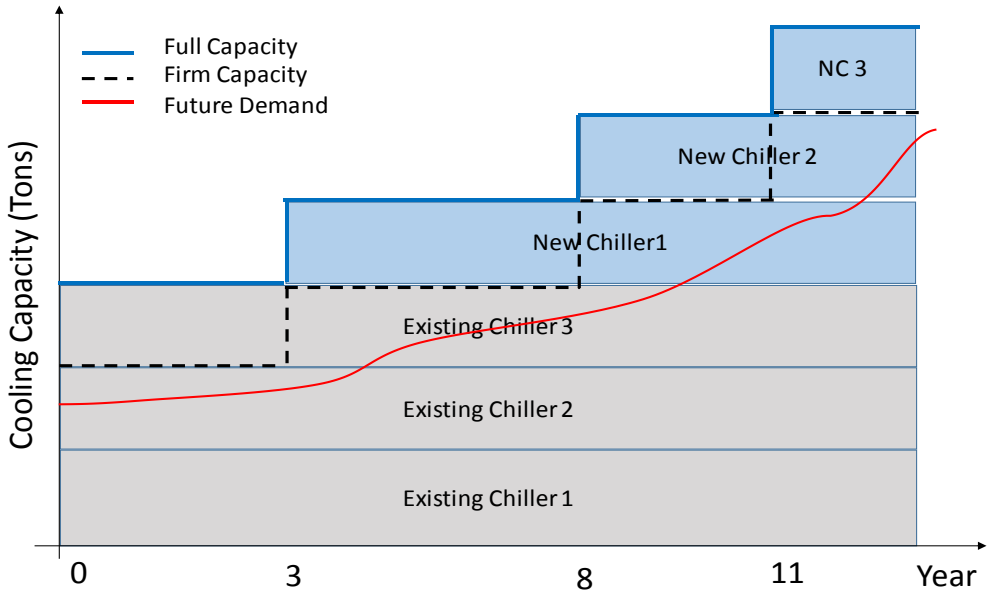


Figure 8 A Notional Chilled Water System Capacity Expansion Plan

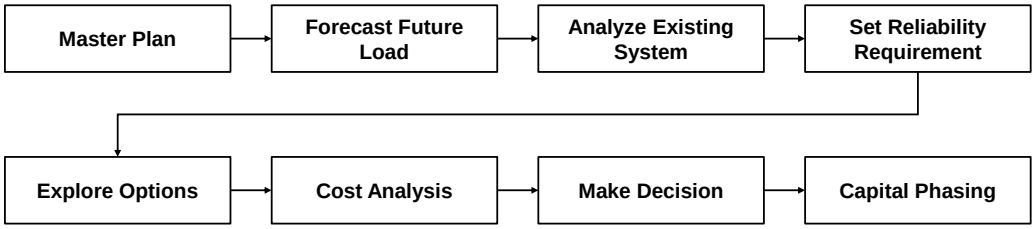


Figure 9 Capacity Expansion Planning Decision Process

Even though this process has been used as the industry standard, there are some shortcomings. First, the current peak load estimation method can lead to the oversized system in many ways. Typically, a safety factor is added to the deterministic estimation of the peak load as Huang et al. pointed out [15] because of lack of uncertainty in demand

estimation. In addition to that, there are many different types of buildings having different load shapes in the system, but it is not considered. The simple summation of their peak loads is not equal to the peak load seen at the plant. Second, the N+1 redundancy requirement is applied to ensure that the system serves the peak load even in a contingency condition when one unit is unavailable due to failure. Despite the fact that this is a simple and effective way to ensure the system reliability, it does not consider the dynamic characteristics of the cooling demand. The peak load condition only lasts a few minutes to hours per year. Moreover, the chance that the failure of one unit will happen at the peak load condition is very slim because the large chillers used for district energy plants are considered highly reliable. Combination of those two leads to a significant gap between the estimation and the actual peak load and many existing district level chiller plants suffers this problem [16]. Having extra capacity may be useful if considering the load growth because it will be needed at some point in the future; however, a chiller is very expensive equipment and having more chillers than needed add additional maintenance and operating cost.

Figure 10 shows how the plant capacity is determined based on the load duration curve of the system load. The load duration curve (LDC) is an illustration of the dynamic load as the load level versus cumulative time duration instead of the time sequential data. The LDC is commonly used to determine the system capacity. For this notional system, the equally sized five chillers are used to meet the minimum load and the N+1 requirement. As can be seen in this plot, the fifth chiller will not be used unless one chiller is down during the full load period indicated as the shaded period. Less than three chillers are needed for the rest of the time. During this period, the plant has more than one redundancy. For example, it can be said that the plant has three redundant chillers during two chillers are operating.

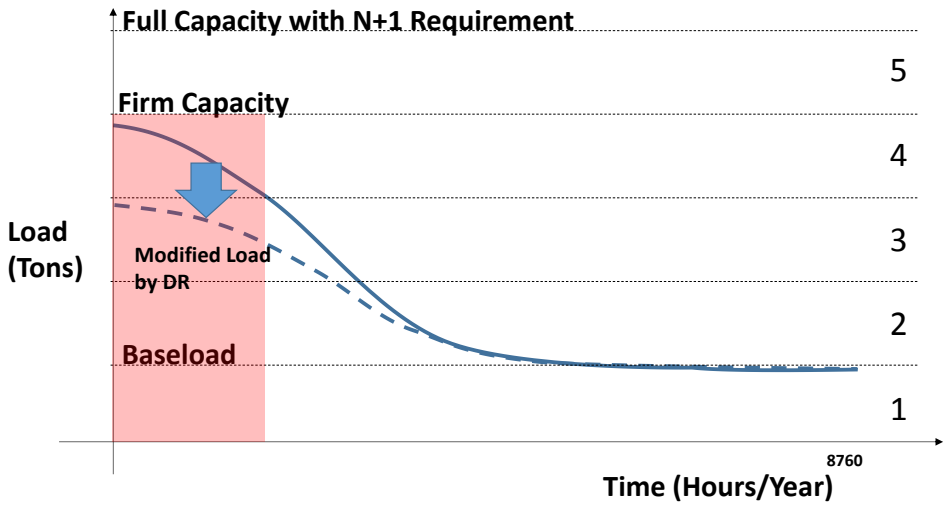


Figure 10 A Notional Load Duration Curve for Capacity Sizing

In an effort to achieve the energy reduction goals, the future buildings will be equipped with more efficient HVAC systems. Even the existing buildings can be improved with software approaches such as advanced operating and control technologies [3]. As the cooling demand is increasing, it will be necessary to improve the efficiency of the whole system. Those technologies can change the load shape permanently or instantly. These attempts can be regarded as a part of demand side management.

As various DSM programs are applied to buildings, the change in the load shape is expected. While the energy efficiency technologies require installation cost, DR can be easily applied at almost no cost if the buildings are equipped with the building automation and communication systems. The dashed line in Figure 10 indicates the modified load duration curve by DR when being applied to the original peak load. This curve may shift upward in the future with demand growth, and DR can delay the need of purchasing new chillers or the construction of a new energy plant.

In addition to the peak load reduction effect, DR can be utilized for preparing the emergency operations instead of having the excessive chillers. As explained above, the redundant chiller is only used in the contingency events. When it happens, by activating DR instead of bringing a redundant chiller online, the system can maintain its functionality. If DR can provide the same level of reliability as a redundant chiller, the plant can keep lower capacity. By deferring the substantial capital investment on equipment, the community can have financial flexibility and save maintenance and operating cost.

Expanding DR utilization can be beneficial for a community having a district energy system as in two ways: peak load reduction and avoiding plant capacity expansion. However, it is not common for a district energy system. The reason why DR is not well appreciated is lack of understanding of its value. A decision maker will be reluctant to adopt DR as a contingency plan without the comprehensive assessment of the performance and reliability of the system. A district energy system planner may pose some questions regarding expanding DR to planning practice such as below:

- Is the capacity of DR enough to affect the chiller plant?
- How much can the redundant capacity of a chiller plant be replaced with DR?
- How likely would it fail to meet the cooling load?
- How much can I save by deferring capacity expansion plan?

These questions cannot be simply answered without an understanding of how the system works and evolves with internal and external changes. A physics based model of the system that can capture the dynamic energy flow among the component connected through a network will allow a better understating of how the system works and how it will evolve with applying additional components and technologies. Wei Gu et al. also

emphasized the importance of accurate dynamic models of the system and the components for the integrated system level performance assessment of the small size energy system planning [17]. There is, however, no generalized models or tools directly applicable to the system level performance assessment of operational technologies and their impact on the capacity expansion planning due to the distinctiveness of buildings and the energy network systems. An important requirement for energy system models is that they should be able to handle its complexity and scale issues [18]. Even though high fidelity building energy system modeling tools such as EnergyPlus [19] are available and widely used for building design and energy retrofit projects, it is not feasible to build all the building models within the community with limited time and resources. Developing a building simulation model is very time consuming and simply collecting those high fidelity building models cannot capture the interactions and collective behavior of the system. To this end, this thesis proposes an integrated solution approach that utilizes data, predictive models, and an optimization technique.

1.2 Problem Statement

Oversizing of a chiller plant is the common problem for a district chilled water system. Use of the safety factor and the redundancy is due to uncertainty in demand prediction and lack of understanding of the probability of failure. DR can alleviate the problem by replacing the redundant chiller which is rarely used. The challenge of adopting DR is the quantification of the capacity value of DR and the prediction of the impact on the plant capacity plan. As the chilled water system consists of many buildings, more efficient and scaleable approach is needed than the high fidelity building simulation tools. Advancement of sensor and measurement for building energy systems allows more realistic assessment and predictions using data.

The objective of this thesis is to develop a decision making framework to evaluate and forecast the impact of DR on the operational reliability and capacity expansion planning decision of the district cooling system using data and predictive models. The expected benefit of using the proposed framework is to widen insight into the design space of multi-year planning problem by providing a set of Pareto optimal solutions.

The proposed decision making framework will be applied to the Georgia Tech's district energy system to test its validity. The real system's geographical information, network topology, and buildings' and plants' data will be used to create a model. Even though it is based on a specific system, the methodology is generic and applicable to other systems having similar characteristics and data. Figure 11 illustrates how the methodology is constructed, which steps address which research questions, and what specific methods are employed to answer the research questions.

1.3 Thesis Organization

In chapter 2, the background of the problem and related research is reviewed. Based on the literature reviews, the problem is defined as research questions and hypotheses in chapter 3. In Chapter 4, the research questions lead to formulating a methodology for the integrated decision-making framework. Chapter 5 validates and verifies the proposed models and the simulation environment and conducts experiments to test hypotheses. The overall methodology is implemented in the Georgia Tech's district energy system as a use case in Chapter 6. Lastly, Chapter7 will summarize the research problem, and the lessons learned and suggest the contributions and the future work.

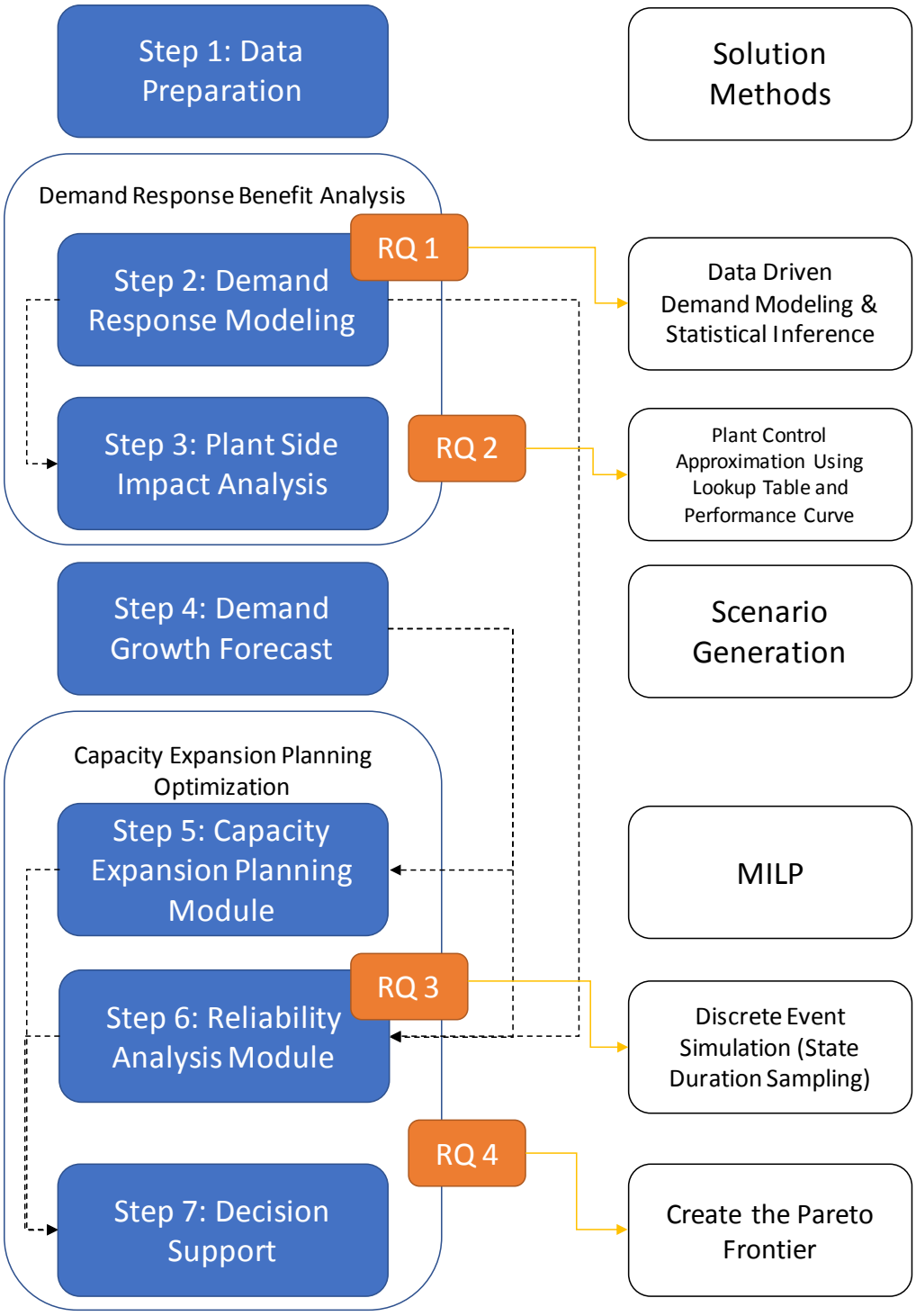


Figure 11 Overview of the Proposed Methodology

CHAPTER 2 LITERATURE REVIEW

This chapter will provide fundamental knowledge about district energy systems and the literature review of the theoretical or practical state of the art in the field of energy capacity planning and demand response analysis.

2.1 District Energy System

A district energy system provides heating and cooling energy in the forms of steam, hot water, and chilled water to a wide area where many buildings are connected to through an underground piping distribution network. As introduced in Chapter 1, district energy systems are analogous to the power grid system:

- They consist of energy generation, transmission and distribution, and distributed end users.
- Both systems are to balance between supply and demand.

While both systems share the same function and structures, their physical characteristics are different, and they face different kinds of challenges. Although the advanced operational technologies developed for power grid systems such as DR are applicable to district energy systems, its unique aspects should be investigated to formulate a right problem.

2.1.1 Characteristics of District Energy Systems

The district energy system spans from the middle to the end users of the entire energy supply chain, and the electrical chiller is an energy conversion machine. The simplified mass and energy transfer through the district cooling system is illustrated in

Figure 12. Electrical energy is transformed into thermal energy by the chillers. Thermal energy is delivered to the end users in the buildings through the medium such as water or air. Heat flux, the rate of thermal energy transfer, is proportional to the water flow rate and the difference between the supply and the return water temperature of load side as in Equation (1).

$$\dot{Q} = C_p \dot{m} \Delta T \quad (1)$$

Where,

$\dot{Q}$: heat flux

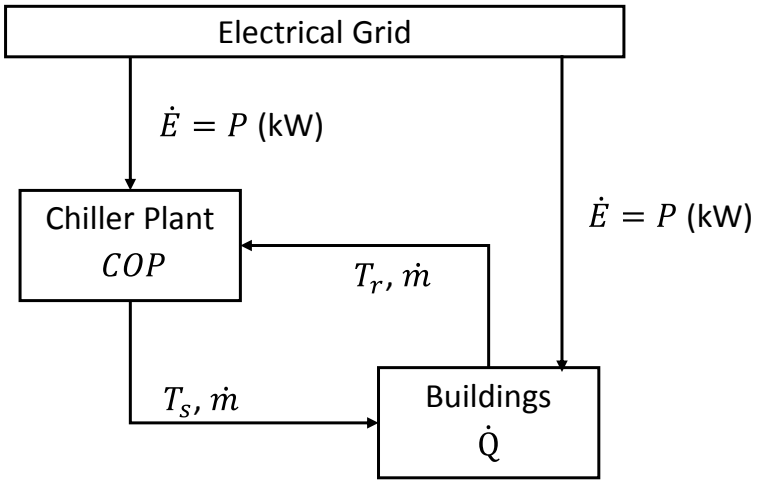
C_p: heat capacity

$\dot{m}$: mass flow rate

ΔT: $T_R - T_s$

T_R: return water temperature

T_S: supply water temperature



$$COP = \frac{power\ output}{power\ input} = \frac{3.517}{kW/ton}$$

Figure 12 Energy Flow through District Energy System

The coefficient of performance (COP) is a metric of measuring chiller's efficiency. The COP is the ratio of the produced cooling rate to the power input. It is used to measure the efficiency of chillers at the full capacity as a non-dimensionalized parameter. Both of terms in the expression are the rate of energy.

$$\mathrm{COP} = \frac{W_{cooling\ output}}{W_{power\ input}} \tag{2}$$

The performance of chillers varies with the loading condition as well as control methods. Chillers with the fixed flow rate typically have a degraded COP at the part load condition. Part load, the percentage of the design load, is the ratio of the actual load on the chiller to the full design capacity. As can be seen in Figure 13, the higher COP can be achieved at the higher part load condition for the fixed speed chiller. However, this trend can be changed by with different control schemes as depicted as the blue circle in Figure 13. System operators try to run their chillers at the highest COP as possible to minimize energy use.

$$Part\ Load\ (\%\ of\ Design\ Load) = \frac{\dot{Q}_{actual}}{Q_{design}} \tag{3}$$

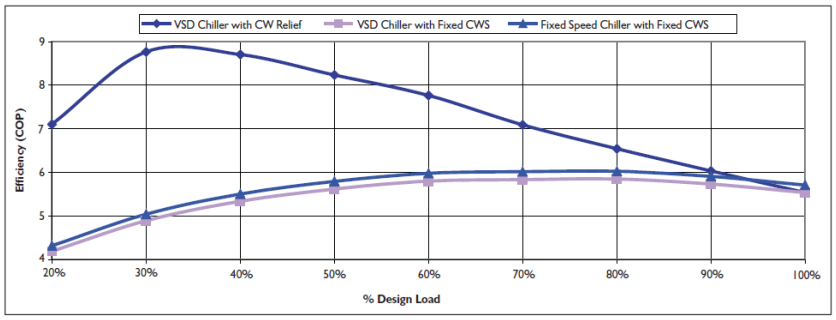


Figure 13 Part Load Chiller Efficiency [20]

The district system chiller plant consists of multiple types and sizes of chillers to meet the dynamic cooling load. This complexity of the system makes the optimal operation of the chilled water plant challenging. The performance of the system is determined by the equipment's efficiency, the system configurations, and the control sequences of chillers [20]. A study shows that the optimal control saves about 2% of the energy consumption compared to the traditional control method for the same system configuration [21]. This could be huge savings for a large chilled water system.

Traditionally, the primary-secondary system, where the chillers' loop is decoupled with the distribution network by the bypass pipe as described in Figure 14, is commonly used for campus level district cooling systems. This separation can eliminate a need for sophisticated controls for chiller staging [22]. The primary loop is maintained to have the fixed flow rate of chillers while the secondary loop pumps are controlled to deliver variable flow to the distribution network by measuring differential pressure at the end of the loop. Chiller staging is determined by the flow rate in the bypass. When the mass flow demand is higher than the flow rate of the primary loop, the secondary pumps draw water from the bypass to mix the return water into the supply chilled water. The amount and direction of the flow rate through the bypass is an indicator of chiller staging decision. By doing this, it is decoupling the controls of energy and mass balance of the system for simpler controls.

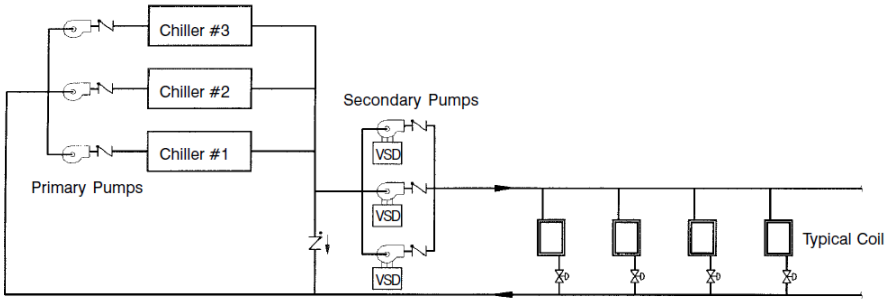


Figure 14 Primary-Secondary Loop System [23]

Despite the simplicity, this system can have undesirable phenomena due to the imbalance between the primary and secondary loop. For the design condition, delta-T (ΔT) is assumed to be constant. Ideally, the thermal energy and mass balances are well maintained without a complicated chiller dispatch control with the constant delta T. In reality, ΔT is varying with many inevitable and avoidable factors such as the set points, the control schemes, low building load, improper configurations, etc. [24] When the system suffers from the lower ΔT than the design condition, it requires more cooling flow rate ($\dot{m}$) for the same cooling demand. This results in more energy consumption on the pumps and requires more chillers to be staged on line only to meet the mass balance. Most of the chilled water system experience this "low delta-T syndrome" and it is unavoidable [24]. When the system is not properly controlled, this can be exacerbated by the flow through the bypass. If the secondary flow is larger than the primary flow, the return water flow is mixed into the supply water to raise the supply water temperature. The warmer supply water temperature causes the cooling coils to require more water and this again pulls more water from the bypass and the supply water temperature gets warmer unless another chiller is staged on-time. This phenomenon is called "death spiral." In this case, the system cannot satisfy the end users even though they have enough capacity. Up to this point, two important observations are made.

Observation 1: The system should maintain the balance of supply and demand in two ways: thermal capacity and mass flow rate.

Observation 2: Operating cost of a chiller plant is mainly electricity bills. This cannot be expressed as a simple and explicit form because the total power consumption of the chiller plant is dependent not only on the nominal chiller performance but also loading conditions, operating methods, etc.

2.1.2 Chilled Water Plant Sizing and Load Prediction Methods

Chilled water system sizing is a process to select the proper chillers and other equipment to provide the desired level of thermal comfort to the buildings connected to the network. It is important to correctly estimate the expected peak load to determine the appropriate size of the chilled water plant. Industrial guidelines provide the standardized process of HVAC system sizing such as CoolTools™ [25] and ASHRAE Handbook [26]. Those guidelines recommend steps and methods from the calculation of load to commissioning. Actions should be included in the design phase are to estimate peak load of buildings, make a list of chiller options, analyze costs, and select the best option. Rules of thumb have been used to estimate the peak load based on types of building and floor area when the specification of buildings is not available [15].

Once the load estimation is done, the selection of capacity generation options is followed. One guideline (CoolTools™) recommend the chiller procurement process as below [25].

- Step #1: Calculate Plant Tonnage
- Step #2: Develop Vendor Short List
- Step #3: Obtain Chiller Bids
- Step #4: Adjust for Other First-Cost Impacts
- Step #5: Estimate Utility Costs
- Step #6: Calculate Life-cycle Costs
- Step #7: Final Chiller Selection

These steps provide a well-defined process to follow for new plant design. When it comes to capacity expansion planning in which multiple times of decision should be made on the capacity sizing over the time, those guidelines cannot provide structured decision-making process. Decisions on when the proper timing of expansion is under the demand growth are made based on the year-to-year prediction and rely on the expert knowledge.

To pursue more accurate evaluations than the rule of thumb, building energy modeling techniques are used. The 2013 ASHRAE Handbook classifies building energy demand estimation methods into two groups: forward models and inverse models [26]. Forward modeling approaches are physics based and descriptive building simulation tools such as TRNSYS, EnergyPlus, and DOE-2. TRNSYS is a well-established tool in the building modeling and simulation community that can simulate transient phenomenon of building energy systems [27]. EnergyPlus is a popular software for a whole building energy system modeling and simulation [19]. It can model a detailed HVAC system of a building and widely used to design the system or estimate energy performance of a building before and after energy retrofit. DOE-2 is a building energy usage and cost analysis tool for the design of buildings [28]. Because the forward models are based on physics such as heat transfer, energy and mass conservations, they can provide the accurate prediction of building energy consumption including unobserved phenomena. For that reason, they are suitable for diagnostic or prognostic purposes [29]. Despite their strength, constructing forward models are challenging since they require a significant amount of modeling effort and time consumption, and extensive information of buildings as the inputs such as floor plan, building materials, schedules, and HVAC equipment, which are usually unavailable to many organizations.

Inverse models are the data-driven parametric models. They use measured building energy performance data. Mathematical correlations between input and output parameters

are determined based on the measurement. They require a small set of parameters so that the models for a large number of buildings can be generated quickly. They are, however, only applicable to the existing buildings with measured data available. The most basic model is a regression model of energy consumption as a function of the outside temperature. Since they are statistical approaches relying on historical data, inverse models may not be accurate outside of observed data range. Zhang et al. compared four different inverse models for the heating demand prediction performance [30]. The models they tested are Change-point Regression Model, Gaussian Process Regression Model, Gaussian Mixture Regression Model, and Artificial Neural Network Model. They showed that the model selection does not make much difference in terms of prediction accuracy.

Table 1 Building Energy Estimation Modeling Approaches

	Forward Model	Inverse Model
Advantages	• Can be used for the future buildings • Can generate data	• Small sets of parameters are needed • Quick and easy • Flexible structure
Disadvantages	• Detailed information is needed • Modeling and computation time	• Only applicable to existing buildings with performance measurement data available

Advantages and disadvantages of two different modeling approaches are summarized in Table 1. Due to its capability of data generation, the forward models are typically used at design phase when the real data is not available to help decision making and data-driven models are used for retrofit analysis for existing buildings. Both

approaches cannot achieve the high level of accuracy because any models cannot capture all the relationship between parameters due to the complexity of building system. The fairly generous criteria of model accuracy defined by the industrial standard reflect this aspect. ASHRAE Guideline 14 defines the acceptance limit of model error as ±10% of Mean Bias Error (MBE) and 30% of Cumulative Variation of Root Mean Squared Error (CVRMSE) for hourly models [31]. BME and CVRMSE are defined as below.

$$BME = \frac{\sum_{i=1}^{N_p}(M_i - S_i)}{\sum_{i=1}^{N_p} M_i} \tag{4}$$

$$\overline{M_p} = \frac{\sum_{i=1}^{N_p} M_i}{N_P} \tag{5}$$

$$CVRMSE = \frac{\sqrt{\sum_{i=1}^{N_p}(M_i - S_i)^2/N_p}}{\overline{M_p}} \tag{6}$$

Where, M_i and S_i are measured and simulated data at time i, p is the time interval, N_P is the number of values at interval p and $\overline{M_p}$ is the average of measured data.

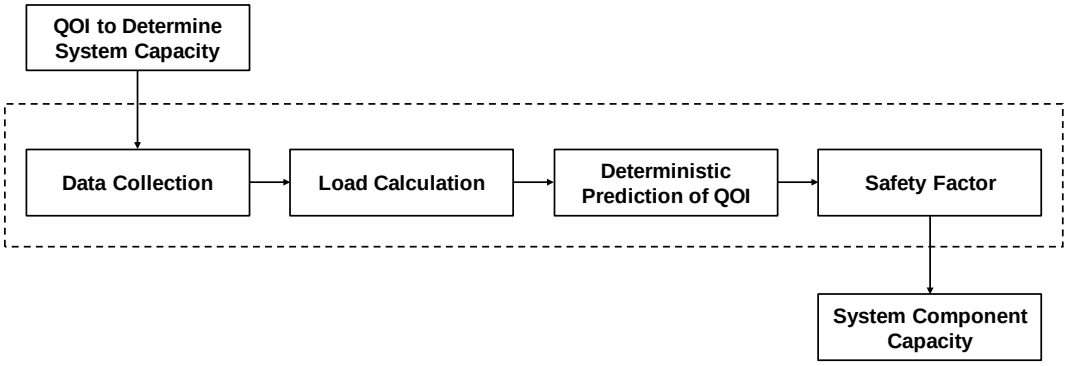


Figure 15 Traditional Sizing Framework (Modified from [32]**)**

The first step of the traditional HVAC system sizing method can be further elaborated as in Figure 15. The quantity of interest (QOI) for the chiller plant sizing is the peak cooling tonnage. Both modeling approaches are basically deterministic and yield large modeling error. To compensate this uncertainty from the building energy model, a safety factor has been used, and this leads to the oversizing of HVAC systems [32]. How the total HVAC system's capacity is determined is illustrated in Figure 16. The dashed line (10,000 tons) indicates the peak load estimation point from a deterministic calculation. Since the distribution (light blue line) around the estimation point is unknown, a safety factor[1] is added to the estimation. This is the required plant tonnage (12,000 tons) at the normal operating condition to meet the peak load. For a district cooling plant with the N+1 requirement, the total plant capacity is determined by adding the largest tonnage of chillers (3,000 tons) to ensure the system works even when the largest chiller is unavailable. As can be seen in this example, there is a very slim chance of failing to meet the peak load with this design. For a large district energy system, the load estimation by combining individual building's peak load with safety factor can lead to a large gap between the plant capacity and the actual load. This gap is also observed at the Georgia Tech chilled water system. The total plant capacity is 15,000 tons for one of their plants while the measured peak load is below 10,000 tons.

---

[1] 20% of the peak load is added in this example. It is recommended 20% to 50% of a safety factor for cooling load calculation in the industry.

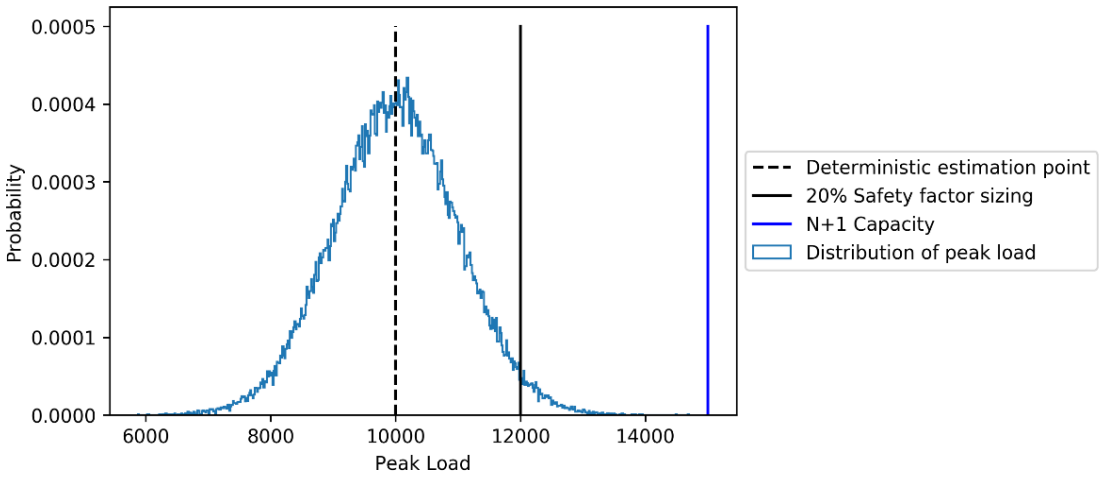


Figure 16 Deterministic Sizing, Safety Factor, and Redundancy

The probabilistic approaches of load prediction and HVAC system sizing are proposed in recent research to deal with the uncertainty and avoid the oversizing problem. Sun et al. developed a modeling and simulation framework to get a probabilistic distribution of the peak load based on the annual weather profile by using random input variables and Monte Carlo simulation (MCS) [32]. They use the value at 99.6% of all load points over the year from simulations instead of the deterministic peak load estimation to determine the HVAC system size. Pei Huang et al. also used MCS and probability of success to find the optimal size of building HVAC systems under uncertainty [15]. For these studies, building simulation models are developed using EnergyPlus, and the random variables for the inputs are used to quantify the uncertainty. Using the building simulation models for a district system is not a feasible approach because developing dozens or hundreds number of buildings' model are time consuming and requires too much information. Even MCS for all the building models is computationally expensive.

2.1.3 The Analogy to Electrical Grid

Although the probabilistic evaluation of load may resolve the oversizing problem due to the deterministic peak load estimation and using the safety factor, there can be a further improvement using demand side management concepts from the power grid system to reduce the peak load [33]. District energy system can be analogous to the electrical grid in terms of function and structure. Both systems consist of energy generation, transmission and distribution, and end users. The primary objective of managing the system is to balance between supply and demand at the least cost. As demand is growing over time, the system operator should cope with the capacity expansion to maintain the functionality. Therefore, the electrical grid system's demand side management and more advanced planning methods can be applied to district energy system.

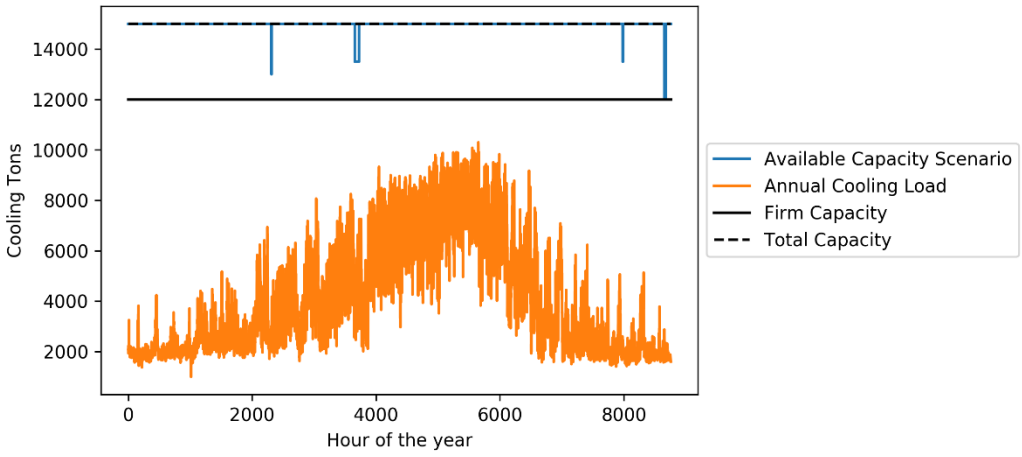


Figure 17 Annual Load Profile and Plant Capacity

Combining the safety factor and the redundancy, the system might have too much unused capacity during a year. It is illustrated how the system runs under the capacity during the entire year in Figure 17. This graph is generated using real data from Georgia Tech chilled water system. Cooling load depicted as orange color unlikely reaches the

design firm capacity, and the plant has a large margin between the available capacity and the actual demand during the entire year. Blue lines indicate the total available capacity considering the failure of one or more chillers. This varies with scenarios. The redundant chiller would never be used unless multiple chillers are unavailable due to failures at the moment of the peak load condition for this system. Therefore, this plant is keeping the redundant chiller for the rare event. Instead of having the redundant chiller, demand response can be utilized as a contingency plan to reduce cooling load instead of bringing another chiller into the network. To have DR as a contingency plan instead of a redundant chiller, comprehensive understanding of its impact on the system performance and reliability is required.

In spite of the similarities, some critical differences restrict the direct application of existing methods in the electrical grid to district energy system. Thermal energy is transferred via a medium such as water and air, and this causes a lag between the moment of actual demand and energy production, unlike the electrical grid where the energy is delivered at the light speed. This physical characteristic confines the scale of the system so that the plants of district energy system provide energy only to the adjacent area within a few miles. Because of the energy storing function of the medium, district energy systems do not experience the "blackout" even when the capacity is less than demand. It still can operate at some degree of performance degradation. Today's grid systems are mostly liberalized, but there is no market structure for district energy systems, and it is more like the traditional vertically integrated utilities. Therefore, one organization manages and operates the entire system within the community. In a modern grid system, energy can be imported from a neighbor rather than load is shed. In district energy system with one plant, the load should be shed when the capacity is insufficient.

Due to these differences, district energy systems do not face the same problems with the modern grid systems such as competitiveness or stability issues. This helps mitigate the problem of complexity, but another challenging aspect should be addressed. Due to lack of awareness of the need for the optimal planning and demand side management, there is no standardized modeling framework and performance or reliability metrics in district energy systems. To develop a more rigorous decision making framework for the district energy system capacity expansion planning, what can be borrowed and need to be modified from other disciplines will be addressed in the next few sections.

2.2 Demand Response

This section will provide more specific definition and history of Demand Response, address its benefits and limitations, and how DR resources are modeled and implemented for electrical grid and district energy systems to answer the following research question.

Primary Research Question: *What is the appropriate approach to analyze DR resources and evaluate the impact on the capacity planning of a district energy system?*

Demand response (DR) is defined as below by U.S. Department of Energy [7].

> *"Changes in electric usage by end-use customers from their normal consumption patterns in response to changes in the price of electricity over time, or to incentive payments designed to induce lower electricity use at times of high wholesale market prices or when system reliability is jeopardized."*

Therefore, it is expected that DR impacts immediately on the system by changing usage pattern temporally. While the main purpose of DR is to improve the operational efficiency and reliability in a short-term manner, it can help avoid the long-term capital investment on generation capacity because the peak load, which occurs only a few hours

per year, can be reduced by DR so that it could remove the necessity of building power plant to serve the peak load [7]. Since the concept of demand side management is introduced in the 1970s, technological barriers such as lack of advanced metering and communication systems have been hindered the deployment of DR for decades. The recent advancement of monitoring and communication technologies resolved the issue so that various DR programs are already in use [34]. One of the remaining challenges of DR, especially for capacity planning is to establish the value. Due to the stochastic nature of demand, DR is regarded as unreliable resources compared to the traditional generation technology, and this obstructs DR to be accepted as resource adequacy to prepare the extreme conditions [35]. To replace the traditional capacity resources with DR, its availability and reliability should be ensured [7].

Since the buildings' energy play the important roles in DR programs, the same practice can be applied to district energy systems. Even though there can be various types of DR applications such as lighting and appliances of residential buildings or commercial and industrial process loads that can be shed or rescheduled, only controlling of building HVAC systems is considered in this thesis. A few efforts to implement DR for district energy system operations to reduce the peak load has found from the literature review. For example, a centralized agent based control scheme is applied and tested for the district heating system in Sweden [36]. They achieved 15-20% of peak load reduction in the field test. Basciotti and Schmidt developed a modeling and simulation environment of district heating systems to test various demand side management methods [37]. They tested different size of thermal storages and a load shifting method and showed that the peak load reduction could be up to 35% from the night setback control and 14% from storages. Li and Wang developed a multi-agent control method for DSM in district heating system. They demonstrated how the load side management benefits to district heating operation by

smoothing out the load profile [38]. These experiments are promising, but the variation of the achievable peak load reduction is too high to be accepted as reliable capacity resources and no research attempt to incorporate DR into the capacity planning problem is found. Unless the chilled water system sizing process evaluates the risk of this uncertainty from DR, a chilled water system manager will not adopt DR as a replacement of redundant chiller. To address this, quantification of DR benefits and risks is essential. In the following section, how the previous research works addressed DR in energy planning problems and various modeling approaches for DR resources will be introduced to identify an appropriate DR assessment method under uncertainty.

2.2.1 Review of Demand Response Assessment Methods

Demand response models have been developed in two different perspectives. The first perspective is for the very high-level policymakers. Price elasticity model is commonly used to represent how customers would react to the price signal to change their usage patterns under the incentive based DR program. The price elasticity, E is defined as the demand sensitivity to the price [39].

$$\mathrm{E} = \frac{p_0}{d_0} \cdot \frac{\partial d}{\partial p} \tag{7}$$

Where, p_0 is the original price, d_0 is the original demand. Large value of the price elasticity means that customers are more sensitive to the price change. For the time varying price, the differential demand at time h corresponding to the price change is expressed as below.

$$\Delta \mathrm{d(h)} = \mathrm{E(h)} \frac{\Delta p}{p_0} \cdot d_0(\mathrm{h}) \tag{8}$$

This concept is widely used to incorporate the market based DR for economic or policy-driven energy planning models. For example, De Jonghe et al. proposed an LP based formulation for the long-term generation investment planning model to integrate DSM into the planning problem using the price elasticity model [40]. A bi-level generation and transmission expansion planning problem is formulated as MINLP to minimize total cost where the peak load is reduced by DR which is modeled using price elasticity in [41]. It is the useful conceptual model for the policy optimization problem to find optimal price to promote participation of customers into DR program. However, historical data of the relationship between the varying prices and demand is required to construct the price elasticity, or it should rely on assumptions. Moreover, its underlying assumption that the customers will respond to the price change is not applicable to the campus level district energy systems where the market structure does not exist.

Huang and Billinton developed a simple algebraic model to study the effectiveness of DSM on the generation adequacy [42]. This model has the targeted value of the peak load and assumes reduced peak load is recovered at the off-peak hours. Using this model, peak clipping and valley filling concepts can be demonstrated by modifying the original load shape. They showed how those DSM concepts would reduce the peak load and improve the system radiality using the simulations of a 48-hour time period on IEEE Reliability Test system. This approach is a good proof-of-concept model, but it does not guarantee the same benefit would be achieved for the real world problem because it does not implicate any specific technology and details of practice.

$$\overline{\mathrm{L(t)}} = \begin{bmatrix} P & t \in \Omega \\ L(t) + A & t \in \Psi \end{bmatrix}$$

$$A = a\left[\frac{\sum_{t\in\Omega}(L(t) - P)}{N}\right]$$

Where,

P: pre-specified peak

$L(t)$: basic load

$\overline{L(t)}$: modified load (9)

ψ : set of off-peak hours during which the energy is recovered

Ω : set of on-peak hours during which the energy is reduced

A : MW load added to each off-peak hour of ψ

N : number of off-peak hours in ψ

a : percentage of the energy reduced during on-peak hours that is recovered during off-peak hours

The price elasticity and load shape models are hypothetical DSM model, and they do not specify any methods or resources of how the system is managed. They are useful for the very high-level decision making to see if DSM would be beneficial for the system and easily incorporated into mathematical optimization models, but cannot provide the accurate estimation of load reduction.

Another stem of DR modeling approaches is the more detailed and technical method using simulation models to control DR resource at the building level. In the same way of the load estimation, forward modeling tools are used to create building HVAC system models and implement control methods for DR such as EnergyPlus in [43]–[45] or Dymola/Modelica [46] in [47], [48], which uses DOE-2 based building component model library. While these are more accurate and reliable models for a specific building, it is very

challenging to apply them to the wide area with a large number of buildings because the building simulation model requires specific building information, and intensive modeling effort and time should be put in developing all the buildings. Therefore, this is not a feasible approach for the district level system with hundreds of buildings in terms of scalability.

To overcome this limitation, more simplified building models are suggested in [49]. They developed EnergyPlus models for various types of commercial buildings and created a piecewise linear regression model to estimate DR potential from the set-point control of the HVAC systems for each type of building. Along with the residential building's two state capacitance-resistance model, the regression models are used to calculate DR potential at the system level. However, this approach still requires developing each type of building model, and the goodness of fit for the regression model is too low - the R^2 value of their commercial building models is between 0.54 and 0.78. Poor fitness of model may result in misleading conclusion. For example, although they claimed that the DR potential should increase as the outdoor temperature rises, even their data does not agree with the claim as illustrated in Figure 18. These plots are generated with data presented in [49]. Measured data is just scattered without any apparent trend unlike the model data. The linear least square fits even show the reverse trends of the model.

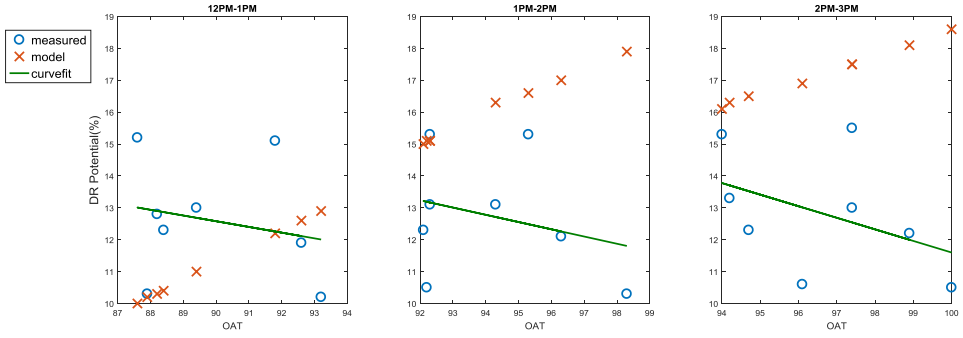


Figure 18 Measured vs. Model Estimated DR Potential (%)

Conceptual models have been used to establish the value of DR when data is not available. The building simulations models provide a capability to generate data based on physical assumptions. In recent years, various DR programs have been deployed, and the measured data enables analysis of DR benefits in real life. There are some efforts to characterize and evaluate the DR benefits using data, but not many because data is usually not available to the public. A group from Lawrence Berkeley National Lab has been conducted a series of research on DR data analysis using DR data in California [50]–[52]. They used a regression model to establish a baseline using historical data as a function of time. Demand shed by DR is defined by the difference between the model and the measurement as a percentage of the baseline load. What should be pointed out from their findings is that the error and variability of DR parameters are huge. They could not find any evidence that DR effectiveness has a correlation with the year, buildings or outside temperatures [52]. This finding agrees with the data validation presented in Figure 18; however, the reason why any trend is observed can be that the large error and variability of DR conceal the actual trend. Therefore, the challenge of predicting DR benefits is the uncertainty from various sources results in error in the estimation which cannot be reduced.

2.2.2 Conclusion

This section reviewed different modeling approaches of DR resources from the literature to address the research question: What is an appropriate approach to analyze DR resources and evaluate the impact on capacity planning for district energy system?

Planning models are simple, but deterministic and should rely on assumptions to estimate DR benefits. Building HVAC simulation models provide a more accurate estimation of DR benefit based on physics. However, they require a lot of modeling effort and computations. It is not feasible to create all the building models to evaluated DR for

the system capacity. From the perspective of plant management and operation, distinguishing buildings is not an effective way because a plant only sees the aggregated load. Building HVAC simulation models are more suitable for a building owner who is willing to opt in DR program and wants to evaluate the benefit.

Lack of data and experiences has made many assumptions on how DR works and be modeled. As data is becoming available from the DR practice, data analytics can bridge this gap. If the measured demand shed by a specific DR method is available for the few sample buildings, it can be extrapolated to other buildings using statistical guess. As the error and variability are unavoidable, the uncertainty of DR impact on capacity planning should be quantified. From the multiple times of DR events, distribution of DR benefits can be obtained. Combining data analytics of DR and a planning DR model, a probabilistic load shape can be created. This section answered partially to the primary research question. More detail of this approach will be discussed in the methodology development chapter.

2.3 Reliability Analysis

The classical definition of the reliability of a system is the probability that it will perform its required function under given conditions for a specified period [53]. Redundancy is the most basic and commonly used in system design to increase the reliability of the system by duplicating critical components. N+1 redundant system means that N number of components provide the normal service capability with one additional backup unit to ensure the system normally being operated even one element is not functioning.

To evaluate the opportunity of adopting DR instead of a redundant chiller, reliability analysis should be performed. Redundancy is a deterministic approach to mitigate risk. It is not fair to compare the potential capacity of DR as demand reduction

and the chiller capacity. Meeting the peak load is important, but the failure does not always happen at the peak time. Evaluation of the system reliability only at the peak load condition may overestimate the risk.

Redundancy cannot capture the dynamics of a district energy system. A function of a district energy system is to supply energy to end users. Cooling demand is varying with time so that the system reliability is dependent on demand not only the plant status. A chiller failure does not mean that the system cannot serve the load because this is an only mechanical failure of the plant side. While it is important that the system be capable of serving the peak load, it is not necessary that all the chillers should be available during the off-peak period. Therefore, redundancy is not enough to evaluate the true reliability of dynamic energy systems. For that reason, in the power system community, the various metrics have been developed to evaluate the reliability of the system at different levels. Since there is no standard metric to measure the reliability of district energy systems, the concepts of reliability evaluation can be borrowed from the power grid system.

2.3.1 Reliability Metrics for Power System

Power system consists of generation, transmission, and distribution systems. Various reliability metrics are defined for each hierarchical system to account for their characteristics and functions. Generation system is required to maintain the required capacity while transmission and distribution systems should maintain the network reliability. A chiller plant of a district energy system is equivalent to the generation system of the power system. Some of the useful adequacy indices for a power generating system are introduced here.

Reserve margin is a basic reliability measure of the generating system, which represents the surplus capacity over the required amount to serve the peak demand. It is

defined by the ratio of the excessive generation capacity and the annual peak system load [54]. This is a deterministic measure of the generating system reliability to ensure the system can serve the load at an emergency event. A redundancy of chiller plant can be converted to an equivalent reserve margin value. For example, if a chiller plant has six 2,000 Tons chillers for the 10,000 Tons of peak load, it is a 5+1 redundant system and can be said to have a reserve margin of 20%. Keeping a high reserve margin causes high capital and operation cost while a low reserve margin may result in high risk of an outage. Typically, 15% is a target reserve margin in the U.S.

$$\text{Reserve Margin}(\%) = 100 \times \frac{Capacity - Load}{Load} \tag{10}$$

Solely reserve margin is not enough to evaluate how the system is functioning because the peak load condition is only a few minutes of a year. The term *adequacy* is used to define the capability of the system that can satisfy the consumer load demand or system operational constraints [55]. To get the generation adequacy metrics, simulation models are required as described in Figure 19 because they are statistical parameters.

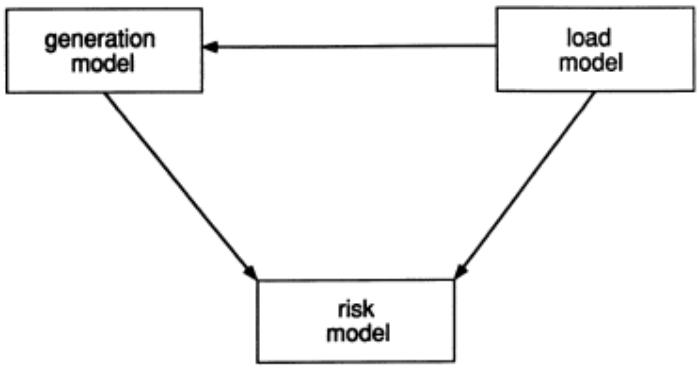


Figure 19 Basic Concepts of Power System Reliability Evaluation [55]

Loss of load expectation (LOLE) or loss of load probability (LOLP) is commonly used for the generation adequacy assessment in a probabilistic way. They are the expected duration of outage in a given period, usually measured as the number of days or hours per year. Once LOLE is defined by Billinton as days (or hours)/year, LOLP is suggested to nondimensionalize as LOLE/N, where N is the number of time increments in the LOLE calculation [54]. LOLE of one day per ten years is usually considered as the desired reliability, which is about 0.0003 in the LOLP metric [56]. There is an attempt to use the concept of LOLE for the design of district energy systems. Gang et al. developed a robust optimal design method for district cooling systems, and they used "the unmet hours" to evaluate the system reliability [57]. To compensate the shortcoming of the LOLE index, which does not indicate the severity of the failure, the LOEE (Loss of energy expectation) can be used. These metrics are calculated from ENS (Energy not supplied) and LLD (Loss of load duration) as illustrated in Figure 20. This represents one sample case from the simulation. The indices are calculated from Monte Carlo Simulation.

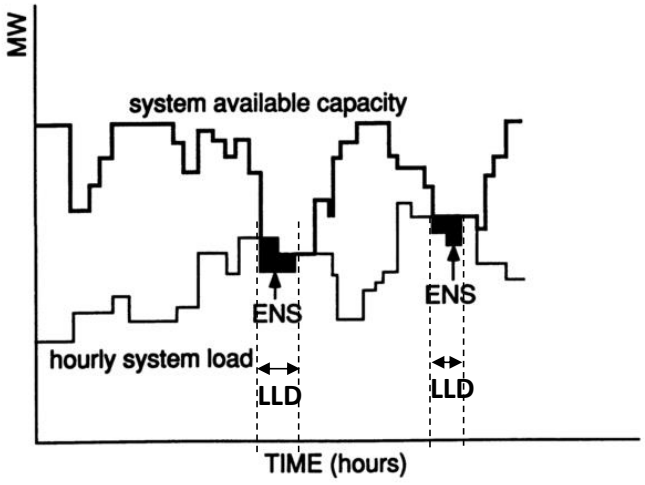


Figure 20 ENS(Energy Not Supplied) and LLD(Loss of Load Duration) are obtained by observing the shortfall of available capacity and load [55]

Effective load carrying capability (ELCC) is another reliability index developed to evaluate an individual generating unit. ELCC is used to estimate the capacity value of power plants, and it decomposes the individual generator's contribution to system reliability [56]. It is defined by the amount of new load that can be added to a system while maintaining the initial LOLE after a new unit is added. The concept of ELCC is illustrated in Figure 21. When an additional generating capacity is added to the original system (the blue curve), 400MW of the peak load can be added to the new system (the green curve) to keep LOLE of 0.1days/year as the reliability target level. Thus, ELCC of the new generation is 400MW. ELCC is used to evaluate the capacity contribution of new types of technologies such as wind power [56] or demand response as the equivalent generation capacity [35].

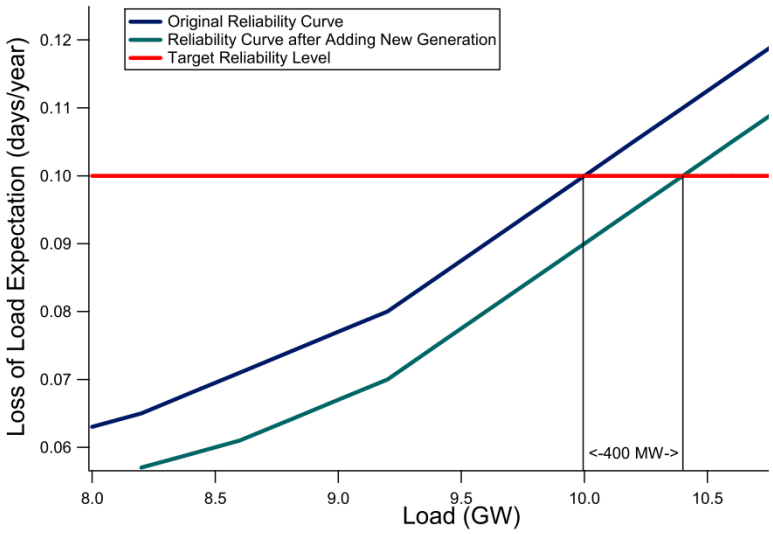


Figure 21 LOLE vs. Peak Load with ELCC of additional generation [56]

2.4 Energy Capacity Planning Optimization

As briefly introduced in Chapter 1, the traditional capacity expansion planning is performed by a top-down process based on simple assumptions and the expert knowledge. There is no rigorous logic behind selecting the best plan especially for determining when new chiller should be installed. It only follows the demand growth curve to find the point when the existing system cannot meet the requirement. This is a myopic approach that only considers one time period, typically a year, to determine if a new chiller is needed. In a longer planning horizon, where multiple times of additional chillers are required, the optimal solution might be different over the entire planning horizon. Without a rigorous method to find the best solution, this process may require several times of iterations based on guesses to get a sub-optimal solution.

This section defines the characteristics of capacity expansion planning problem and reviews the decision making methods widely used in the energy sector to select a suitable method for capacity expansion planning of a district energy plant.

2.4.1 Energy Planning

The energy supply chain starts from the primary energy, which is a natural energy resource such as oil, gas, solar, etc. The primary energy is transformed into another form of energy multiple times through the energy supply chain until it is delivered to the end users as presented in Figure 22.

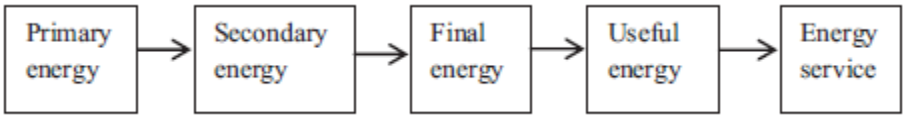


Figure 22 Energy Supply Chain [58]

Throughout the chain, there can be many options to obtain and deliver energy to the next level. Prasad et al. collected definitions of the energy planning from different authors and summarized them as below [58]:

"Energy planning is a process to find the optimal solutions that can meet the future demand while taking into account many different aspects of the community such as political, social, and environmental concerns."

Energy planning problems are categorized by the planning time horizon as described in [59]. Short term planning is to determine the operational decision for the next hours or up to one year such as how to dispatch energy generators. The purpose of short-term planning is to maintain the reliability of the system with available resources at times. Medium-term planning is to make a tactical decision for the next several years to meet the demand. For this term, introducing new technologies can be considered [58]. The purpose of long-term planning is to prepare new infrastructure and technologies to serve the increasing demand of the future in time. Uncertainty involving with technological and political changes in the future makes this problem challenging [59].

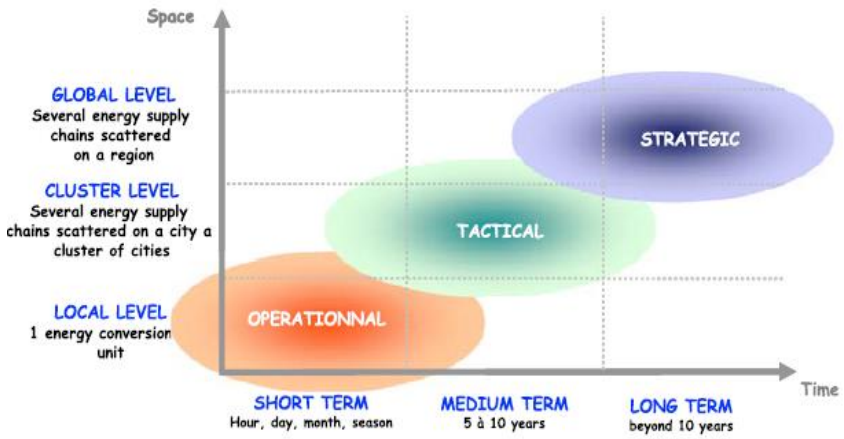


Figure 23 Levels of Energy Planning Decision Making [59]

Energy planning is conducted by multiple entities from national governments to small communities throughout the energy supply chain to manage resources and demand within their boundaries and exchange energy with other entities. Decision making structure and its temporal and spatial scale are illustrated in Figure 23.

For long-term energy planning, the investment decision on generation capacity of various energy technologies to maintain the balance between demand and supply is one of the most important problems. This is called capacity expansion planning. In general, capacity expansion planning is a process to determine the future expansion's timing, sizes, locations, and the types of product families [60]. A basic capacity expansion process is illustrated in Figure 24. For a given demand growth projection, optimal timing and size of additional capacity are determined to minimize total cost. As a mathematical formulation, it can be expressed as Equation (11).

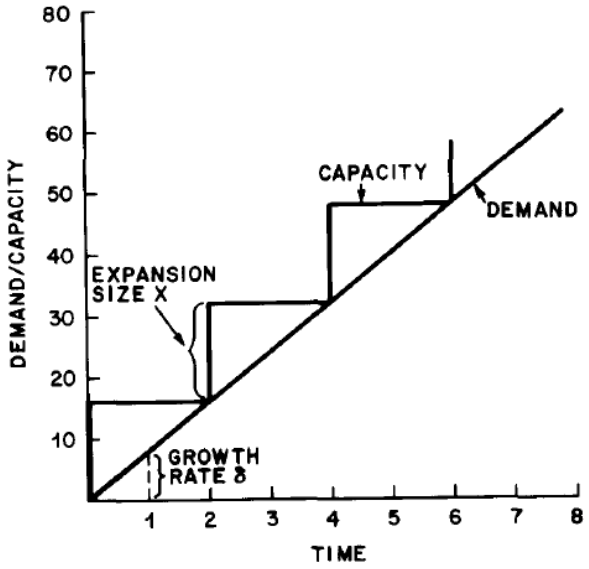


Figure 24 Capacity Expansion Process [60]

$$\min \sum_{t=1}^{T} \sum_{i=1}^{N} f_{i,t}(x_i) \times C_{i,t}(x_i) \quad (11)$$

$$\text{s.t. } D_t < \sum_{i}^{N} C_{i,t}(x_i) \text{ for } \forall \text{ t} \quad (12)$$

where,

x_i: ith Capacity option

$f_{i,t}$: Cost function of the unit capacity of ith technology at time t

$C_{i,t}$: Capacity of ith technology at time t

D_t: Demand at time t

A district energy system is typically below the local level because of its limitation of network size while planning period covers all across the time frame because its management and operations are independent of power grid systems. A chilled water plant expansion problem fits into the basic capacity expansion planning problem.

2.4.2 Review of Capacity Planning Decision Making Methods

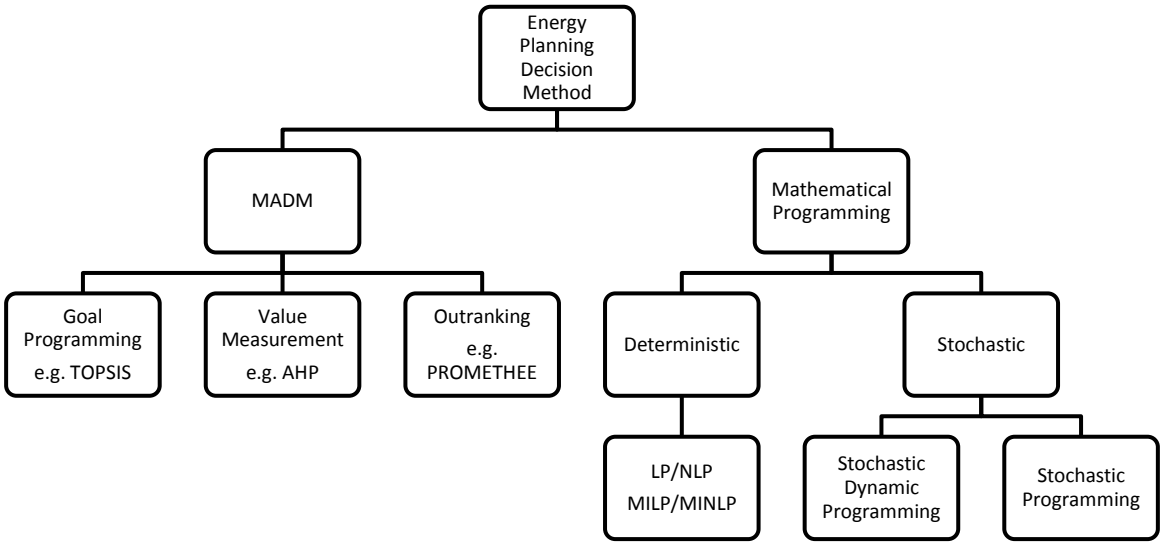


Figure 25 Decision Making Method in Energy Planning

Besides the practical static planning introduced in Chapter 1 that uses a static and deterministic estimation of the future peak load and relies on the expert knowledge to select the best candidates, more formal decision making methods found from the literature review for energy planning problems are summarized as presented in Figure 25.

In the early years of generation expansion planning research since the 1950s focused on linear programming (LP) to find the least cost solution [61]. Assumptions behind LP formulation is that the objective function and constraints are expressed as linear functions of parameters. As the grid system is becoming more complex, research on generation expansion planning has branched out in many directions to overcome the limitation of LP. The nature of discrete capacity increments makes LP invalid [54]. To handle nonlinear and discrete variables, various nonlinear programming (NLP) and mixed integer programming (MIP) techniques are employed. Environmental concerns and

different stakeholders created conflicting objectives. Multi-criteria decision making techniques (MCDM) became popular since the 1980s when the environmental impact was considered important in energy planning to deal with conflicting objectives [62]. MCDM is further divided into multi-attribute decision making (MADM) and multi-objective decision making (MODM).

Løken identified three categories of MADM techniques that are applicable to energy planning problem [63]. They are value measurement models, goal, aspiration and reference level models, and outranking models. Value measurement models are scoring methods to evaluate alternative solutions quantitatively. Evaluation criteria are summarized into the single scoring function with some weights by various techniques in these types of methods. A representative value measurement model is the analytical hierarchy process (AHP) and commonly used in energy planning problem. AHP has been used for the optimum location selection of renewable resources [64], the optimal chilled water network design selection [65], and the selection of space heating systems in an industrial building [66]. Next class is goal programming, which identifies the ideal solutions and finds the closest one as the best solution among alternatives. The technique for order preference by similarity to ideal solution (TOPSIS) is one of them. The last one is the outranking model that compares alternatives pairwise to select preferred one for all the criteria. The elimination and choice translating reality (ELECTRE) and preference ranking organization method for enrichment evaluation (PROMETHEE) are the most popular methods in energy planning problems. PROMETHEE is used to find the best energy resources among four renewable energy options for different stakeholders for district heating system in Vancouver [67]. In this research, different results are drawn depending on the communication scenarios between stakeholder groups. More extensive review of those methods and applications are available in [62] and [63].

MADM techniques are simple and transparent tools to find the best solution among a small set of candidates by preferences of decision makers and capable of evaluating alternatives based on both of quantitative and qualitative criteria. To use MADM techniques, all the alternatives are precisely defined, and decision makers should be aware of them [61]. In many cases, alternatives are not defined at the planning phase, and the number of alternatives is too large especially for long-term planning. They are suitable for static design optimization problems with a limited number of down-selected alternative options rather than multi-period optimization problems such as capacity expansion planning.

As stated earlier, variant mathematical programming techniques have been developed to solve energy planning problems and various industrial applications of capacity planning problems [68]. Most of the previous research focused on how to formulate complex problems as mathematical programming and find solution approaches for that specific problems. Using mathematical programming, the optimization problems can be formulated as either single objective or multi-objective problems. To overcome the limitation of LP that cannot deal with discrete time intervals and capacity unit sizes, mixed integer programming (MIP) models have been used for various situations. For examples, mixed integer linear programming (MILP) models are developed for generation expansion planning with mid-term scheduling [69] or unit commitment [70].

Since the concept of DSM is introduced, many researchers have attempted to incorporate DSM into planning optimization problem. Hobbs developed a simple mixed integer linear programming (MILP) based on the existing LP model to include DSM options for resource planning [71]. In his model, binary variables indicate whether a particular DSM program will be implemented at a specified year to minimize total investment and operating costs over the planning horizon and the impact of each DSM

program is models as demand reduction by the program. Antunes et al. formulated power generation expansion planning as multi-objective mixed integer linear programming (MOMILP) to incorporate environmental impact and cost into objective functions [72]. Like in Hobbs' model, DSM is represented as a simple load reduction program. Lohmann and Rebennack used mixed integer nonlinear programming (MILNP) to incorporate short-term demand response into a long-term planning problem [73]. In their model customers are assumed to respond to electricity price to change usage patterns, and the demand function is modeled as a nonlinear function. Similarly, the price responsive model is commonly used to incorporate DR into existing energy planning optimization problems [40], [74], [75]. Above mentioned models used the simple assumption of demand reduction by DSM programs or price elasticity model to find the optimal pricing policy.

Uncertainties and risks are inevitable for long-term energy planning because of the stochastic nature of demand. The high penetration of renewable energy resources into the generation mix has brought more uncertainty to the supply side. This situation urges to take uncertainty into account for energy planning. Some of the commonly used approaches are scenario analysis, two-stage stochastic programming, and stochastic dynamic programming.

Scenario analysis or sensitivity analysis is an approach to evaluate what-if scenarios using a deterministic optimization model. Input variables are created based on scenarios, and decisions are made for each scenario. It has been widely used to manage the risk of the decision in energy planning problems under uncertainty from uncontrollable factors such as economy or policies. For example, fuel prices and CO2 prices uncertainties are formulated as scenarios and sensitivity analysis is performed using a deterministic MILP formulation for Greek power system planning problem in [69]. While it is simple and easily

implementable with an existing model, it does not provide the single best solution. To overcome this, more complex optimization problems are formulated.

The two-stage or multi-stage stochastic programming technique is used to find the robust optimal solution under uncertainty. It is commonly used for co-optimization problems for different time scales such as capacity planning and operational decisions. There are abundant attempts in academia to solve co-optimization problem using two-stage stochastic formulation (e.g., the capacity planning and dispatch optimization under high wind energy penetration [75], the day ahead robust unit commitment optimization [76], scheduling of building energy system operation [77], and other applications are available in [78]–[80]) because the assumption behind the logic is that the decision can be made with certain information 'here and now' and then corrective actions can be made once the uncertainty is realized. The first stage is to solve a longer period planning problem such as the investment optimization problem and the second stage is to solve the generation dispatch problem with the fixed capacity plan. The realized uncertainty scenario at the second stage allows the problem as a deterministic formulation. This is, however, the main drawback of the stochastic programming because the scenarios of uncertainty realization should be constructed, which requires knowledge or historical data, and the number of scenarios becomes intractable for large scale problems. This hiders the stochastic programming approaches to be used in real life problems in practice [81].

Another stem of stochastic mathematical programming approaches is stochastics dynamic programming. Dynamic programming is a classic problem formulation method to decompose a large problem into small sub-problems to get the optimal solution by combining the solutions of sub-problems. It is a similar idea to multi-stage stochastic programming as a sequential optimization method, but the difference is that this method is

based on Bellman's Principle of Optimality. Bellman's Principle of Optimality is as follows [82]:

> *"An optimal policy has the property that whatever the initial state and initial decisions are, the remaining decisions must constitute an optimal policy with respect to the state resulting from the first decision."*

If the transition from one state to another state is stochastic and the stochastic process has the Markovian property – the conditional distribution of the future state depends only on the current state regardless of the past, then the problem can be formulated as a stochastic dynamic programming or Markov decision process.

While there is some research to demonstrate the applicability of stochastic dynamic programming for energy planning problems [83][84], it is not widely used in long-term energy planning problems compared to the stochastic programming approach. It is more suitable to the case with a small set of decision (or actions) and infinite planning horizon such as maintenance optimization problems.

2.4.1 Challenges of the Problem Formulation

The objective of capacity expansion planning optimization is to minimize total cost over the planning horizon while ensuring the system serve the load. Decision variables are logically discrete because it should be chosen from available commercial chillers that have specified capacity size and the decisions are made at discrete time step. Therefore, mixed integer program is suitable to formulate the capacity constraint and the capital cost function. A deterministic investment optimization can be formulated as mixed integer linear program with demand satisfaction constraint. The Equation (11) can be rewritten as a linear form with redundancy constraint as below.

$$\min \sum_{t=1}^{T} \left(\frac{1}{1+\gamma}\right)^{t} \left[\sum_{i=1}^{N} (IC_i xn_{it}) + OC_t\right] \tag{13}$$

Subject to

$$\mathrm{D}_t + \max(C_i) < \sum_{i}^{N} C_i(xe_{it}) \text{ for } \forall \mathrm{t} \in [1, \mathrm{T}] \tag{14}$$

$$xe_{it} = xn_{it} + xe_{it-1} \text{ for } \forall \mathrm{i} \in [1, \mathrm{N}] \text{and } \forall \mathrm{t} \in [1, \mathrm{T}] \tag{15}$$

$$xe_{it} = xe_{i0} \tag{16}$$

$$\sum_{i=1}^{N} xe_{it} \leq n_{max} \text{ for } \forall \mathrm{t} \in [1, \mathrm{T}] \tag{17}$$

where,

xn_{it}: binary decision variable for a new chiller at time t

xe_{it}: discrete variable for the number of existing chiller

IC_i: investment cost of ith chiller option

C_i: the catpacity of ith chiller option

D_t:p eak demand at year t

γ: discount factor

N: number of chiller options

T: planning period

n_{max}: the maximum number of chillers for a plant

Recall the two observations made in Section 2.1. Those characteristics of the system make MILP formulation infeasible for district energy capacity planning problem.

Observation 1: The system should maintain the balance of supply and demand in two ways: thermal capacity and mass flow rate.

Observation 2: Operating cost of a chiller plant is mainly electricity bills. This cannot be expressed as a simple and explicit form because the total power consumption of

the chiller plant is dependent not only on the nominal chiller performance but also loading conditions, operating methods, etc.

If one wants to add operating cost (OC_t) for the MILP formulation, it should be expressed as an explicit linear function of decision variables. The observation 2 indicates that it is very challenging. The COP is a nonlinear characteristic curve of chillers and it is changed by the operating method. Combination of different chillers make it impossible to formulate the operating cost as an explicit function of the decision variable xe_{it} and xn_{it}. To calculate operating cost of a plant, a simulation model is required with a control design of chiller dispatch. However, designing control method is out of scope for this thesis and it is another extensive research topic.

The demand satisfaction constraint only accounts for the peak load condition. From the observation 1, it is concluded that demand satisfaction should be evaluated for both of thermal and mass balance. This will affect formulating constraints and calculation of reliability metrics. Since they are not independent of each other, care should be taken when designing and operating a chiller plant. This will be further investigated in the next chapter.

Incorporating DR brings more difficulties to the problem formulation. First, none of the existing DR representation models reviewed in Section 2.2 can be directly used. Since a district level energy planning is conducted for much smaller scale than a national energy planning problem, the conceptual model without a realistic estimation of DR capacity may result in failures because the solution will be more sensitive to the error of prediction for the smaller scale system. The price responsiveness model is not applicable to the district energy system where no market structure exists. Second, having DR for capacity options, reliability metrics are required to evaluate system adequacy. Calculating

reliability metrics for system adequacy requires stochastic simulations, and this cannot be incorporated into the given mathematical program formulation.

Due to the above-mentioned challenges, instead of formulating a single optimization problem to include all the requirements, the step-by-step methodology will be formulated to optimize capital investment in capacity expansion plan and evaluate different contingency plants.

2.5 Conclusion

In this chapter, the concept of a district energy system is introduced as well as its characteristics and the limitations of the current sizing and capacity planning process. Analogies to a power grid system and its own challenges are investigated. Demand response concepts and the state of the art of DR modeling approaches for power grid system are reviewed from available literature. Due to lack of data, most of the existing models are at the conceptual planning level. The recent advancement of sensing and measurement, and communications technologies initiated DR programs and data is becoming available. There is a handful of research conducted to use DR data to evaluate and predict its benefit, but not many. Reliability metrics for power system generation adequacy test are reviewed. The same metrics can be used to evaluate district energy system reliability because they evaluate the balance between supply and demand. Finally, capacity expansion planning optimization methods are examined for their suitability to establish a decision making environment to yield a consistent plan without a human decision maker. The next chapter will synthesize information reviewed in this chapter to formulate a research problem and develop a methodology.

CHAPTER 3 PROBLEM FORMULATION

In the previous chapter, the need for research was addressed through the literature review, and the technical challenges were identified. This chapter will formulate the research problem by asking research questions and proposing solutions as forms of hypotheses to scope down to the several research topics and identify the solution approach.

The motivation of this thesis originated from the observation of the measured data of a campus energy system. Findings from the literature review supported the observation that the oversized chiller plant is a common problem. The main cause of the problem is the inherent uncertainty in demand prediction, but it is not well characterized. The nature of sizing makes it worse. The total capacity is designed to meet the peak load, which only lasts a few hours per year, and a redundant capacity is added to ensure the system to operate even at a contingency event. As a result, the plant runs at part load condition with a huge margin for the rest of the time. Demand response can alleviate this problem by allowing the system to have less safety margin if it is used during the contingency events. In order for DR to be considered as a realistic option, its equivalent capacity compared to a chiller and contribution to the system reliability should be quantitatively evaluated. Although DR models are developed for the grid system planning or the building HVAC systems, they are not directly applicable to a district energy system because of the different characteristics as explained in Section 2.1. Therefore, suitable modeling approaches and evaluation methods are investigated, and they will be the foundation of the proposed methodology.

3.1 Demand Response Modeling for District Energy System

As introduced in the previous chapter, the motivation of DR is mainly to reduce peak load and improve system reliability from the demand side. Most of the studies about DR focused on control methods of individual buildings or very high-level planning with simple models. A district energy system employed in a community such as Georgia Tech has both sides of the perspectives. As presented in Figure 7, a chiller plant consumes electricity to produce thermal energy. The required cooling energy for campus buildings is all generated at the plant so that the electrical energy savings by DR from the building HVAC systems of the individual buildings cannot represent the benefit of DR for the entire system. Currently, a typical campus having a district energy system participates in DR programs only as a reactive user to the price signal with the expectation of small monetary benefit or avoidance of penalty at the peak times. The Benefit at the district plant level with the aggregated cooling load reduction is not well appreciated.

Recall the primary research question presented in Section 2.2.

Primary Research Question:

What is the appropriate approach to analyze DR resources and evaluate the impact on the capacity planning of a district energy system?

Where data is available, a data-driven approach is suitable to avoid extensive modeling of building HVAC systems. Assume that a small number of campus buildings have participated in a pilot DR program. The goal of study in this section is to investigate an efficient way of utilizing historical building energy data with DR event logs to convert raw data into actionable information and widen the perspective of DR from a passive user

to more active manager to exercise DR at its own need. The two perspectives expect different benefits: Energy cost savings and the improvement of plant operation and management. Anticipated benefits of DR are in two perspectives are listed below:

- Participating DR program will reduce the peak cooling load of buildings for a community
- Reduced load by DR can impact on a chiller plant operational performance and reliability by changing the required capacity of the plant

Having data including DR events, these statements can be proved statistically. However, it cannot be done with raw data. Appropriate methods to utilize data should be investigated. This section will formulate a research problem to address the first perspective, and the next section will expand it to the second perspective. The first research question is formulated as below.

Research Question 1:

How can the benefits of DR for the entire campus buildings be assessed with a small number of sample buildings?

To answer this question, available data should be thoroughly examined.

3.1.1 Preliminary Data Analytics for Model Development

Three buildings at Georgia Tech campus are identified as pilot buildings of DR program in which the global HVAC system's set point temperature is automatically adjusted by the fixed degree of temperature to reduce load when the price signal rises over the specified threshold. In 2016 summer, from June to August, the eighteen DR events have occurred. Activated and deactivated times of DR events are stored in the building

automation system database. For each building and chilled water plants, historical chilled water meter and electricity meter data at the building level are available for 15 minute interval up to past five years, but the floor or room level data is not stored. The building i's cooling load ($\dot{Q}_i$) is measured by the temperature difference and cooling flow rate as in Equation (18) and the sensors are located at the tertiary level, which is the interface between the building and the network as presented in Figure 26. Weather parameters such as temperature and relative humidity are also available from the on-site weather station with the same resolution of the time interval.

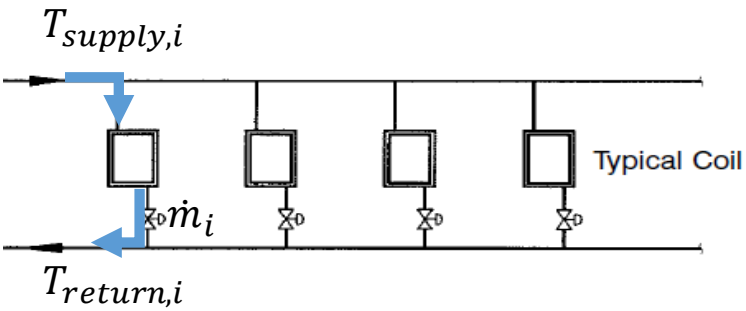


Figure 26 Building Cooling Load Measurement Point

$$\dot{Q}_i = C_p \dot{m}_i (T_{return,i} - T_{supply,i}) \quad (18)$$

Basic information about the buildings with DR events is listed in Table 2. Three buildings are assigned to different departments for the mixed-use such as classrooms, offices, and research labs like most of other university campus buildings. On the DR days, the buildings have much higher peak load than the non-DR days. This makes sense because the DR events typically occur when grid system is in danger of outages or requires the expensive generation to meet the peak load. On the DR days (the red line), the average and high temperatures were higher than those of the non-DR days (the blue line) as presented in Figure 27.

Table 2 Sample Buildings with DR Events

Name	Types	Floor Area (Gross sqft)	Average Peak Load (Tons) DR Days	Average Peak Load (Tons) Non-DR Days
Building A	Mixed Use	139,914	141	109
Building B	Mixed Use	136,092	268	200
Building C	Mixed Use	417,576	432	332

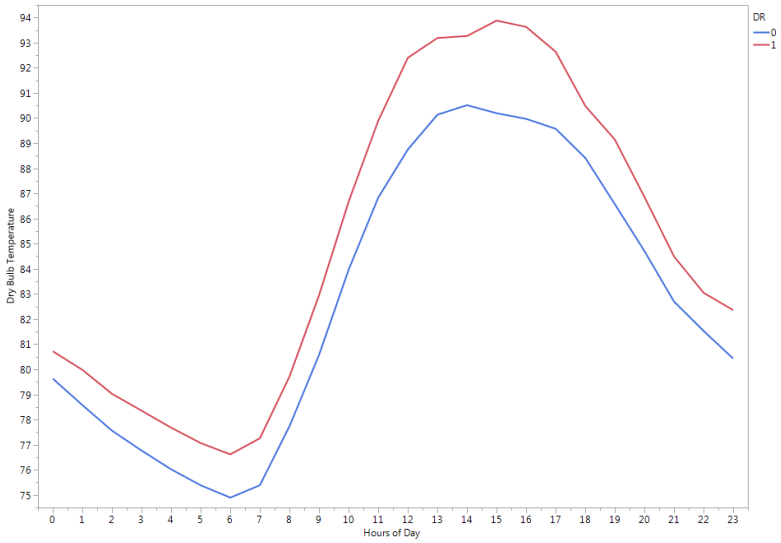


Figure 27 Hourly Temperature Profile on DR and Non-DR Days

Sample data for three days with a DR event from one of the pilot buildings is presented in Figure 28. This dataset indicates the cooling load was reduced by DR. Red dots creates a valley while the normal days did not have such valleys in the load curve. Unfortunately, the amount of load shed cannot be measured because the baseline load shape that would have been without DR is unknown. Therefore, an effective way to

estimate the original load to calculate the amount of DR load shed is needed. This leads to the Research Question 1.1.

Research Question 1.1:

How can we create a credible enough model to quantify the cooling load curtailment by DR where small sets of data for DR events exist?

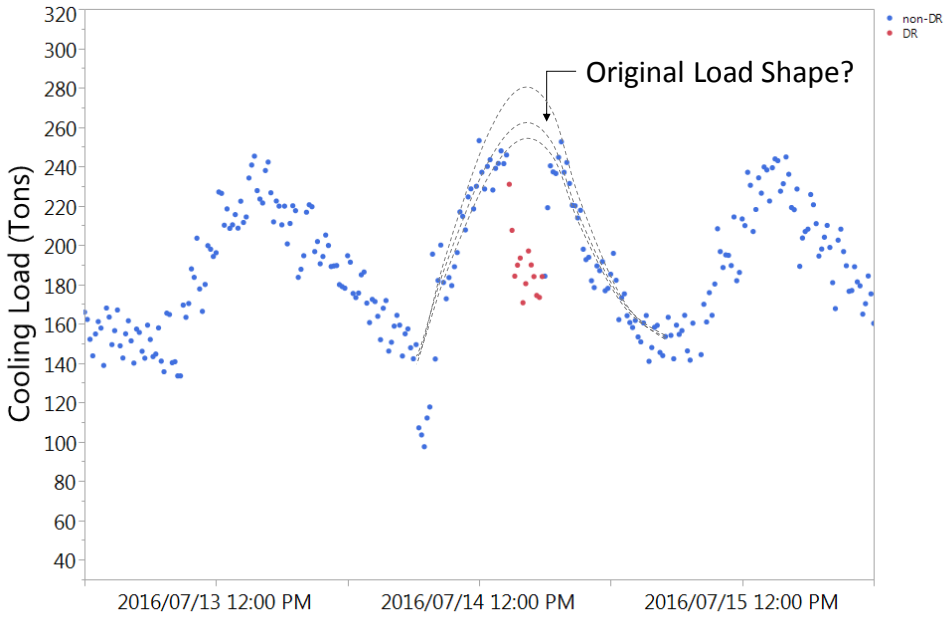


Figure 28 A Set of Sample Data with a DR event

Having data, a data-driven building modeling approach is a reasonable choice to develop a load prediction model. The motivation to develop a building energy model is to identify the original load shape as depicted as dashed lines in Figure 28 and calculate the difference between the original load and DR load to estimate DR benefits.

The cooling load is mainly affected by weather conditions and the occupancy schedule. Buildings have their own unique schedule as illustrated in Figure 29. It shows that three buildings have different usage patterns throughout a week from Monday to

Sunday. Building A and C seem to have an automated HVAC control schedule to reduce night time and weekend energy consumption and relatively flattened pattern during the daytime. This might cause the early peak time before noon even in the early morning when the schedule changes. On the other hand, Building B is following the outside temperature pattern. The model should be able to capture these effects of weather and schedule.

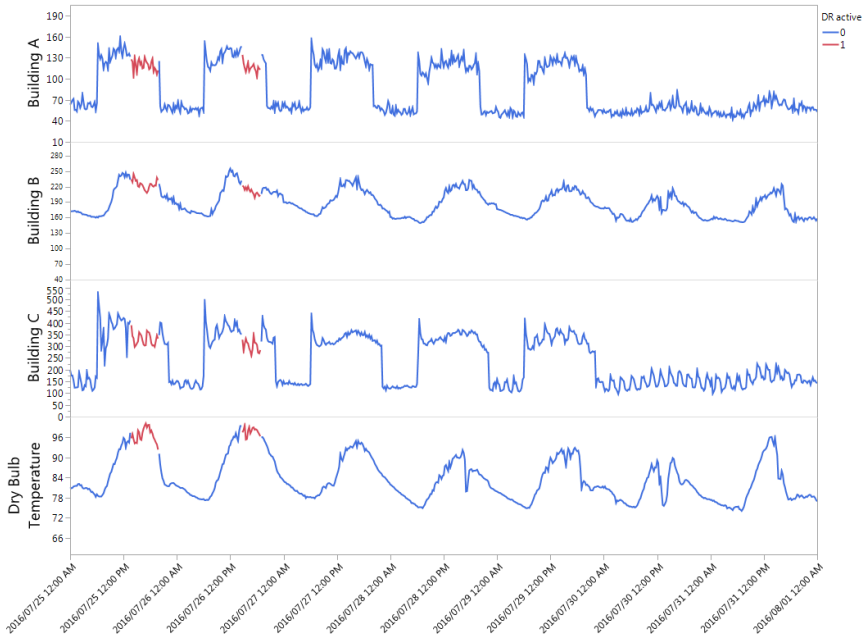


Figure 29 Load Patterns of Sample Buildings

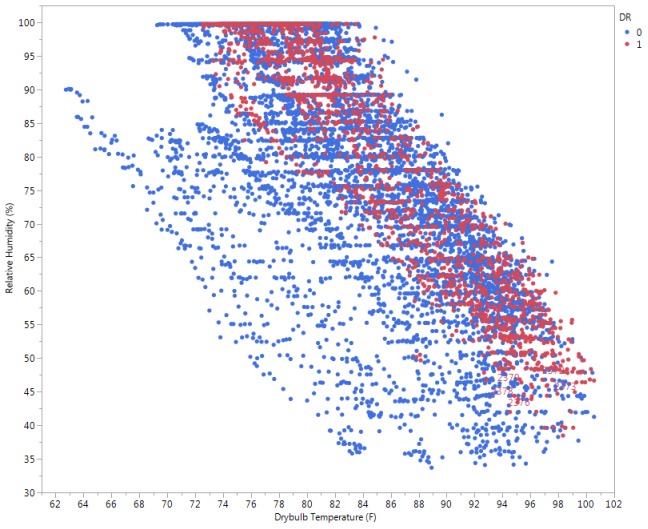


Figure 30 Temperature vs. Relative Humidity in Summer

Dry bulb temperatures and relative humidity of 2016 summer are plotted in Figure 30. Even though DR days have higher average temperatures and the narrow area in the plot, similar environmental conditions can be found from non-DR data. Using this non-DR data set, the original load shape can be predicted. From this observation, Hypothesis 1.1 is formulated.

Hypothesis 1.1:

If a parametric model is built with non-DR data from similar environmental conditions as a function of weather and time parameters, then the fitted model will predict the original load during the DR active period. DR load shed can be discerned by the difference between the model predicted and measured load.

Once a predicted model is built using non-DR data set, the amount of load shedding for each DR event can be calculated using Equation (19) by the difference between the measured value and the model predicted value. It is normalized by the model predicted value because the model represents the original load when DR would not have occurred so

that the normalized value can be used to compare its effectiveness for the different buildings.

$$Normalized\ DR\ Load\ Shed:$$
$$\Delta Load_{DR} = \frac{y_{model} - y_{measurement}}{y_{model}} \quad (19)$$

3.1.2 Generalization of DR Model

There are from dozens to hundreds of buildings in a district energy system. Those buildings are all distinguishable, and cannot be represented in a uniform model like an airplane or car. As pointed out in Section 2.1, it is not feasible to build high fidelity building energy simulation models for all the buildings with limited time and resources. The inverse modeling approach can reduce the modeling burden if measured data is available. It is a kind of the surrogate modeling approach of a real system, and the same formulation of the equation can be used for other buildings. As in the load prediction model, if the amount of load shed can be modeled like an inverse model, the observation from the sample buildings can be generalized for the same DR method – the global setpoint temperature adjustment of the HVAC system. This thought leads to Research Question 1.2.

Research Question 1.2:

Is it possible to build a generalized DR model to predict the DR effectiveness of other buildings with sample data?

From Equation (19), the modified cooling load by DR can be derived as in Equation (20).

$$\dot{Q}_{DR} = (1 - \Delta Load_{DR})\dot{Q}_0 \quad (20)$$

The relative amount of load shed by DR ($\Delta Load_{DR}$) may depend on the building type, the change of set-point temperature, and other factors such as occupancy's activities or environmental variables while it is independent of the size of the original load because it is normalized. If the same DR scheme is applied to all the buildings, the factors are reduced to external variables such as building schedule, time of the day, or weather parameters. Since the load model is a function of time and weather variable, the normalized DR load shed $\Delta Load_{DR}$ can be approximated as a function of the external variables if there are any patterns. If $\Delta Load_{DR}$ can be modeled as a parametric function, then it allows to estimate the reduced load by applying DR using Equation (21).

$$\dot{Q}_{DR} = (1 - \Delta Load_{DR}(t, w))\dot{Q}_0(t, w) \quad (21)$$

Where,

$\dot{Q}_{DR}$: Reduced cooling load by DR

$\Delta Load_{DR}$: Demand Response Effectiveness model

t: time variable

w: weather variable

$\dot{Q}_0(t, w)$: the original load model

There were a few attempts to characterize DR load shed for HVAC system as the described way. Dyson et al. used smart meter data to estimate demand response potential of residential buildings by changing A/C set point temperature [12]. They used the slope of the outside temperature vs. power consumption plot during cooling days. They did not have data for the actual DR event, but the assumption behind their method is that the change in indoor temperature by A/C and the change of outside temperature will have the same effect on load reduction. If the setpoint temperature change is fixed, then the load shed by DR becomes a constant value for a given building regardless of the outside temperature. Another research group claimed that the amount of load shed for A/C is correlated with

outside temperatures and time of day [49]. However, even their data shows contradicting results as presented in Figure 31, which represents the measured vs. the predicted DR potential calculated as a percentage of the original load plotted by the outside temperature and time. The measured data is randomly scattered and shows the opposing trend to the predicted model. The reason for the contradiction may be the poor fitness of models, uncertainties that are not considered in modeling, or just because their modeling assumption is wrong.

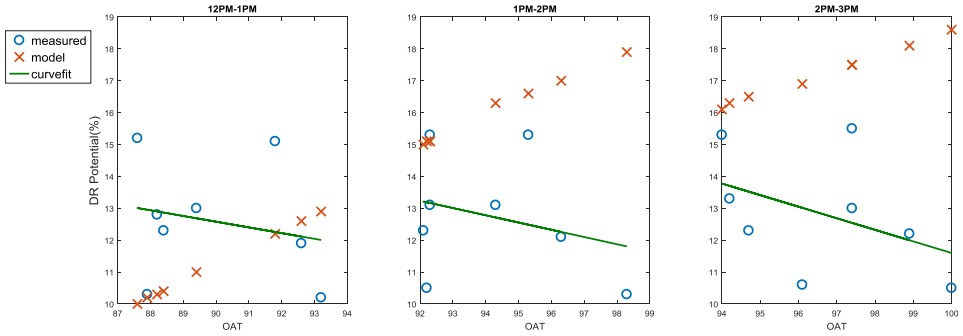


Figure 31 Measured vs. Model Estimated DR Potential (%)

Whether the DR load shed is constant or dependent on time and weather, if uncertainty is too large, then the uncertainty will conceal the actual trend. Both of the approaches mentioned above did not consider the uncertainty in the DR load shed. However, this uncertainty is not reducible by modeling because of the nature of randomness of demand. This observation leads to Hypothesis 1.2.

Hypothesis 1.2:

No apparent correlation will be observed regardless of the true relationship due to high uncertainty.

If a strong correlation is observed, DR can be modeled as a parametric function of temperature and time. If Hypothesis 1.2 is valid, then DR can be modeled as an additive random variable to modify the original load as in the price elasticity model as below. A probability distribution can be inferred from data.

$$\dot{Q}_{DR} = (1 - \Delta Load_{DR})\dot{Q}_0(t, w) \tag{22}$$

$$\Delta Load_{DR} \sim X$$

X: the continuous random variable with

a probability distribution function $f(x)$

By validating two sub-hypotheses, Research Question 1 can be answered by developing the load model and the DR model.

3.2 Quantification of DR Impact on Chiller Plant

Now that the research problem is formulated from the first perspective of DR, the DR benefit for the second perspective will be investigated. The second perspective is for a system operator as a thermal energy producer. The focus of this study is to examine if DR can change the operational pattern of the chiller plant. The expected benefits for the chiller plant are electricity cost saving and the improvement of availability if DR can reduce the operating number of chillers. Therefore, the required metrics to evaluate the impact of DR on the chiller plant is the operating cost and chiller staging status. Both metrics are not independent because the electricity consumption of a chiller plant is closely related to the operating methods as pointed out in Section 2.1.1.

The aggregated load will be reduced if the entire campus buildings participate in DR at the same time. However, it is not certain that it will impact the plant's operational state. If the number of the available chillers in standby is increasing, then it can be

qualitatively said that reliability is improved. Does DR change the status of chiller staging? Answering this question is not straightforward. A conceptual chiller staging logic is illustrated in Figure 32. As the cooling load is continuously increasing, the required number of chillers is increasing in a stepwise manner. An operator (or an automated chiller dispatching program) switches on or off another chiller if cooling load is detected in the switching zone. The switching zone differs by a specific logic, perception, or sensing and it affects the efficiency of a chiller plant. Depending on where the original load was located on this staging curve and the amount of load shedding, DR can trigger to switch off a chiller or not.

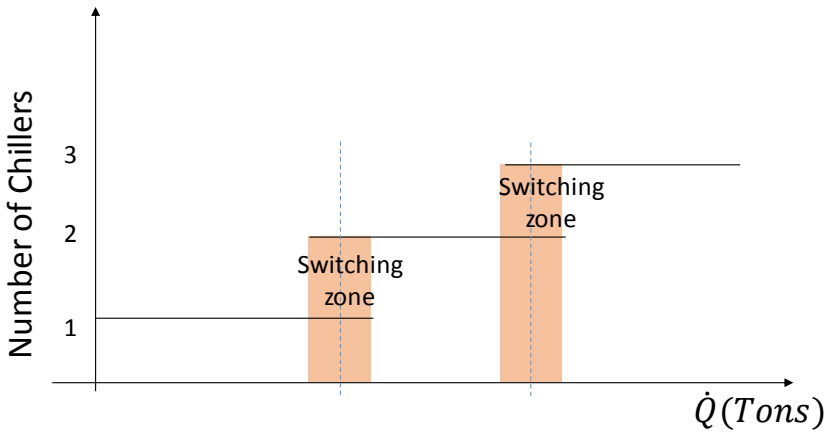


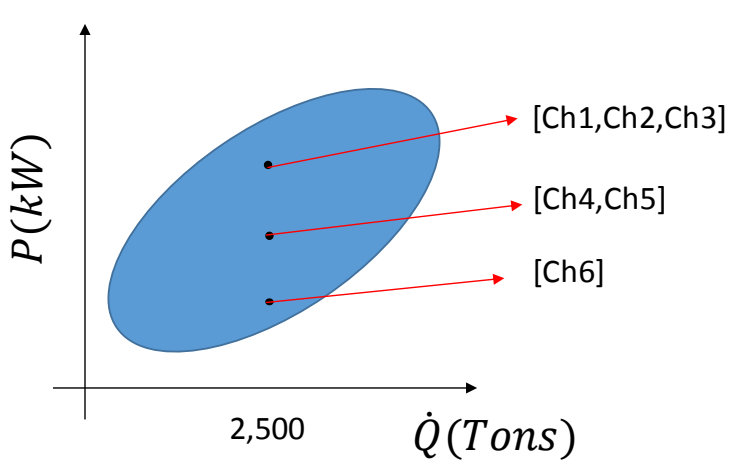
Figure 32 Chiller Staging Control

Due to the dynamic nature of the system, a simulation model of plant with the capability of automated staging that responds to varying load is required to calculate the operating cost and examine if DR improves the availability. This requires developing a physics-based dynamic model of the chiller plant and designing a control system. However, designing an optimal control logic is very complex and time consuming work because it is a discrete combinatorial problem and beyond the scope of this thesis. Therefore, the Research Question 2 is formulated to address this challenge.

Research Question 2:

How can be the benefits of DR for a chiller plant be assessed without implementing a complex control logic?

To answer this research question, a hypothesized situation is given. Where the multiple chillers are used to meet varying demand, there can be many different possible combinations of chiller settings for one load condition. A notional system is illustrated in Figure 33. This plant has six chillers of three different sizes. To serve 2,500 tons of cooling load, the three different combinations of chillers presented in Figure 33 and more options can be possible. If all the combinations that meet the load requirements are plotted, then a design space for this chiller plant control can be generated as presented in the blue shade.



Chiller 1	1,000
Chiller 2	1,000
Chiller 3	2,000
Chiller 4	2,000
Chiller 5	2,000
Chiller 6	3,000

Figure 33 Chiller Plant Control Design Space

Assuming a chiller plant can be operated at the most efficient combination of chiller settings, the optimal chiller combinations will be located at the lowest edge of the design space. If the optimal chiller combinations are found dominating other combinations for all load conditions, then chiller staging can be emulated using the optimal combination of the

settings for varying load conditions without a sophisticated control logic. This observation leads to the Hypothesis 2 as an answer to the Research Question 2.

Hypothesis 2:

If the optimal control design can be approximated as a set of the best combinations per loading conditions to minimize power consumptions, then it can predict changes of operational status of a chiller plant without a sophisticated control logic.

To prove this statement, a chiller plant model using campus data should be developed. Once the model is validated, then the optimal setting combinations can be generated by simulation, and it enables to test if DR can change the status of the chiller plant staging status.

3.3 Reliability Analysis

The primary goal of this thesis is to evaluate DR as the replacement of the redundant chiller. A contingency plan should be well documented to run system smoothly when the event happens to prepare an emergency such as equipment failure. Keeping a backup unit is one way of the contingency plan. The redundancy alleviates the risk by ensuring the system operate properly at the peak load condition even when the largest chiller is unavailable. DR can do the same function by reducing the peak load. However, the DR impact cannot be formulated as an equivalent amount of redundancy. Detecting changes in the required number of the chiller and counting the number of chillers in standby is an only proxy way to evaluate system reliability. To persuade a decision maker to consider DR as a contingency plan, a well-defined and quantitative evaluation metric should be provided. By quantifying its potential capacity and evaluating its impact on the system reliability, DR can be used as a contingency plan. To this end, the Research Question 3 is formulated.

Research Question 3:

How can the performance of DR as a contingency plan for a district chilled water system be measured and fairly compared to the deterministic redundancy?

One may be tempted to calculate the potential capacity of DR using the DR model proposed in Section 3.1. It is, however, not a suitable metric to evaluate and compare DR and a redundant chiller. The capacity of the redundant chiller only tells how much margin the system has beyond the required capacity and cannot represent the dynamic characteristics of the system. Instead of formulating a testable Hypothesis, the traditional and probabilistic reliability metrics will be explored to determine the proper metrics.

3.3.1 Reliability Evaluation Metrics

The purpose of a contingency plan is to mitigate the risk of failure. The risk comes from uncertainty. To evaluate the performance of a contingency plan considering the uncertainty, a probabilistic metric is needed. The N+1 redundancy requirement can be interpreted into a mathematical expression as in Equation (23). X is the status vector of the chillers indicating the availability. It means that the total capacity should be larger than the peak load plus the largest capacity of chillers to satisfy the N+1 redundancy requirement.

$$\dot{Q}_{peak} < \sum_{i=1}^{N} Cap_i\, x_i \; and \; X = [1, \dots, 1, 0] \quad \text{Or } \dot{Q}_{peak} + \max_{i \in n} Cap_i < \sum_{i=1}^{N} Cap_i \tag{23}$$

Where,

$\dot{Q}_{peak}$: the peak load

Cap_i : the capacity of chiller i

$X = [x_1, \cdots, x_N]$: the status vector of chillers

x_i: the binary random variable, 1: available, 0: unavailable

N: the total number of chillers

$n = \{1, \cdots, N\}$: the set of chiller index (in order of size)

In reality, the peak load and the chiller availability status are stochastic variables. The safety factor and the redundancy area used because the distributions of those variables are unknown. Considering this, the N+1 requirement can be converted into a probabilistic measure as a conditional probability of the system failure as express in Equation (24).

$$P(fail\ to\ meet\ peak\ load \mid the\ largeset\ chiller\ is\ unavailable\ and\ others\ work) = P\left(\dot{Q}_{peak} > \sum_{i=1}^{N} Cap_i\, x_i \;\middle|\; X = [1, \cdots, 1, 0]\right) \quad (24)$$

This expression reveals the flaws of the N+1 redundancy requirement. It assumes that when the largest chiller fails, all other chillers are available. This cannot assess the risk properly because it only represents one special case. Therefore, it should be fixed to include all other failure cases to estimate the probability of failure at the peak when the largest chiller is broken as in Equation (25).

$$P(fail\ to\ meet\ peak\ load \mid the\ largeset\ chiller\ is\ unavailable)$$
$$P(\dot{Q}_{peak} > \sum_{i=1}^{N} Cap_i\, x_i \mid x_i = 0) = P(\dot{Q}_{peak} > \sum_{i=1}^{N-1} Cap_i\, x_i) \quad (25)$$

Another flaw of the redundancy requirement is that it exacerbates the risk. A system may be able to meet the peak load even the largest chiller is unavailable. In more general form, the probability that a system cannot meet the peak load can be expressed as in Equation (26).

$$P(fail\ to\ meet\ peak\ load) = P(\dot{Q}_{peak} > \sum_{i=1}^{N} Cap_i\, x_i) \tag{26}$$

It is obvious that Equation (25) is always larger than Equation (26) if the random variable x_i is independent and identically distributed. This probability function can be used to evaluate either a redundancy or DR as a contingency plan of a system at the peak load condition.

By adding a contingency plan to the original system, which is designed to meet the peak load with N number of chillers, and calculating the change in the probability of failure, one can fairly measure and compare the performance of either a redundancy or DR. As the mathematical expressions, Equation (27) is devised for a redundancy and Equation (28) is for DR. Therefore, the difference of failure probability can be used as a risk mitigation capability metric of contingency plans.

Given a system where $\mathbb{E}[\dot{Q}_{peak}] < \sum_{i=n}^{N} Cap_i$

$$\Delta P = P\left(\dot{Q}_{peak} > \sum_{i=n}^{N} Cap_i\, x_i\right) - P\left(\dot{Q}_{peak} > \sum_{i=n}^{N+1} Cap_i\, x_i\right) \tag{27}$$

$$\Delta P = P\left(\dot{Q}_{peak} > \sum_{i=n}^{N} Cap_i\, x_i\right) - P\left(\dot{Q}_{DRpeak} > \sum_{i=n}^{N} Cap_i\, x_i\right) \tag{28}$$

As pointed out in the literature review section, the dynamic reliability metrics of power systems can be adapted to evaluate a district energy system to compensate for the shortcomings of measuring only the peak load performance. Those metrics are rewritten for a district cooling system.

The loss of load Probability (LOLP) is the probability that the system cannot serve the load over the unidimensional period. If the time interval is one hour, LOLP is expressed as the loss of load expectation (LOLE) which means the expected duration of the unserved load per year. The loss of energy expectation is the average amount of unserved energy per year.

$$LOLP(t) = P\left(\dot{Q}_t > \sum_{i=n}^{N} Cap_i\, x_{it}\right) \tag{29}$$

$$LOLE = \int_{t=0}^{T} LOLP(t)\, dt = \sum_{t=0}^{8760} LOLP(t) \tag{30}$$

$$LOEE = \sum_{t=0}^{8760} \left(\dot{Q}_t - \sum_{i=n}^{N} Cap_i\, x_{it}\right) LOLP(t) \tag{31}$$

ELCC is the horizontal distance between two systems' LOLE vs. load curves as presented in Figure 21 to measure the capacity value of added generation to maintain the system reliability. However, the district energy system society does not have the standard for this kind of measure. Thus, the vertical distance at the original peak load estimation point. This is the same idea with the risk mitigation metric for the entire year.

In conclusion, a contingency plan should be evaluated by how well it will recover the original functionality when a system fails. Therefore, the differences of the proposed metrics before and after of the system modification can be used to evaluate and compare DR and the redundancy as the contingency plans.

3.3.2 Definition of Failure

District chilled water systems are rarely exposed to failures such as an outage. Typical reliability values of the mature systems are higher than 99.98% [26]. The definition of failure should be specified to calculate the reliability metrics. Failure means that a system cannot perform the specified function at the desired level. Improper staging of chillers, other equipment failures such as pumps or cooling tower may result in malfunction of cooling load distribution even if the available capacity at the plant is enough to satisfy demand. These failures are not considered in this problem. Assuming the chiller plant is optimally operated all the time without the auxiliary equipment failure, the shortfall of production of cooling is defined as a system failure. If a chilled water plant cannot produce thermal energy enough to supply demand at any time due to the chiller failure for some reason, it will be said the system fails.

Definition:

If maximum cooling capacity is less than the cooling load, then the system fails.

To satisfy demand, two physical criteria should be met: the energy conservation and the mass conservation. As explained in the previous chapter, the low delta T syndrome may fail the system even when the thermal capacity is enough. Notional examples are given to show how the low delta T causes the degradation of the system performance by asking more cooling flow than the actually required.

The notional system consists of the same sized chillers of 3,000 Tons. The design delta T is 12°F, and the chiller's mass capacity is 6,000 GPM (Gallon Per Minute). Water's specific heat is 1BTU/(lbm°F). As a refrigeration ton is approximately equivalent to 12,000 BTU/hr, the required cooling flow rate is calculated as below.

$$\dot{m}_{requied} = \frac{\dot{Q}_{required}}{C_p \Delta T} = \frac{24 \times \dot{Q}_{required}}{\Delta T} \ (GPM) \qquad (32)$$

As illustrated in Figure 34, the required cooling flow rate is increasing as delta T is decreasing for the same cooing demand. A system operator has to bring more chillers into online than the actually required thermal capacity as presented in Figure 35. When delta T is lowered to below 6°F, this system needs more than the double number of chillers compared to the system with the design delta T only because of the mass flow rate requirement.

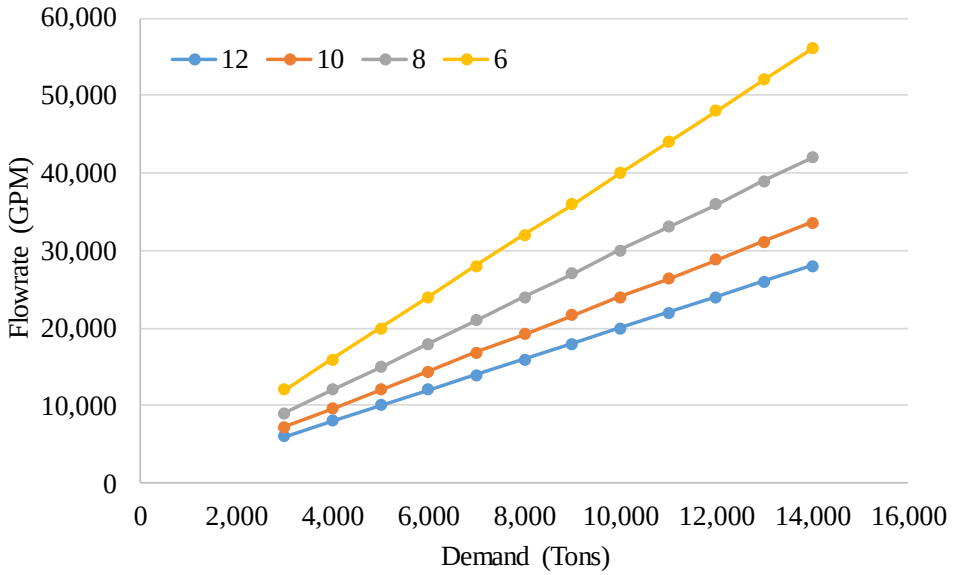


Figure 34 Required Cooling Flow Rate by Delta T – Notional System

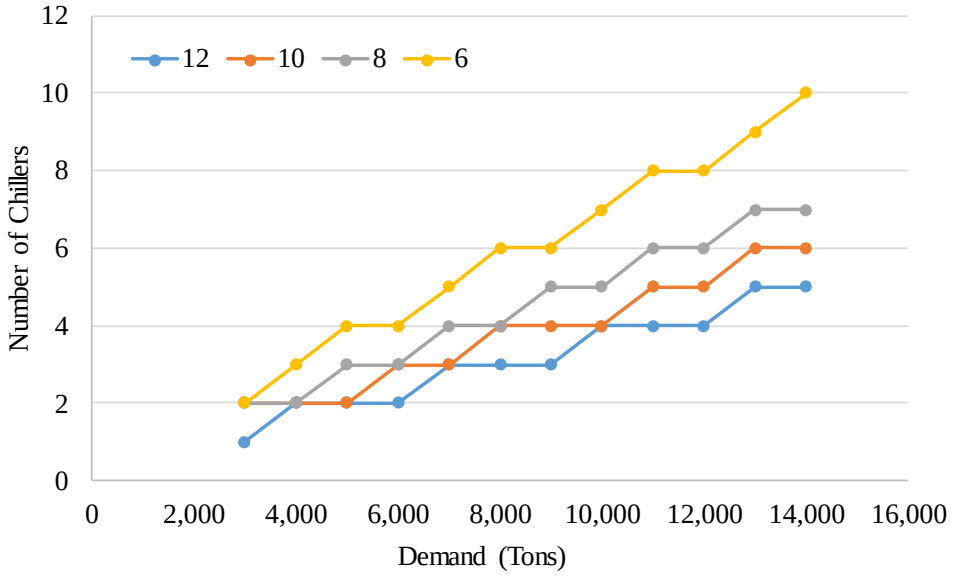


Figure 35 Required Number of Chillers by Delta T – Notional System

A real system typically can achieve the design delta T at the high load conditions while delta T is degraded at low load conditions. Data collected from the Georgia Tech chilled water system is presented in Figure 36. Dispersion of data is due to many factors such as sub-optimal operations of the plant, environmental variation, or the demand side performance. The trend of delta T according to the loading conditions can be approximated using the second order polynomial equation as depicted as the red line.

$$\Delta T = c_0 0 + c_1 \dot{Q}_{normed} + c_2 (\dot{Q}_{normed} - c_3)^2 \tag{33}$$

The required cooling flow rate and the number of chillers can be calculated with the approximation model of delta T in Equation (33). The actual system's performance would be in between the two extreme cases of a notional system as presented in Figure 37 and Figure 38.

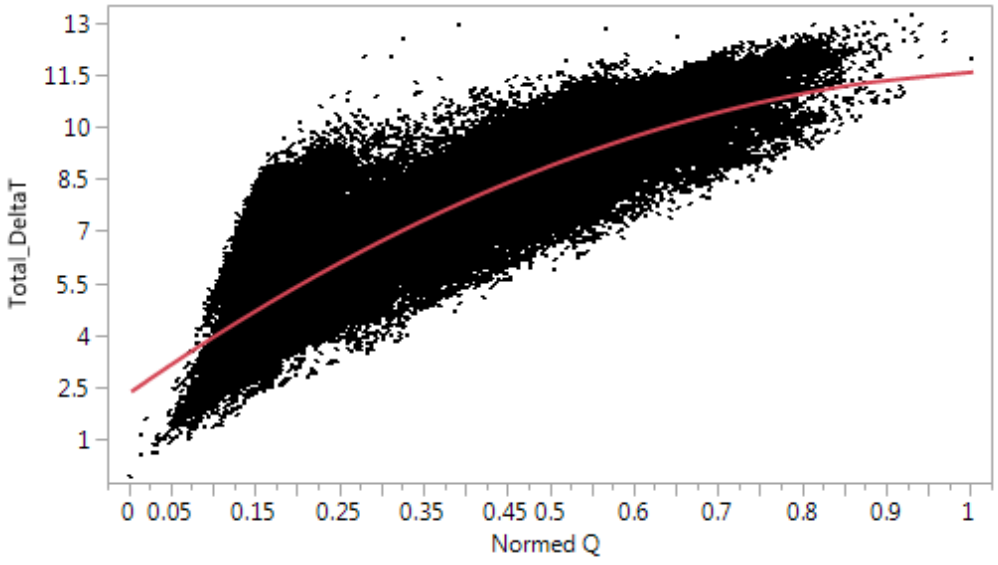


Figure 36 Delta T vs. Load Condition

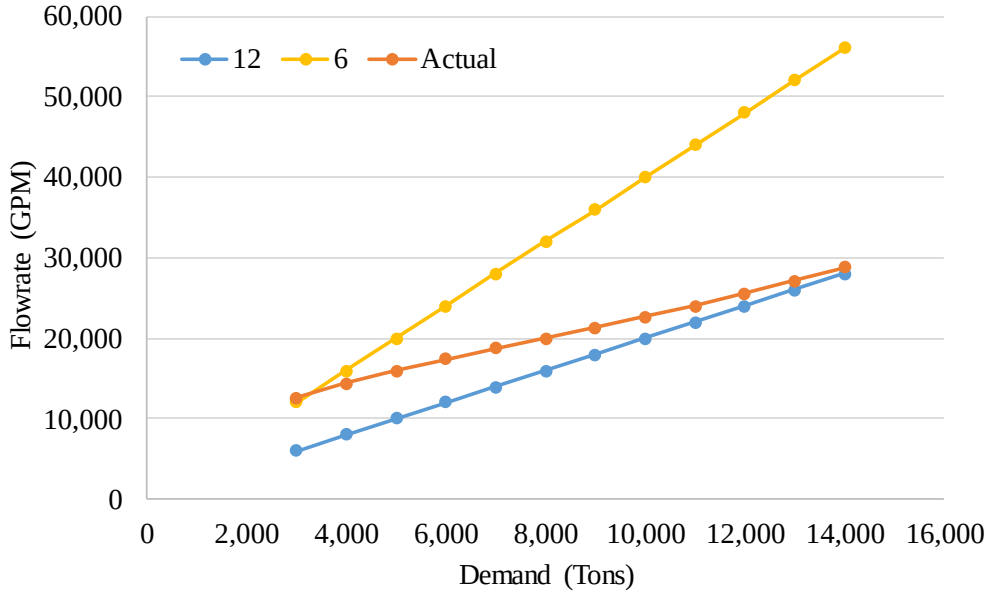


Figure 37 Required Cooling Flow Rate by Delta T – Real System

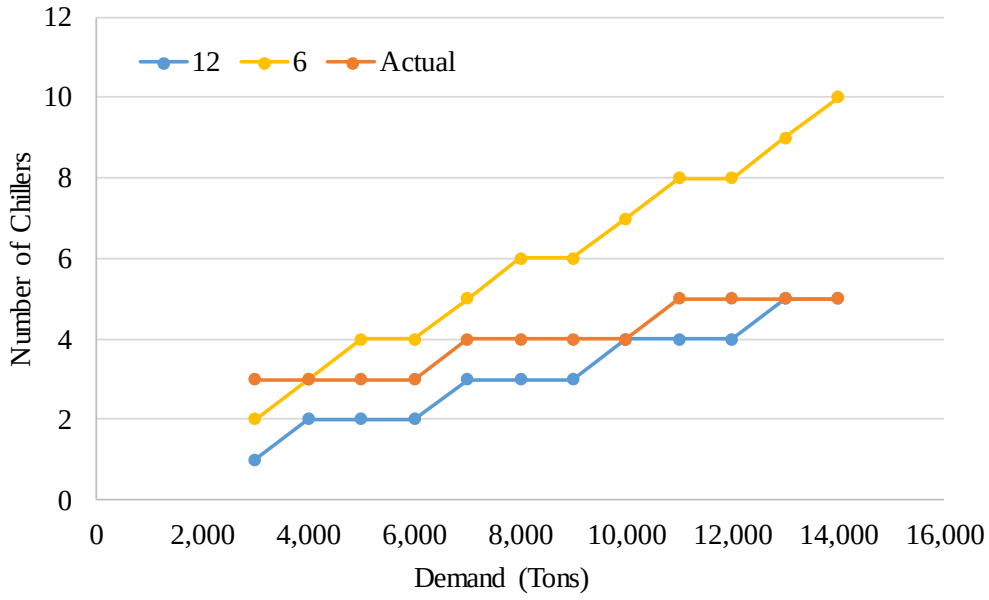


Figure 38 Required Number of Chillers by Delta T – Real System

Looking at this trend, one can notice that the system always needs more chillers than the number of chillers by required cooling load except the full load condition. It can be interpreted into that the system failure is determined by the cooling flow rate balance rather than the cooling load balance at part load condition. Therefore, the system failure should be determined by cooling load discrepancy AND mass flow rate discrepancy. However, at part load condition when delta T is lower than the design value, the probability of failure due to mass balance will dominate thermal load balance as expressed in Equation (34).

$$\begin{aligned} &P(\dot{Q}_{produced} < \dot{Q}_{required} \cap \dot{m}_{produced} < \dot{m}_{required}) \\ &= P(\dot{m}_{produced} < \dot{m}_{required}) \end{aligned} \tag{34}$$

3.4 Capacity Expansion Planning Problem Formulation

The previous sections investigated the modeling and evaluation approaches for the existing system. This section will address how to forecast the capability of DR to defer or avoid the future capacity expansion plan. The last research question is formulated to evaluate the idea of using DR as a contingency plan in replacement of a redundant chiller.

Research Question 4:

How can we predict if DR is capable of avoiding or deferring capacity expansion without sacrificing reliability?

This research question motivates the formulation of the optimal capacity expansion planning problem. To test if DR can avoid or defer adding capacity, a baseline plan should be given. A consistent and rigorous planning method is required to compare how different

strategies affect the plan. The next section will build up the optimal capacity expansion planning problem formulation to answer this research question.

3.4.1 Planning Schedule Optimization

Capacity expansion plan of an existing plant is to determine a schedule of purchasing and installing new chillers to meet the increasing demand. Decisions are made for when and how much size of chiller should be purchased. It is driven by the demand growth forecast. Demand Response can affect a long-term capacity plan by reducing the peak load or releasing the constraint on the minimum capacity by removing the redundancy requirement so the capital investment cost can be avoided or deferred [7]. The purpose of formulating the optimal capacity expansion planning problem is to establish a rigorous decision-making process without a judgment by a human decision maker. Therefore, it should yield a consistent result under the same condition.

Chiller options are selected from a limited number of commercial products, and decisions are made at the discrete time interval. Therefore, decision variables should be formulated as the integer or binary variables to account for this characteristic of the problem. The MILP (Mixed Integer Linear Programming) is the most commonly used mathematical programming formulation for capacity planning for industrial manufacturing facilities and can be solved with various commercial solvers [68]. In Section 2.4, the challenges of formulating a mathematical programming problem were addressed. To resolve those challenges, the capacity expansion planning problem will be formulated as Mixed Integer Linear Program as the least cost-capital investment plan under these assumptions.

Assumption 1:

The chiller plant is operated at the most efficient staging scheme for all load condition as explained in Figure 33. This enables the approximation of the operating cost as a function of demand for a fixed plant configuration. If adding new chiller do not differ the approximation curve enough to affect total cost, then the operating cost can be ignored for the capital investment optimization.

This statement is an assumption, not a hypothesis. However, to validate this assumption, the experiment will be performed later in Section 5.4.1.

Assumption 2:

To enable DR for the building HVAC system, no additional cost is required where the building automation systems and communication infrastructure are already established.

With the above assumptions, the objective function can be formulated as a single objective to minimize the total capital investment cost over the planning horizon. The basic MILP formulation can be written by Equation (35)-(39). The total capacity of the plant should be constrained by demand growth projection D_t at every year t. The N+1 redundancy can be easily incorporated into the demand satisfaction constraint as $\max(C_i)$ in Equation (36). This means that the total capacity should be larger than the expected peak load plus the capacity of the largest chiller. If DR is an alternative contingency plan option, this constraint can be relaxed. Equation (37) constrains the number of chillers to be updated every year by the sum of existing chillers and new chillers. Equation (38) is the initial plant condition and Equation (39) restricts the total number of chillers to be installed at the plant.

$$\min \sum_{t=1}^{T} \left(\frac{1}{1+\gamma}\right)^{t} \left[\sum_{i=1}^{N} (IC_i x n_{it})\right] \tag{35}$$

Subject to

$$\mathrm{D}_t + \max(C_i) < \sum_{i}^{N} C_i \times xe_{it} \text{ for } \forall \mathrm{t} \in [1, \mathrm{T}] \tag{36}$$

$$xe_{it} = xn_{it} + xe_{it-1} \text{ for } \forall \mathrm{i} \in [1, \mathrm{N}] \text{and } \forall \mathrm{t} \in [1, \mathrm{T}] \tag{37}$$

$$xe_{i0} = ne_{i0} \tag{38}$$

$$\sum_{i=1}^{N} xe_{it} \leq n_{max} \text{ for } \forall \mathrm{t} \in [1, \mathrm{T}] \tag{39}$$

where,

xn_{it}: binary decision variable for new chiller type i at time t

xe_{it}: discrete variable for the number of existing chiller type i

ne_{i0}: initial number of chiller type i

IC_i: investment cost of chiller type i

C_i: the capacity of chiller type i

D_t: peak load at year t

γ: discount factor

N: number of chiller types

T: planning period

n_{max}: the maximum number of chillers for a plant

In summary, if two conditions are met, the proposed MILP formulation can be used as an optimal capacity expansion planning.

- Operating cost is negligible compared to capital cost.
- No additional cost is required for DR deployment to campus buildings.

3.4.2 Trade-Off Analysis

The MILP formulation is only to optimize the discounted capital cost (NPV) and yield the optimal plant configurations per year. In fact, there is a trade-off in capacity sizing like any other decision-making problem. Cost and reliability are two competing objectives of the sizing problem. Even though DR does not involve an additional capital cost for a campus level district energy system, frequent load shedding is undesirable. If the adjustment of the HVAC set-point temperature lasts too long, then occupants may feel uncomfortable and complain. This can be considered as a cost of DR. Instead of converting this into monetary cost, the reliability metrics suggested in the previous section can be used as another objective function to minimize. The improvement of LOLE from the baseline by DR implies the average time requiring load shed. Stochastic simulation is needed to calculate these metrics. The MILP formulation cannot handle this as an objective function because it requires every element be explicitly defined as linear forms. For the sake of simplicity of the problem formulation, the reliability analysis will follow when the capacity expansion planning is done and the problem will be analyzed in two steps as presented in Figure 39. By adjusting the capacity constraint, the optimal capacity expansion schedule can be generated for the N+1 system and the DR system. For the DR enabled the system, once a baseline plan is generated from the capacity expansion planning optimization, DR is applied to the resulting baseline plan. And the reliability metrics are calculated of the year for the planning horizon.

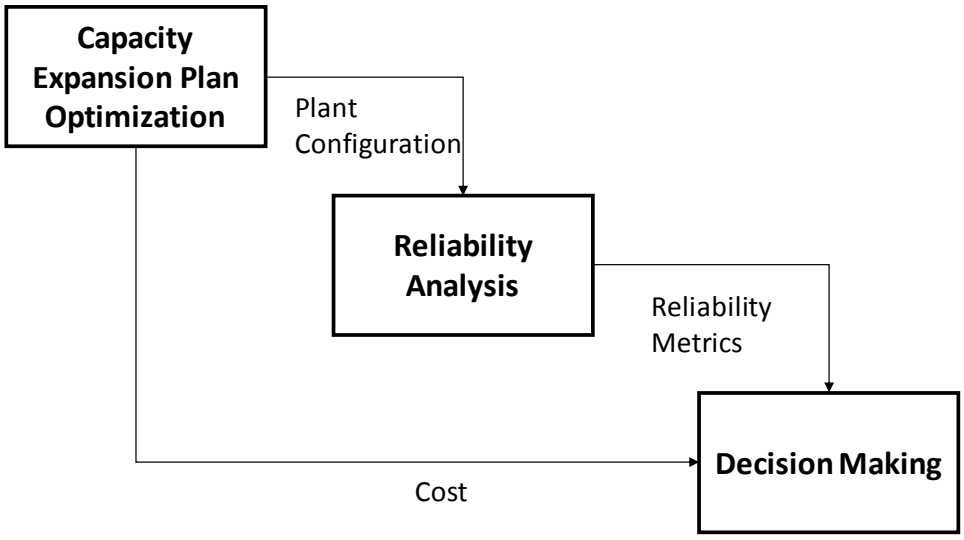


Figure 39 Information Flow between Optimization and Analysis Blocks of the Proposed Decision Making Framework

Since the reliability analysis is performed separately from the capacity expansion scheduling optimization, a decision maker might wonder if the solution is optimal. Regarding this, the last research question is formulated.

Research Question 4.1:

How can we know the optimality of the solution for the capacity expansion plan with the proposed decision making framework?

As a multi-objective decision making problem, there is no single optimal plan. The N+1 system is the most conservative design for the reliability at the peak load condition. At the opposite point, the minimal system without any safety margin will be positioned as illustrated in Figure 40. The minimal plan can satisfy the peak load at normal operation, but it cannot if any failure occurs at the peak time. The optimal plans will not be found as the exact size as presented in Figure 40 because the optimization problem has integer decision variables. These two lines indicate the constraints for the two extreme cases. Small

changes of a constraint can enforce the optimal solution moving to other position for an integer programming problem. By relaxing the constraint, less reliable but lower cost plans can be found between the N+1 and the minimal system at a discrete distance. The capacity constraint Equation (36) can be modified with the safety margin sf as in Equation (40).

$$\mathrm{D}_t + sf < \sum_{i}^{N} C_i \times xe_{it} \text{ for } \forall \mathrm{t} \in [1, \mathrm{T}] \tag{40}$$

Where, $sf: safety\ margin, 0 < \mathrm{sf} < \max(C_i)$

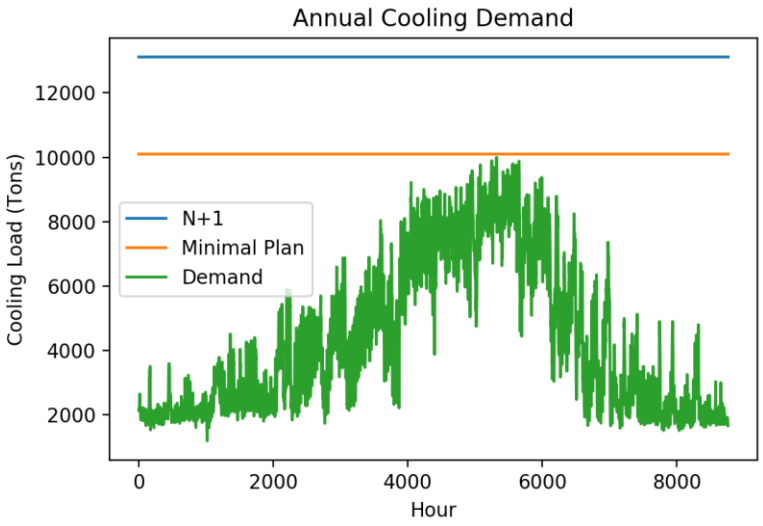


Figure 40 Conceptual Sizing by Margin

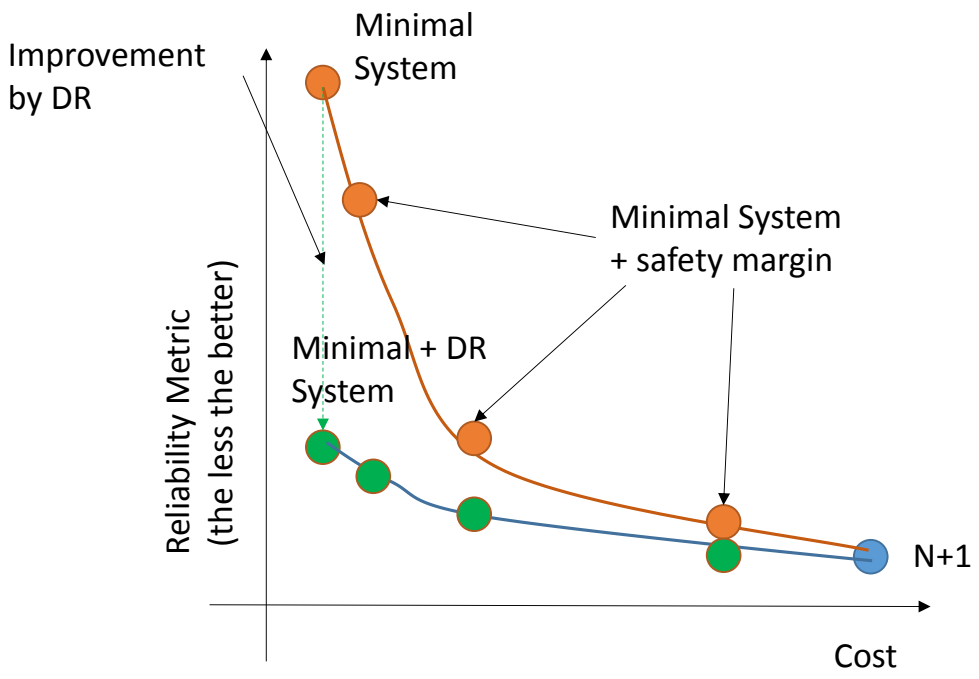


Figure 41 Pareto Frontier and Improvement by DR

With the adjustment of the safety margin on the capacity limit to be between zero and the largest capacity of chiller options, the multiple Pareto optimal plans can be generated depending on the preferences. The concept of the Pareto frontier is presented in Figure 41. The Pareto frontier is the collection of solutions that cannot improve in any direction in the multi-objective decision making problem. The solutions from the MILP capacity expansion planning problem is the Pareto optimal in the cost direction given scenario. Once the reliability metrics are calculated from the simulation for all the design points, it can be presented like a Pareto frontier as in the orange line in Figure 41. From this, DR can improve the reliability metrics to the points on the green line. How much will be improved is determined by the ratio of buildings with DR to the entire campus. If the entire campus is enabled with DR, then the point cannot be improved further for the given system. This thought is summarized as Hypothesis 4.1.

Hypothesis 4.1:

With the two extreme points of the N+1 constraint and no safety margin, the Pareto frontier can be generated by adjusting safety margin on the demand satisfaction constraint between the two points and the fraction of DR buildings.

The formulation of the optimal capacity expansion problem in this section will answer the Research Question 4 as in Hypothesis 4.

Hypothesis 4:

The proposed two-step decision making framework can yield multiple Pareto solutions and the corresponding optimal capacity expansion planning schedule. Using those, one can predict the impact of DR on the capacity expansion planning and the system reliability.

This hypothesis can be tested by formulating a methodology which integrates all the modeling and metrics into a unified decision making framework.

3.5 Summary

In this chapter, the research questions for each topic are addressed, and the potential answers are provided as the hypotheses or the metrics. By answering research questions and test hypothesis, the foundation for the proposed decision making framework will be established.

CHAPTER 4 METHODOLOGY OVERVIEW

The four research questions lead to the essential elements of the methodology. The solution approaches are linked to the corresponding research questions. The methodology has seven steps, and the main elements can be grouped into two big analysis and optimization blocks. The proposed methodology is presented in Figure 42. This methodology enables the quantification of the impact of DR on the district energy system in two perspectives. Once the modeling and analysis of DR are completed, its capability of avoiding or deferring expensive capacity expansion is explored through the capacity expansion schedule optimization and the simulation based reliability analysis. The remainder of this chapter will describe how each of the steps is implemented.

This methodology can be applied to a specific type of problems that meet the following conditions.

- Campus-level district energy systems
- The building automation system is already established
- Historical DR data should be available
- The central plant has an expansion plan

These conditions restrict the applicability of the methodology, but many university campuses, where the district energy system is employed for heating and cooling, fit in this category of the problem. This methodology does not result in a single optimal solution. Instead, it provides an insight into the more flexible capacity planning options by visualizing the trade-off space.

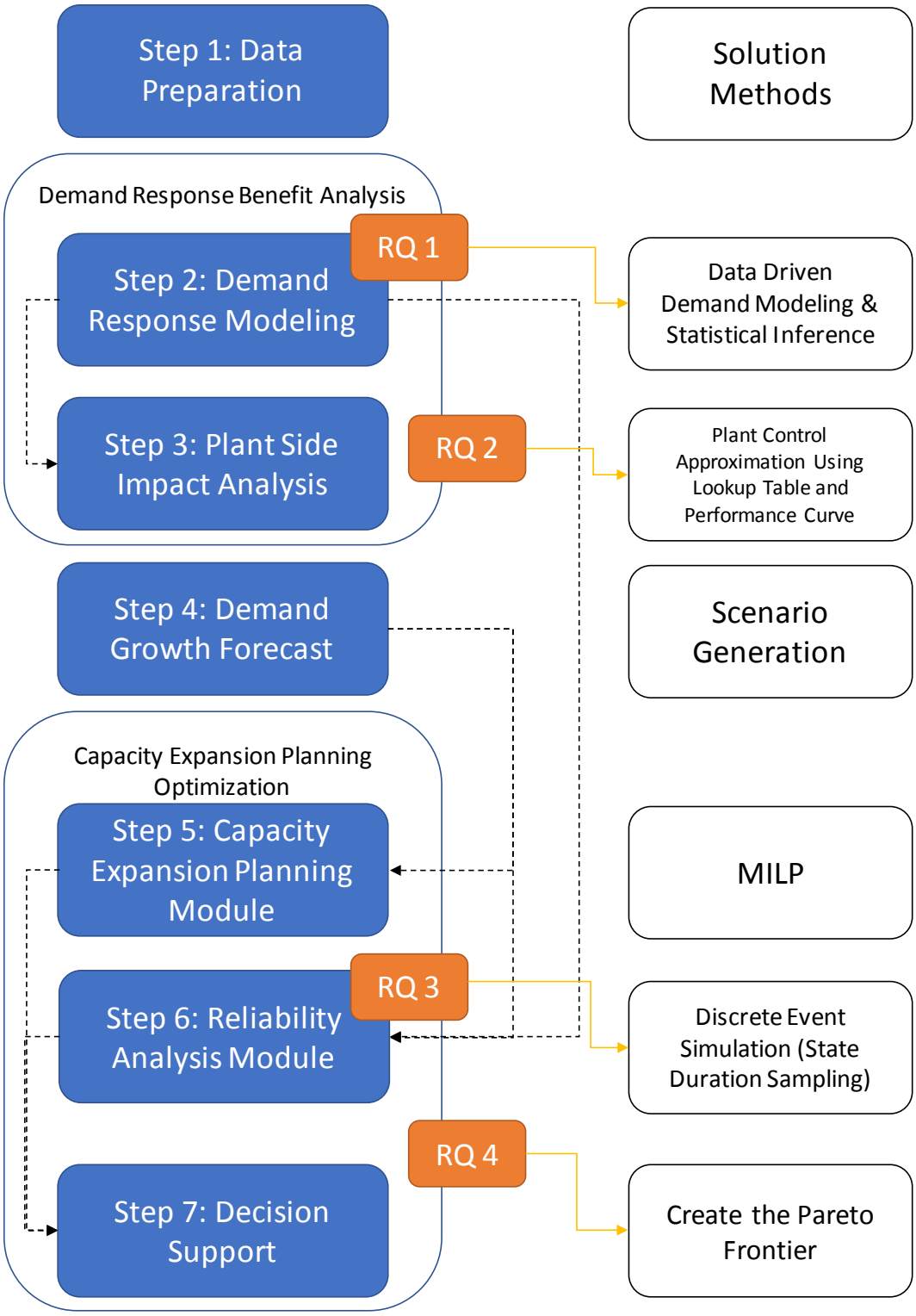


Figure 42 Proposed Methodology and Solution Approaches

4.1 Step 1: Data Preparation

The main prerequisite of this methodology is the data availability. Each of the models established in the following steps requires various types of data. Most of them are the historical measurement data while some of the steps require the descriptive data such as a master plan or the failure rate. The measured data is sometimes corrupted, missing, or containing errors. Bad data should be filtered before use. In the first step, data is collected and prepared from the various sources to be ready to use in the next steps. All measured data should be time-stamped. This type of data is accessible from the database of the building automation system. Most modern commercial buildings are equipped with the building automation and HVAC control systems. The required data types and descriptions are listed in Table 3. Data types are categorized into the measured and descriptive. If it is introduced in this thesis, the variable name is listed, or a just description is provided.

Table 3 Required Data for Each Step

Steps	Data Type	Variable	Description
Step 2	Measured	$\dot{m}_i$	Building chilled water flow rate
	Measured	$T_{S,i}$	Building supply temperature
	Measured	$T_{R,i}$	Building return temperature
	Descriptive	t_{Start}	DR event started time
	Descriptive	t_{End}	DR event ended time
	Measured	T_{Dry}	Dry Bulb Temperature
	Measured	RH	Relative Humidity
Step 3	Measured	$\dot{m}$	Plant chilled water flow rate
	Measured	T_S	Plant supply temperature

	Measured	T_r	Plant return temperature
	Measured	$T_{CHW,S}$	Chilled water supply temperature
	Measured	$T_{CHW,R}$	Chilled water return temperature
	Measured	$T_{CW,S}$	Condenser water supply temperature
	Measured	$T_{CW,R}$	Condenser water return temperature
	Measured	$\dot{Q}$	Cooling load of chiller
	Measured	P	Chiller power consumption
	Descriptive	$\dot{Q}_{ref}$	Chiller thermal capacity
	Descriptive	$\dot{m}_{ref}$	Chiller mass capacity
	Descriptive	P_{ref}	Chiller reference power
	Measured	T_{Dry}	Dry Bulb Temperature
	Measured	RH	Relative Humidity
Step 4	Descriptive		Year of new buildings
	Descriptive	$\dot{Q}_{NB,t}$	Estimated peak load of new buildings
	Measured	$\dot{Q}$(h)	Historical cooling load
Step 5	Descriptive		Purchasing & installation cost of the chiller
Step 6	Descriptive	MTBF	Mean time between failure
	Descriptive	MTTR	Mean time between repair

4.2 Step 2: Demand Response Modeling

Once data is prepared, the next step is to analyze the impact of DR and develop a data-driven DR model. Research Question 1 seeks the appropriate modeling approach when DR sample data is available. It is repeated here:

Research Question 1:

How can the benefits of DR for the entire campus buildings be assessed with a small number of sample buildings?

This question is too broad to find an answer. The raw data were examined to characterize the building cooling load dynamics and DR events. The observation of data leads to the sub-questions. The first sub-question arose from the fact that the original load during the DR active period cannot be measured, so the need for predictive load model development is addressed.

Research Question 1.1:

How can we create a credible enough model to quantify the cooling load curtailment by DR where small sets of data for DR events exist?

From the preliminary observations, the similar environmental conditions are found between DR and non-DR data. This observation formulates Hypothesis 1.1. It states that it is possible to create a predictive load model using non-DR data set to estimate the DR load shed by observing the difference between the predicted and measured load.

Hypothesis 1.1:

If a parametric model is built with non-DR data from similar environmental conditions as a function of weather and time parameters, then the fitted model will predict the original load during the DR active period. DR load shed can be discerned by the difference between the model predicted and measured load.

The purpose of the DR modeling is to evaluate the system level impact. If the DR data from the small number of sample buildings can be manipulated to represent the entire

system, then the goals can be fulfilled. The sub-question 1.2. asks the possibility of expanding the statistical characteristics of DR load shed to the entire campus buildings.

Research Question 1.2:

Is it possible to build a generalized DR model to predict the DR effectiveness of other buildings with sample data?

Contradictory opinions found from the literature survey and the thought experiment lead to the assumption that high uncertainty is involved in load prediction and DR activity. This thought summarized into Hypothesis 1.2

Hypothesis 1.2:

No apparent correlation will be observed regardless of the true relationship due to high uncertainty.

This is important because where the high uncertainty conceal the true trend (if exist), it may result in a large error in the prediction. Instead of developing a parametric model that predicts the relationship between the input parameter and the output, a probabilistic approach that treats DR load shed as a random variable can be more robust.

Through investigating Hypothesis 1.1 and 1.2, the DR model for the system level impact evaluation can be developed. The resulting model is passed to the next step to assess the impact on the chiller plant operation.

4.3 Step 3: Plant Side Impact Analysis

The previous step characterized the DR load shed of the building level cooling load as a random variable, and its distribution can be inferred from the load modeling and the

residual calculation. This step analyzes the impact on the plant level savings and explores the possibility of DR capability. Two questions depending on the perspectives can be answered in this step.

- What if we expand DR to the entire campus building, how much can we save? (As the energy consumer)
- What if we control the system in the same way as the current DR, can we increase availability? (As the energy producer)

The system level assessment requires a plant model that can respond to dynamic loads and yield the corresponding chiller staging and power consumptions because historical plant operation data do not exhibit the impact of DR. The small number of sample buildings' cooling load reduction is too small compared to the entire system's load. The dynamic plant model can fill this gap.

4.3.1 Modeling Tool Selection

The building energy simulation modeling tools are introduced in Section 2.1. Among the tools reviewed, Modelica Buildings Library is selected to model the district chilled water system. Modelica Buildings Library is developed by Lawrence Berkeley National Laboratory (LBL) to model building energy and control system [46]. Modelica is an object-oriented and equation-based modeling language. This property enables users to develop a model as modularized and hierarchical structure. Any component model can be developed as a physics-based equation. Using the Modelica language and the component model library, a campus energy network model can be created as another library. A modularized system model having the components as sub-models provides flexibility and extendibility. These are appropriate characteristics to build an evolving system model.

Since it is an equation based language, any physical phenomenon can be modeled. The LBL's Buildings library provides the basic district level component models, and this can be a good starting point [46]. Any fidelity level can be modeled as desired and the sub-models are replaceable if the inputs and the outputs are compatible.

4.3.2 Operation Modeling Approach

As Hydeman pointed out, developing the efficient control system for a chiller plant is a very complex problem [20]. There are too many parameters that affect the plant operational efficiency. Most of the existing systems employ the rule-based manual operation of chiller sequencing, which cannot be used to evaluate the DR impact because it does not yield the consistent result. Even though a dynamic control system can be developed using the library and the Modelica language, it is very time consuming and beyond the scope of this research. The purpose of this research is to explore the capability of DR if it can affect the chiller plant operation. Research Question 2 seeks a simple and efficient way of representing the plant operations.

Research Question 2:

How can the benefits of DR for a chiller plant be assessed without implementing a complex control logic?

The dynamic modeling and simulation environment can produce data for any conditions. If the loading conditions can be linked to the chiller combinations resulting in the minimum power consumptions, then it can represent the optimal operating points. This thought leads to Hypotheses 2.

Hypothesis 2:

If the optimal control design can be approximated as a set of the best combinations per loading conditions to minimize power consumptions, then it can predict changes of operational status of a chiller plant without a sophisticated control logic.

Once the operation model is created as the approximated curve and the lookup table, it can be incorporated into the assessment framework with the demand model. The assessment framework will be further described in Section5.2 with the detailed modeling and assessment process. This step concludes the DR impact assessment for the current system. The following steps expand the analysis to the future system.

4.4 Step 4: Demand Growth Forecast

This step is the preparatory step for the capacity expansion planning optimization block. The capacity expansion plan is confined to the future demand growth projection. In the energy sector, predicting energy demand is an essential activity to operate and manage the energy system. Load forecasting methods have been studied in the energy sector to get a more accurate prediction. It is commonly categorized into three terms as below [85].

- Short term: from one hour to one week
- Medium term: a week to a year
- Long term: longer than a year

Each of the planning terms requires different accuracy and time resolutions. Shorter term requires more accurate and finer resolution of the load forecast. Typically, the annual peak load projection and total energy demand are needed for the capacity expansion

planning while hourly dynamic load profile is required for the daily operation. To incorporate DR into planning problem, one should consider this discrepancy.

Time series forecasting is a commonly used technique for short term or long term demand forecast. It uses time-stamped data to extract time variant trends or seasonality. It doesn't need to consider input and output parameters to model demand. Chen developed a multi-scale time series forecasting method for power plant fleet management [86]. She used twenty years of data for the future demand forecasting. Espinoza et al. used the Periodic Autoregressive model to extract seasonality from five years of data for a short term forecasting [87]. As can be seen from those examples, time series forecasting requires a long period of data to get accurate forecasting results. In addition, it is a purely statistical method and dependent only on historical data so it is not suitable for the evolving system that may have a new feature unobserved before.

The main driver of the demand growth for the district level energy system is the connection of new buildings to the network because the relative size of the new building is much bigger for the district level system than the national system. The growth path will be like having a shock rather than the gradual increase. A detailed construction plan is specified by the campus master plan, which describes the upcoming building types, the years to be built, the estimated areas, or the locations. Traditionally, only the peak load estimation is used to forecast the future demand growth in capacity expansion planning for the central district chiller plant. However, the peak load is not enough to evaluate the system reliability. The annual cooling demand is characterized by two attributes: the peak load and the load shape. The peak load is used for sizing, and the load shape is used to evaluate the operational performance and reliability. The load shape is mainly affected by the weather profile and the occupancy schedule. As can be seen in Section 3.1, the load shape is the unique property of a building. The campus load is the aggregation of those

unique buildings, and the campus will have its own load shape. Assuming the campus operation schedule would not change in the future (business hours, semester schedule, etc.), the current annual load profile will maintain while the peak load is increasing. The additional buildings are assumed to contribute the peak load growth. Based on this assumption, the current load shape can be normalized and scaled up with the expected peak load of the future. The current load shape can be obtained by the data-driven model. As in the building load modeling, ANN is used to create the aggregated plant load model using the time and the weather variables as the inputs. Once the model is created, it is normalized by the peak load and baseload using Equation (41) as presented in Figure 43. This normalization transforms the annual load profile into the range between 1 and $\dot{Q}_{min}/\dot{Q}_{max}$. The base load, which is the relative value of the peak load, is used to prevent the minimum load from becoming zero. The normalized load is stored as an array of the vector with the length of 8760 hours so that it can be scaled up by multiplying the peak load of year t as in Equation (42). The future peak load at year t $\dot{Q}_{peak,t}$ is estimated from the master plan. The additional building is treated as an additive variable to the peak load projection. The peak load at year t is calculated using Equation (43). Once the peak load estimate is updated over the planning horizon, the peak load projection can be obtained as an array of the vector with the length of the planning horizon T.

$$\dot{Q}_{normalized}(h) = \frac{\left(\dot{\mathrm{Q}}(h) - \dot{Q}_{min}\right) \times \left(1 - \dot{Q}_{base}\right)}{\dot{Q}_{max} - \dot{Q}_{min}} + \dot{Q}_{base} \qquad (41)$$

Where, $\dot{Q}_{base} = \dot{Q}_{min}/\dot{Q}_{max}$

$\dot{Q}_{max}$: the annual peak load, $\dot{Q}_{min}$: the annual minimum load

$$\dot{Q}_t(h) = \dot{Q}_{peak,t} \times \dot{Q}_{normalized}(h) \tag{42}$$

$$\dot{Q}_{peak,t} = \dot{Q}_{peak,t-1} + \dot{Q}_{newb,t} \; \forall t \in [1, T] \tag{43}$$

$$\dot{Q}_{peak,0} = \dot{Q}_{max}$$

$\dot{Q}_{newb,t}$: the estimated peak load of the new building of year t

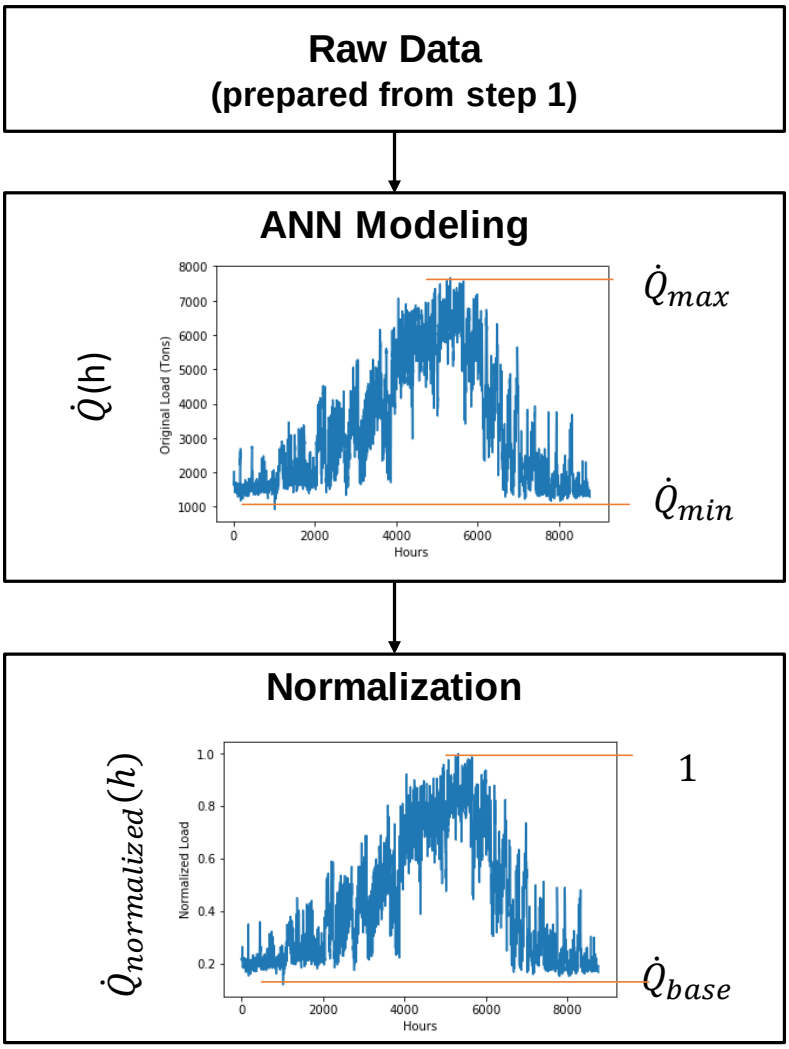


Figure 43 Load Profile Normalization to Extract the Shape Feature

4.4.1 Scenario Generation

The estimated peak load and the probable year to be connected can be known from the master plan, but it is still subject to change. Sources of the uncertainty in the long-term load forecasting are diverse. The new buildings' peak load can deviate from the estimation, and the construction can be delayed or rushed. Climate change can raise the cooling load for the existing system. If the climate change is significant, then the load forecast based on the historical weather profile may become invalid. The fuel cost or market uncertainty, which is important in power sector, is not significant for a campus level district energy system. Even though fuel cost will affect operating cost, it would not change the decision because all the options are the electric chillers. Regardless of the electricity price, the plant consumes the same amount of energy for the same demand. Therefore, the deviation of cooling demand from the projection will be the main concern in the load forecasting.

Capacity planning optimization approaches under uncertainty were reviewed in Section 2.4. While the stochastic programming technique provides a robust decision-making framework, it requires to enumerate all the scenarios and is more suitable to the problem trying to optimize multi-time scale decisions at the same time. Since this research is not aiming to control DR in real-time but attempts to evaluate its capability in the planning phase, scenario analysis with a deterministic optimization framework is enough to quantify the uncertainty in the planning. The advantage of scenario analysis over the stochastic programming is that it is easily incorporated into the deterministic optimization framework and allows a transparent decision making by providing the multiple scenarios and design space exploration capability and leave the selection to a human decision maker.

Based on assumptions of the future building peak load distribution, demand growth rate, and the construction year's delay or advancement, the future load growth path is

determined. Monte Carlo Simulation is used to quantify the uncertainty in the future load growth path. Uncertainty on the supply side is also important such that the reliability of the system is dependent on the system capacity availability. This will be further covered in Step 6.

4.5 Step 5: Capacity Expansion Plan Optimization

Given a scenario of the future demand growth forecast, optimal size and schedule for the new chiller installation plan are determined in this step. The objective of the capacity expansion plan is to minimize the capital cost over the planning period while meeting the demand satisfaction constraints. The problem is solved using the MILP formulation introduced in Section 3.4. To maintain the redundancy requirement, total plant capacity should keep excess capacity as the amount of the largest chiller's capacity. For the DR option as a contingency plan, no additional capacity is needed. However, to account for a decision maker's preference, a small safety factor is introduced. The safety factor is varying in the range between zero and the largest chiller capacity. For the same demand forecast scenario, multiple plans are generated by adjusting the safety factor. The resulting plans and the corresponding plant configurations of each year are transferred to the next step to evaluate those system's reliability and the improvement of reliability metrics by DR. The safety factor will relax the strict constraint and allow to explore the design space of capacity expansion plan.

4.6 Step 6: Reliability Analysis

Once the optimal capacity expansion schedule is determined, this step calculates the reliability metrics introduced in Section 3.3. Research Question 3 seeks the suitable method to compare the redundant system and the DR-enabled system.

Research Question 3:

How can the performance of DR as a contingency plan for a district chilled water system be measured and fairly compared to the deterministic redundancy?

This was answered by the investigation of the existing reliability metrics and definition of the failure. The probability of failure at the peak load is derived from the redundancy definition to account for uncertainty in the failure mechanism. To evaluate the overall system reliability through the entire loading condition, the dynamic reliability metrics are borrowed from the grid system reliability. Forecasting the future system's reliability relies on various assumptions either from supply and demand side. The stochastic simulations models for both demand and supply side are needed. The future demand profile for the particular year is generated in Step 4. In this step, the plant simulation model should be developed.

4.6.1 Simulation Methods for Reliability Assessment

A plant availability simulation model is required to calculate the probabilistic reliability metrics. A chiller plant is a parallel system of repairable components. The simulation methods commonly used in reliability assessment for a multi-component system are simple point Monte Carlo Simulation, Discrete Event Simulation, and Markov Model. In power system, they are named as state sampling, state duration sampling, and system state transition sampling [55].

- Point Monte Carlo Simulation / State sampling

A system state is defined by the combination of component states. The component state is a binary random variable and determined by sampling with the probability that a component is at the state.

$$x_i = \begin{cases} 1 & if\ u \in [0, fr_i] \\ 0 & if\ u \in [fr_i, 1] \end{cases}$$

$u \sim U[0,1]$: u is a random number with a uniform distribution between [0,1] (44)

fr_i: failure rate of component i

The system state X consisting of N number of the component is expressed by a state vector.

$$\mathrm{X} = [x_1, \cdots, x_N] \quad (45)$$

This method is simple and requires minimal information, but it cannot be used to calculate the dynamic reliability metrics.

- Discrete Event Simulation / State Duration Sampling

In this approach, the state is determined by the probability distribution of time duration that a component stays at that state. By simulation, a state transition occurs after random time passes. Using this approach, a time sequential array of states of a component can be obtained without continuous time simulation. A system state is determined by the combination of component states. Typically, the duration of a component staying at a state has an exponential distribution. The reliability metrics can be calculated using this approach.

- Markov Model / System State Transition Sampling

The system state transition sampling approach simulates the whole system's state and its transition to another state. Since the system state is defined as a whole, this approach does not require sampling of a component's state. It assumes that the time duration of a component staying at a state follows an exponential distribution, so the duration of a system at a specific state also follows an exponential distribution. If a system has a large number of components, this approach can save the storage by eliminating the necessity of recording time sequential state vector of components. Instead, it should define the multiple system states, and the transition probability should be known prior to the simulation.

For a chiller plant, which does not have too many chillers, the state duration sampling method is most suitable to calculate dynamic reliability metrics. Each chiller can be modeled as a two-state component as presented in Figure 44.

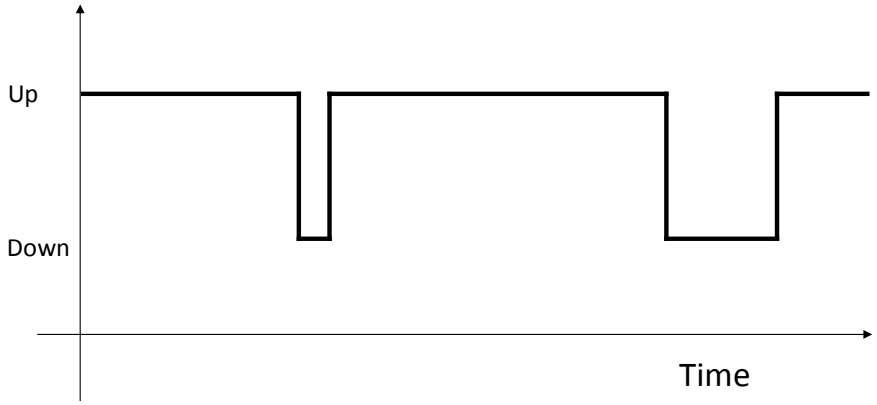


Figure 44 A sample sequence of a chiller's state over time

Discrete event simulation (or the state duration sampling technique) samples the time duration in each state as in the sequence of events. Simulation collects the sequence of the duration of the interchangeable states until the specified simulation time ends (e.g.,

8760 hours for one year). When the simulation is done, the chronological state vector can be generated from the result. Once the vectors for each chiller are created, the total available plant capacity can be calculated by Equation (46).

$$plant\ capacity = \sum_{i}^{N} Cap_i \times S_i \tag{46}$$

$$S_i = [x_1, \cdots, x_{8760}]$$

4.7 Step 7: Decision Support

The final step wraps the previous three steps and provides a visualized design space as the Pareto frontier. This step is formulated based on Hypothesis 4.1 as the answer to Research Question 4.1.

Research Question 4.1:

How can we know the optimality of the solution for the capacity expansion plan with the proposed decision making framework?

Hypothesis 4.1:

With the two extreme points of the N+1 constraint and no safety margin, the Pareto frontier can be generated by adjusting safety margin on the demand satisfaction constraint between the two points and the fraction of DR buildings.

The two-step analysis yields a solution with the discounted capital cost and the reliability cost in terms of the required load shedding time over the planning horizon. Varying constraints and scenarios produce multiple solutions. These will create a Pareto

frontier on the capital cost vs. the reliability cost plot. Once the Pareto frontier is generated, then it will support Hypothesis 4 to answer Research Question 4.

Research Question 4:

How can we predict if DR is capable of avoiding or deferring capacity expansion without sacrificing reliability?

Hypothesis 4:

The proposed two-step decision making framework can yield multiple Pareto solutions and the corresponding optimal capacity expansion planning schedule. Using those, one can predict the impact of DR on the capacity expansion planning and the system reliability.

CHAPTER 5 MODEL DEVELOPMENT AND EXPERIMENTS

This chapter will describe how each of the research questions leads to a specific solution by developing modeling and simulation environment and testing hypotheses. Once the resulting modeling approaches are validated, all the elements are integrated into a methodology to answer the primary research question and demonstrate the capability of the proposed solutions. Each section will demonstrate how the constituent models are developed and validated through the experiments.

5.1 Demand Response Assessment

Research Question 1 led to the investigation of a suitable modeling approach to the baseline demand and load shed by DR using a small set of data to evaluate the benefit of DR at the campus level. A data-driven demand modeling approach and statistical inference for DR load shed estimation are proposed as the solution approaches. From the primary research question, two sub research questions were derived, and hypotheses are formulated for each of the sub research questions. A methodology for the DR assessment is developed to test hypotheses in this section.

5.1.1 Test Methodology

A data-driven DR modeling methodology is developed to answer the Research Questions 1. Experiments are conducted to see if the predictive model can discern the load reduction by DR. This methodology uses data and statistical methods to analyze a specific type of DR method. The example DR method is the global setpoint temperature control of the building HVAC system to reduce the cooling load of the building. The proposed method is only applicable to the same DR scheme. The methodology is presented in Figure 45. By conducting tasks through the process, hypotheses will be tested and validated with

the observation of the results. The first step requires historical data from the building energy meters and the DR event log. The collected data should be grouped into DR and non-DR data with timestamps. A parametric model is created by fitting to the non-DR data set. Once the baseline model is developed, the residuals are calculated using Equation (19) for every data point. Resulting data will be used to test Hypothesis 1.2. Lastly, a generalized DR model will be created. The proposed methodology and testing hypotheses will answer the Research Question 1.

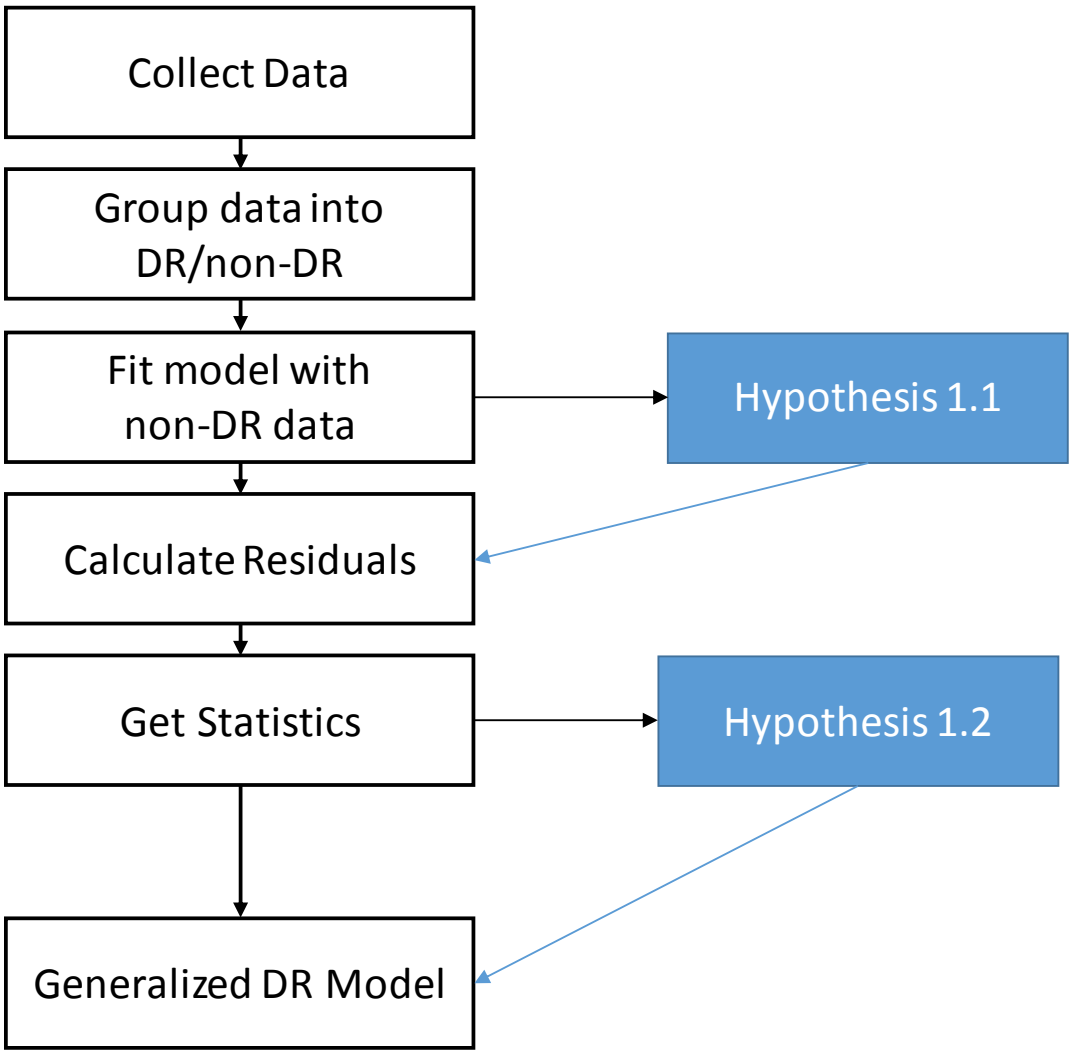


Figure 45 Data-Driven DR Modeling Methodology

5.1.1 Data Preparation

Data used for the DR modeling and analysis were collected from the database, which was recorded between 6/1/2016 and 8/31/2016. All data have the 15-minute interval resolution. During this period, eighteen automated DR events had occurred including one weekend event. Samples are available for every weekday and a Saturday. Days of the week when DR events were called are summarized in Table 4. A DR event typically was activated around 2-3 pm and inactivated around 7 pm. The longest DR event lasted about six hours, and the shortest event lasted about one hour.

Table 4 Number of Samples per Day of Week

Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
4	4	2	4	3	1

Cooling load for three pilot buildings and weather data were collected for the same period. The model input parameters and the ranges are summarized in Table 5. Data are divided into non-DR and DR groups. Unfortunately, data are sometimes corrupted and recorded as invalid numbers. These data were excluded. The cooling load for the three buildings and the weather profile, the dry bulb temperature (°F) and the relative humidity (%), are presented in Figure 46 and Figure 47, respectively. The red colored fraction of data indicates the DR active period. Building A and C show the apparent pattern that the buildings are automatically controlled to reduce load during nights and weekends while Building B is naturally following the temperature pattern.

Table 5 Model Input Parameters and Range

Variable Name	Non-DR (8570 samples)		DR (244 samples)	
	Min	Max	Min	Max
Day of Year	153	244	165	244
Day of Week	0 (Monday)	6 (Sunday)	0	5
Minute of Day	0 (12:00am)	1425 (11:45pm)	825 (1:45pm)	1185 (7:45pm)
Temperature	62.74	100.55	72.59	100.55
Relative Humidity	33.69	99.8	38.48	99.78

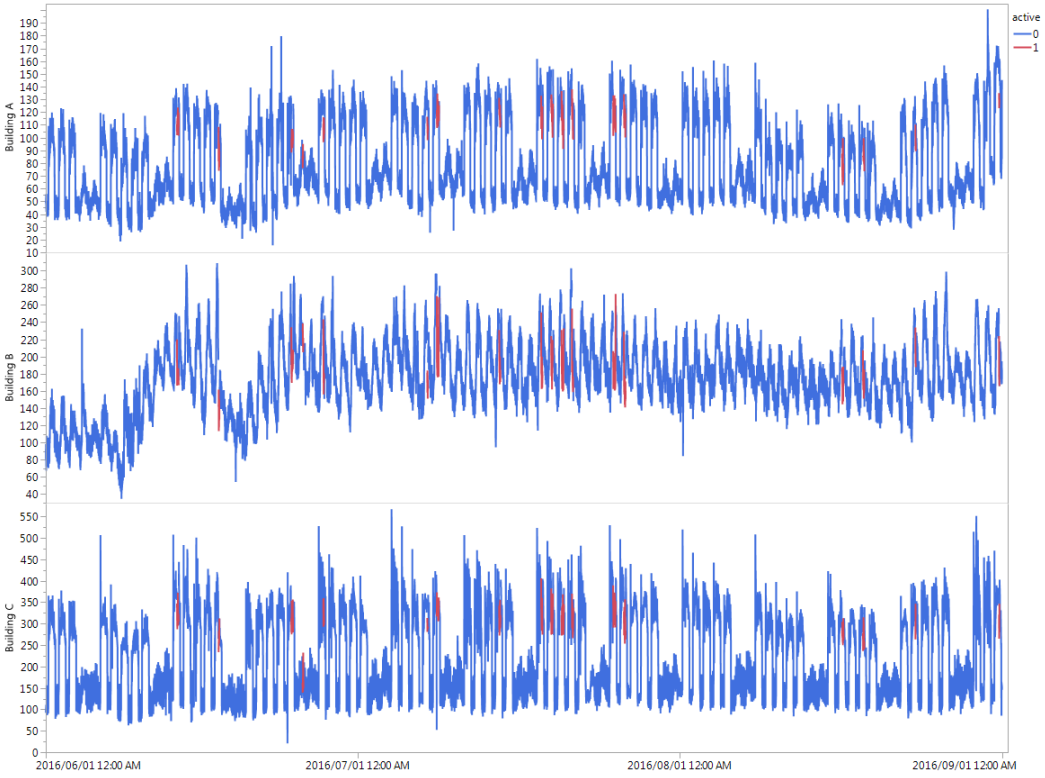


Figure 46 Cooling Demand Pattern for Buildings During Summer

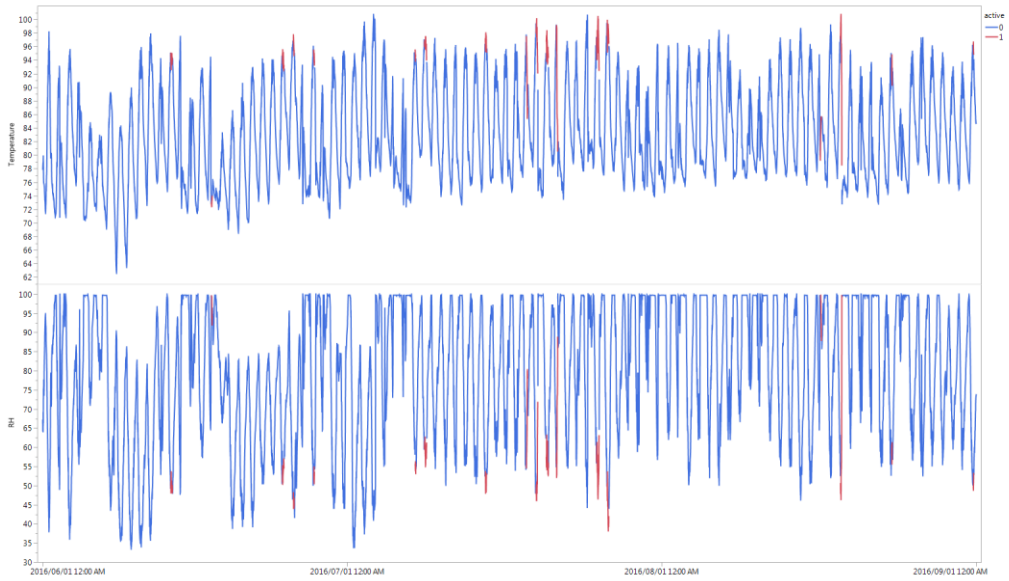


Figure 47 Weather Profile

As explained in Section 3.1.1, three buildings are not distinguishable by the building type even though they are different in size. Three is not enough number to differentiate the DR effectiveness based on the building type. Therefore, the building type is not considered in this study, and all the campus buildings will be assumed to be the same type.

5.1.2 Experiment 1-1: Model Validation

Research Question 1.1 is formulated to figure out how the baseline model of cooling load should be created. It is repeated here:

Research Question 1.1:

How can we create a credible enough model to quantify the cooling load curtailment by DR where small sets of data for DR events exist?

The preliminary data analysis led to Hypothesis 1.1, which suggests using non-DR data having the similar environmental condition to construct the predictive model.

Hypothesis 1.1:

If a parametric model is built with non-DR data from similar environmental conditions as a function of weather and time parameters, then the fitted model will predict the original load during the DR active period. DR load shed can be discerned by the difference between the model predicted and measured load.

By fitting a model to non-DR data and testing the goodness of fit of the model, Hypothesis 1.1 will be tested. From data observations, it was concluded that the model should be able to capture the building schedule and the weather effect. Artificial Neural Network (ANN) is used to capture the nonlinear relationships between cooling load and weather, occupancy schedules and activities inside of the building. The reason why I chose Artificial Neural Network model is that it can capture the nonlinearity of the relationship between parameters, not to promote ANN model as the best approach. As Zhang pointed out, the prediction accuracy is not differentiated among the data-driven models [30].

The used neural net has a multi-layered structure with an input layer, two hidden layers, and an output layer. The input layer has five input nodes for weather and time parameters. Dry bulb temperature and the relative humidity are used to represent weather effects and three time parameters are used to capture the occupancy schedule and seasonality. These time parameters are a day of a year (0 to 364), a day of a week (0 to 6), and a minute of a day (0 to 1440). Each hidden layer contains 15 hidden nodes with a tangent sigmoid as the activation function. The single output node is a linear unit to model each cooling load measurement. The weight training has been done for a batch of 75% of the full data point set using the Lavenberg-Marquardt algorithm with backpropagation error

for gradient calculations. The validation error is monitored during the training process for 15% of the entire data points in order to prevent overfitting.

The predicted vs. the actual plots for three buildings are presented in Figure 48, and the model error statistics are summarized in Table 6. They show that the models are good enough in terms of the error. Data points on the scatter plot are well gathered along the diagonal line for all the three buildings. The building C has fatter and slightly asymmetric data point cloud compared to two other buildings. However, its error statistics are in the acceptable range. The coefficient of determination (R^2) is commonly used to evaluate a predicted model's goodness of fit and defined as in Equation (47). The CV-RMSE, which is introduced in Section 2.1 and commonly used to evaluate the building energy model's accuracy, is also listed. The R^2 for all three buildings are higher than 0.9 and the values of CV-RMSE are much higher than the acceptance limit of ASHRAE guideline (30%). Therefore, the fitted model is good enough to predict the cooling load.

$$R^2 = \frac{\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{47}$$

Where, $\hat{y}_i$: the model predicted value

y_i: the measured value

$\bar{y}$: the mean of the measured value

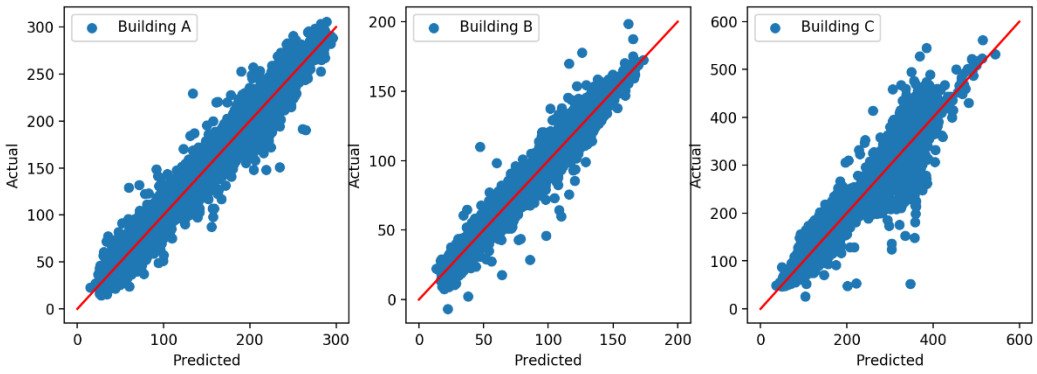


Figure 48 Predicted vs. Actual Plot

Table 6 Error Statistics

	R^2	CV-RMSE (%)
Building A	0.96	8.19
Building B	0.97	7.29
Building C	0.94	11.55

Figure 49 shows how the individual building model captures the cooling demand on DR days and non-DR days. On this plot, there were DR events on the first four days, which are depicted as different colors. The dashed lines indicate the model predicted cooling load that would have been without DR, and the solid lines are the actually measured cooling load. The actual measurements have fluctuations, which might be one of the sources of error in the DR load shed estimation. The building C has the largest fluctuations, and it is the reason for that its goodness of fit is the worst among three buildings. The model prediction does not show this fluctuation. Instead, it smoothes out the demand profile to uncover the trends. The baseline model can even capture the morning peak on Building A and C. As can be seen in Figure 49, the reduced load during DR active periods is well captured by the gap between the predicted and the measured load.

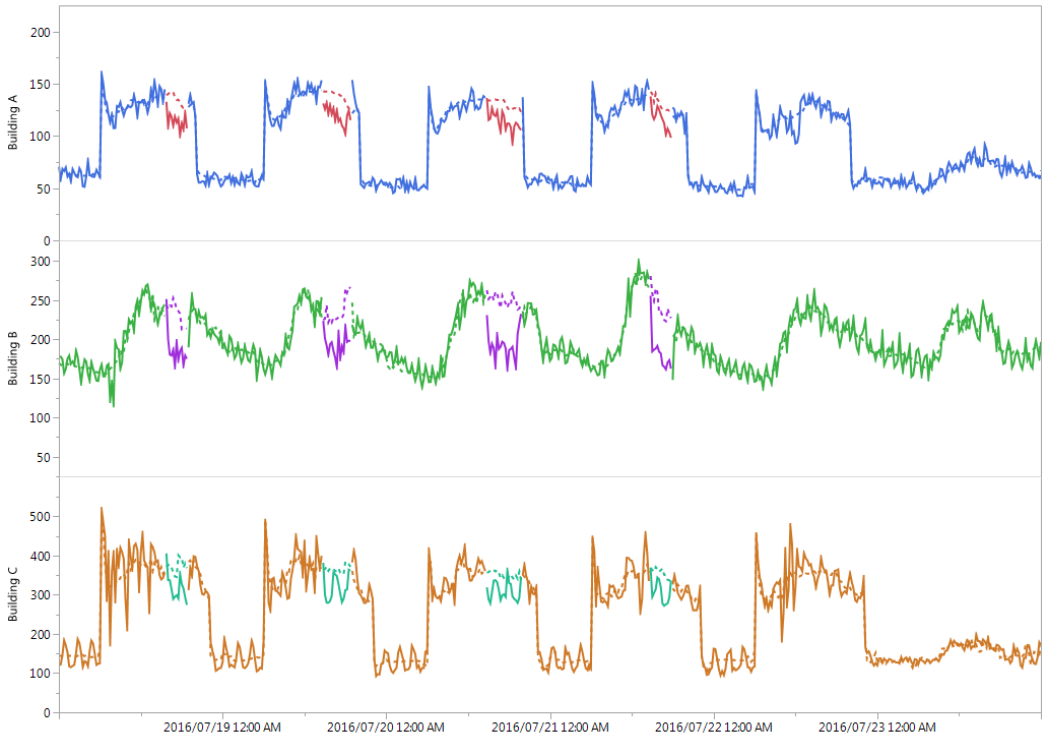


Figure 49 Model Predicted vs. Actual Demand Profile

The averaged load profile on DR and non-DR days for the 24-hour period are presented in Figure 50. The model can capture the characteristics of cooling load pattern for the three buildings, and the discernible gaps on DR days indicate the load shed by the DR operation. The building A and C have the morning peak due to the night setback control, which shifts the peak time to early in the morning around 6 am and reduces the daytime cooling load. On the other hand, the building B has the afternoon peak. Due to the morning peak of the building A and C, no other peak is observed larger than the morning peak, and the load profile is flattened compared to Building B. That is why the building B has the largest gap during the DR event. Since the DR events were not synchronized with the buildings' peak times, the peak load reduction effect by DR is not attained for these buildings.

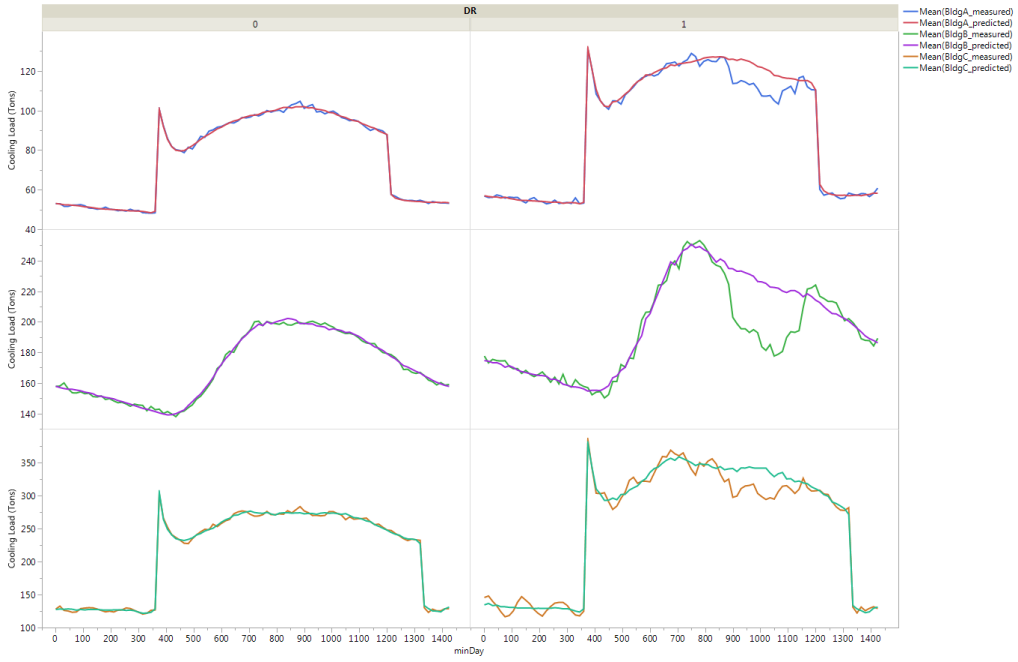


Figure 50 Average Load Pattern on DR and non-DR days

In summary, the data-driven models fitted to non-DR data are good enough to predict the cooling load and discern the load shedding by DR. Calculating the goodness of fit and data visualization support the claim. Therefore, Hypothesis 1.1 is valid. Now that the DR load shed is distinguished, the next step will follow to produce data to test Hypothesis 1.2.

5.1.3 DR Load Shed Calculation

Visual inspection proved that the data-driven model could disclose the load shed by DR. In this section, the difference between the measured and the model predicted value will be calculated and normalized for all data points during the DR active period using the Equation (19). Data are indexed with time variables. A sample DR day of the building B is illustrated in Figure 51 to help to understand the calculation method. There are nine

sample points of DR in this example. The relative amount of load shed ($\Delta Load_{DR}$) will be calculated for all the 244 sample points of three buildings.

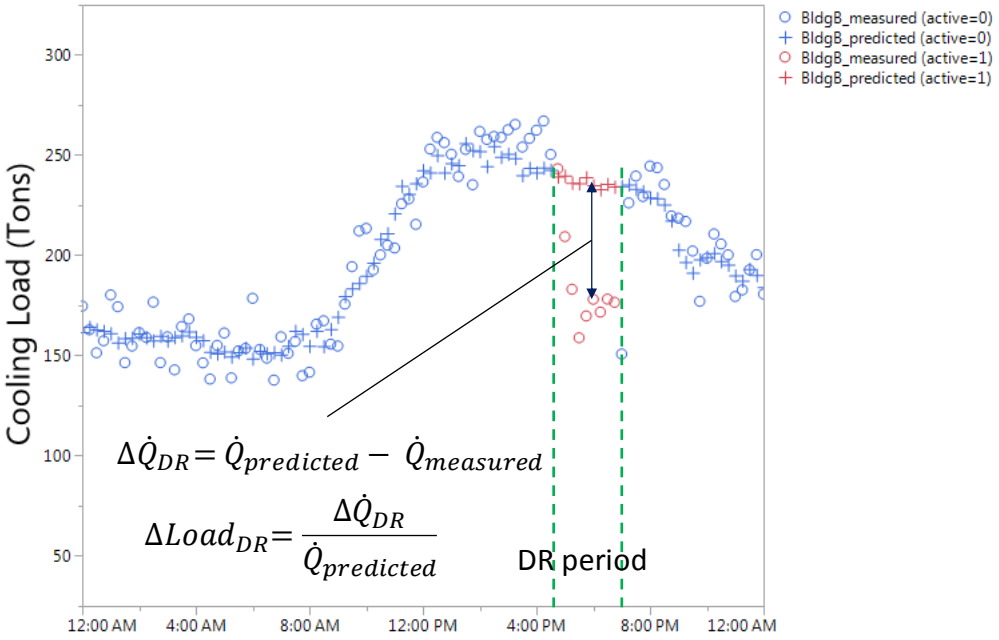


Figure 51 DR Load Shed Estimation Using the Predicted Model

The daily average $\Delta Load_{DR}$ with a standard error bar for the eighteen DR days are presented in Figure 52. It shows variability across the events and buildings. The negative $\Delta Load_{DR}$ values were observed on the fourth event, which is Saturday. On holidays and weekends, there are few occupancies in campus buildings, and some of the buildings are operated with a pre-programmed schedule. The automatic setpoint control in responding to DR signal may not be functioning on this day, but it could not be confirmed because setpoint data is only available for the recent 24 hours. Whatever the reason of the negative value in DR load shed is, it is more likely the model prediction error than the actual increase in load. Therefore, data points on this day were excluded for analysis.

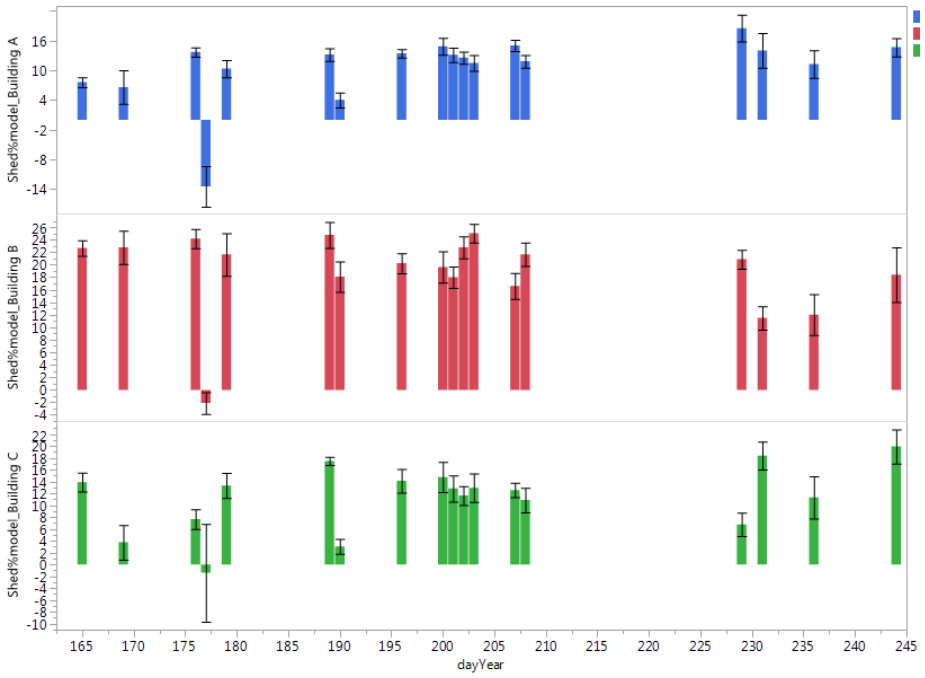


Figure 52 Average DR Load shed per Event

In the next step, multivariate analysis will be performed on the calculated DR load shed and weather and time parameters to quantify the DR benefit and investigate Research Question 1.2 if the sample data can be generalized.

Research Question 1.2:

Is it possible to build a generalized DR model to predict the DR effectiveness of other buildings with sample data?

5.1.4 Experiment 1-2: Data Analysis of DR Load Shed

This step will conduct statistical analysis of DR load shed data to test Hypothesis 1.2. The question arose from the observation that a couple of previous research did not agree on the load shed model parameters and their relationship. Hypothesis 1.2 suggests

that the large uncertainty is the cause of the disagreement and it will conceal the true trend if any.

Hypothesis 1.2:

No apparent correlation between load shed and the model input parameters will be observed regardless of the true relationship due to high uncertainty.

To test this statement, the relationships between the calculated load shed and the environmental parameters are thoroughly examined using statistical methods including visual analytics. Multivariate plots for the buildings are presented in Figure 53, Figure 54, and Figure 55 for each building. These plots are generated from the seventeen weekdays' DR load shed data. No building among the three shows a strong correlation between the load shed and temperature or minute of the day. The correlation coefficients (r) are marked on the plot to show the quantitative correlations between the load shed and the input variables. Most of them are very small, which means that there are weak or no linear relationship between load shed and temperature or time. Thus, the linear relationship cannot explain the relationship between the load shed and the time of day or the temperature.

The distribution of the load shed seems to be close the normal distribution. They have a symmetric distribution (the building A and C) or slightly skewed toward the right (the building B). Since the building A and C have similar schedules while the building B is sensitive to the temperature trend, the distribution might be associated with the building type or the schedule. However, it is too hasty to conclude that relationship given the small number of samples. Moreover, the negative values on the estimation might cause the skewness. Whether it is symmetric or skewed, it seems that the distribution around the mode (the most frequent value) is caused by random uncertainty.

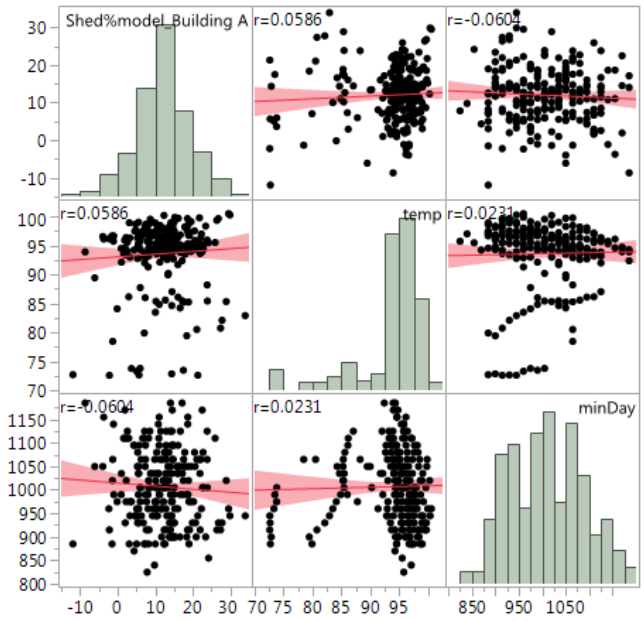


Figure 53 Multivariate Plot for Building A: Load Shed (%) vs. Temperature and Minute of Day

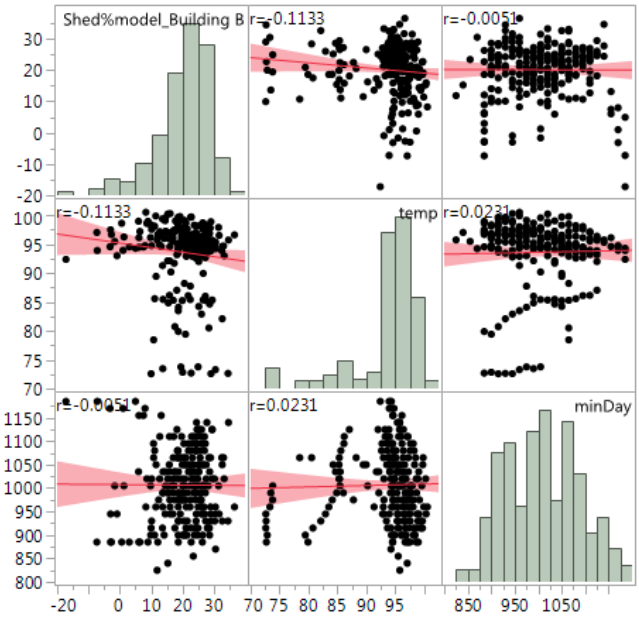


Figure 54 Multivariate Plot for Building B: Load Shed (%) vs. Temperature and Minute of Day

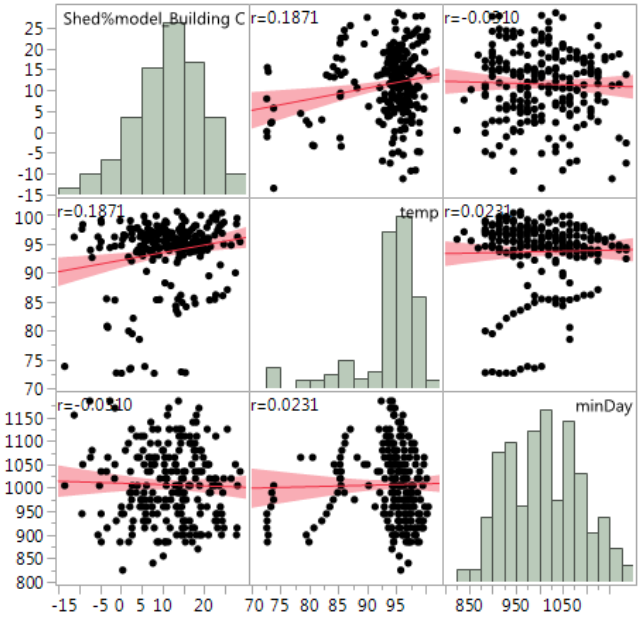


Figure 55 Multivariate Plot for Building C: Load Shed (%) vs. Temperature and Minute of Day

Most of DR days have similar weather conditions except two days that appear to be apart from the group as highlighted in Figure 56. They were unusual days having lower average temperature and higher relative humidity compared to other days. Note that the DR method studied in the research is the price based DR. The electrical utility company sent the dynamic price signal, and the building's HVAC system automatically adjusted the setpoint temperature when the price was higher than the threshold. The mechanism to determine the price is unknown. Hot weather, which causes the high cooling demand, may drive up the price, but it is not the only reason to increase the price. Whatever the pricing mechanism behind is, it cannot be controlled by the community that has enrolled in the DR program. Excluding these two days, most of the data points are under the similar condition.

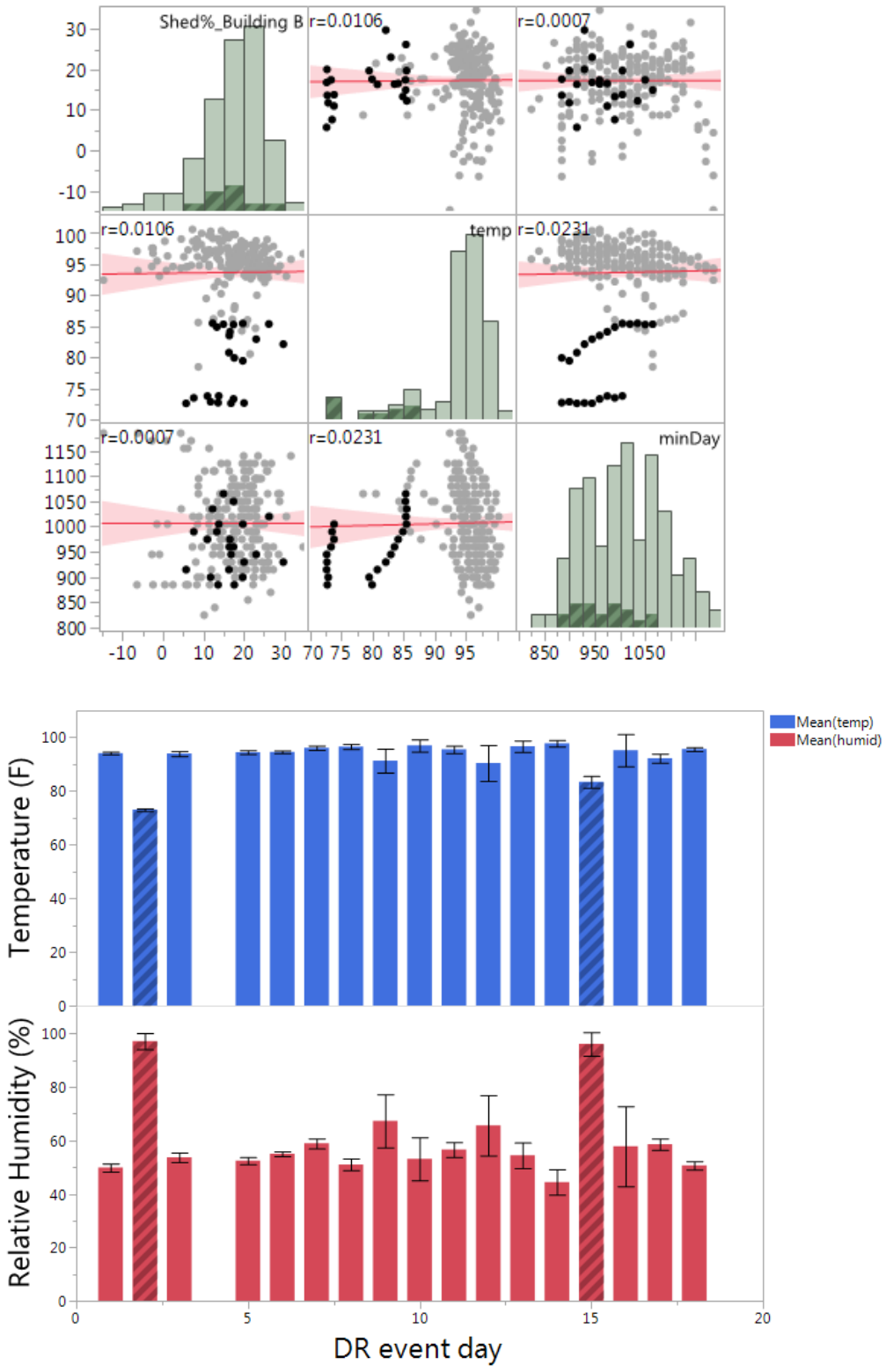


Figure 56 Environmental Conditions during Active Period on DR days

The amount of load shed on the days with lower temperature are not smaller than other days. They are dispersed across the range. This observation is another evidence that DR load shed is independent of the weather variables. If the dry bulb temperature and relative humidity are transformed to the wet bulb temperature, then the unusual days reduced to one day as presented in Figure 57. The red boxed scatter plots show that no building has the significant correlation with wet bulb temperature, and the highlighted data of the unusual day are evenly dispersed. Similarly, no significant correlations were observed between the load shed and the minute of a day as presented in Figure 58.

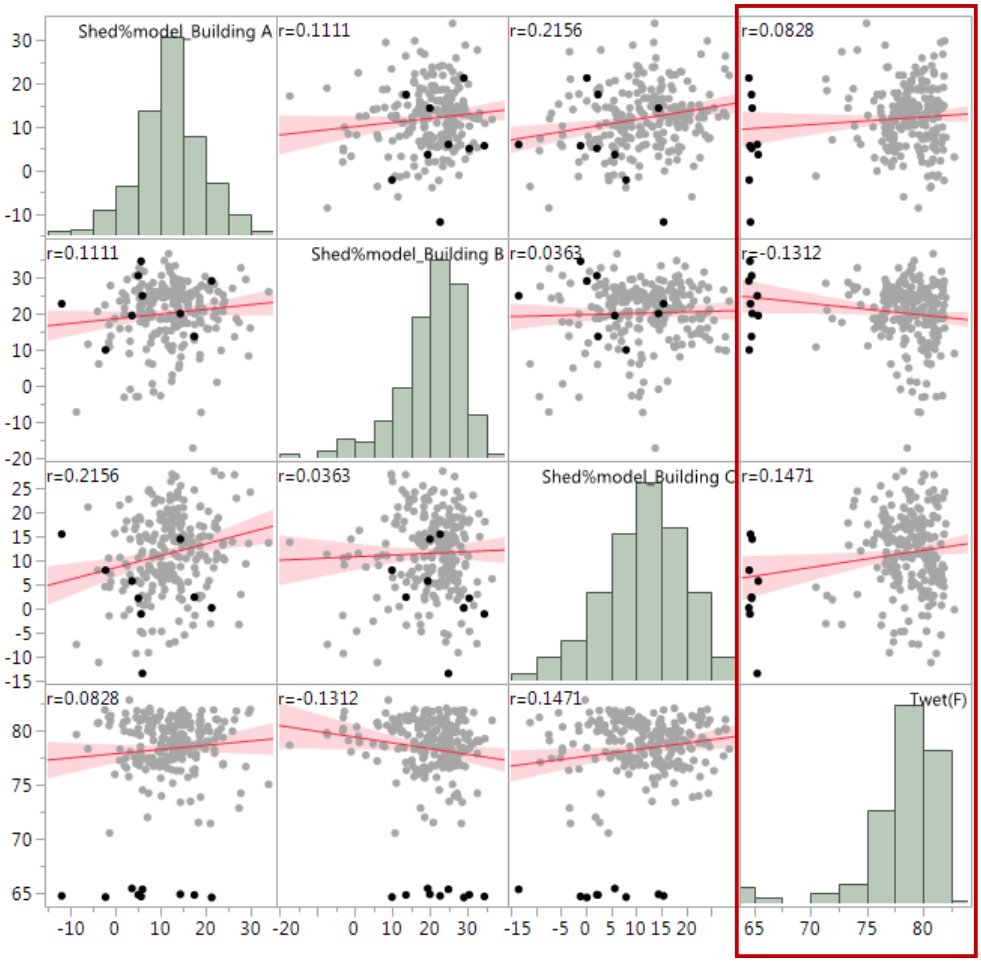


Figure 57 Correlations of Load Shed (%) and Wet Bulb Temperature

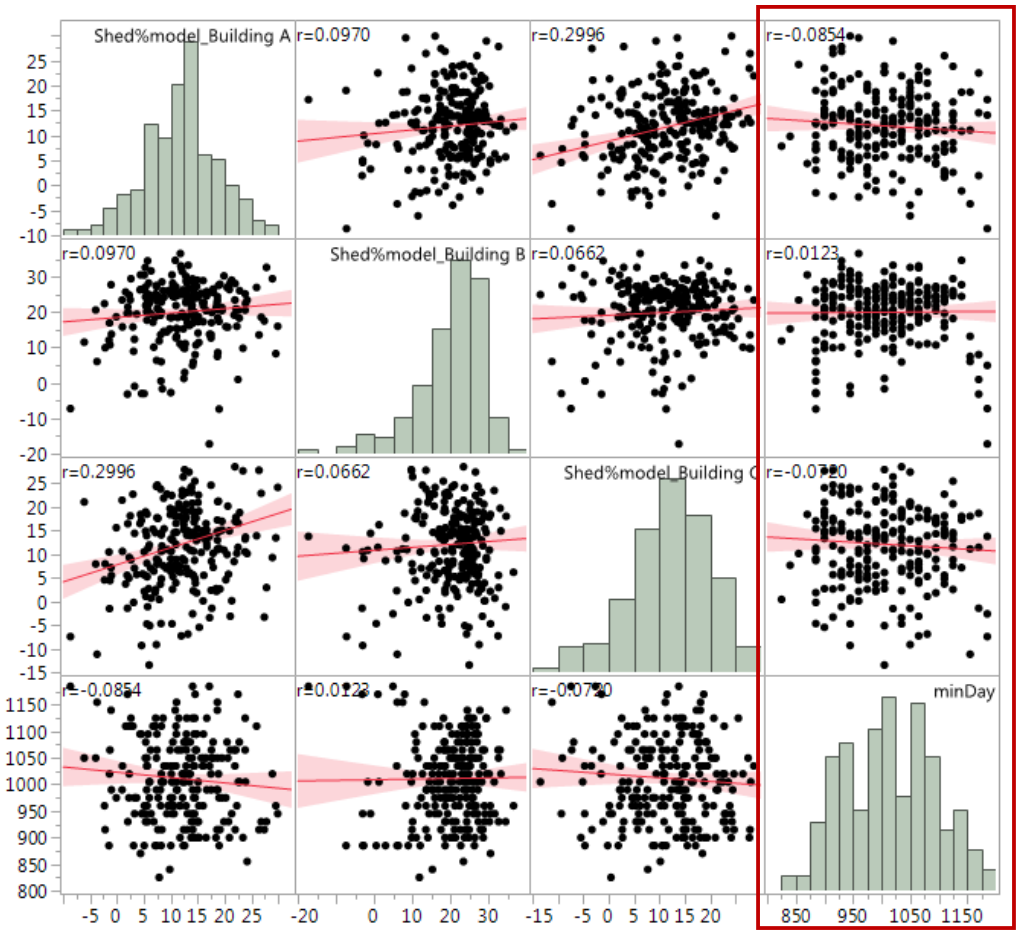


Figure 58 Correlation of Load Shed (%) and Minute of Day

No evidence is found to prove that there is a correlation between the load shed and the time or the outside temperature, which supports Hypothesis 1.2. There might be some degree of the trend as the correlation coefficient is not completely zero. However, the deviation from the estimation is much bigger than the difference between the two extreme conditions on the trend line. Unless it considers this uncertainty in the prediction, developing a parametric model as a function of the weather and time variables is not suitable to estimate the DR benefit.

The histograms of the load shed for each and the aggregated amount of the three buildings, and their summary statistics are presented in Figure 59. As pointed out earlier, the distribution has a mode and symmetrically distributed. The building B is slightly skewed to the maximum. If the negative values are truncated, then it becomes close to the symmetric distribution. Its mean value of the load shed is higher than other two buildings. As can be seen in Figure 50, the building's schedule made the difference. The sum of three buildings results in about 15% of load shed on average.

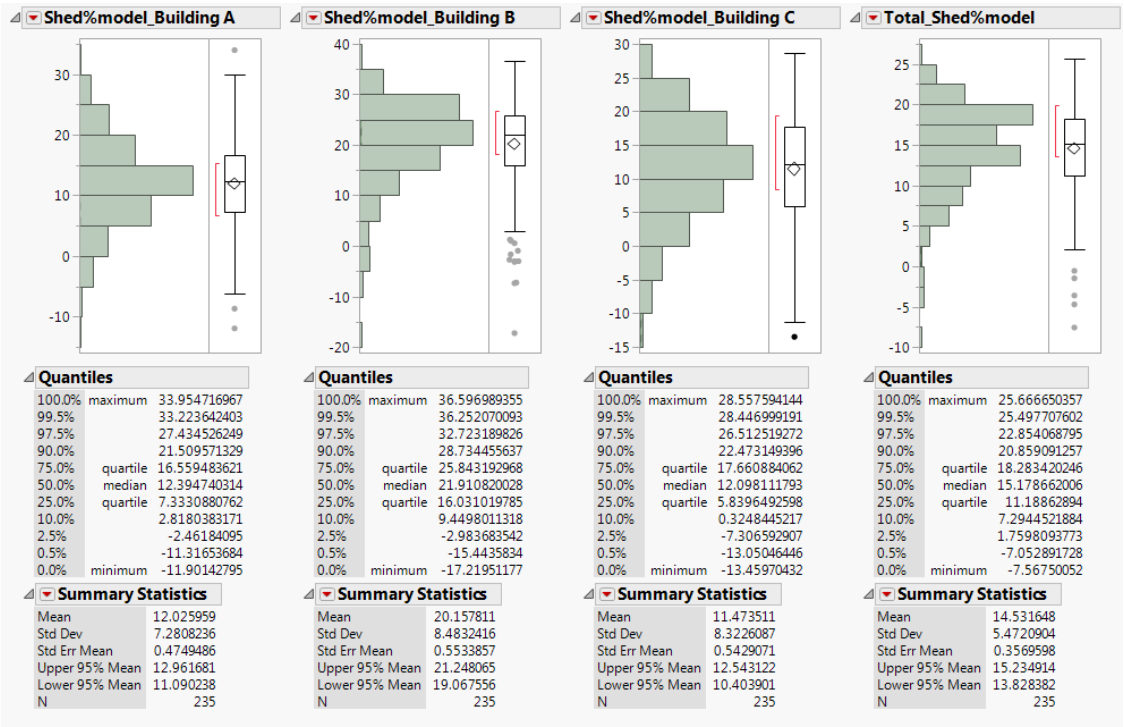


Figure 59 Summary Statistics of the Load Shed

In summary, Hypothesis 1.2 holds because no evidence is found to reject the hypothesis. The relative amount of load shed (%) can be considered as a random variable regardless of time or weather condition. The distribution of each building shows the same shape: unimodal, and maybe bounded or normal distribution. If a true value of the load

shed is constant when the conditions are controlled, then the normal distribution will be the logically right distribution. Naturally, noise variables form the bell shape.

5.1.5 Generalization

The distribution of the load shed for three buildings can be assumed to be the same while the mode can be different per buildings. Even though three buildings have different modes of the distribution, more samples are needed to draw a conclusion regarding building types. Therefore, the normalized DR load shed will be assumed to not be associated with building types, or size for this study. Regardless of the individual building's distribution, the mean value of aggregated DR load shed at the campus level will be close to the normal distribution because it is the sum of random variables if the amount of DR load shed for the buildings has the same distribution.

Central limit theorem (CLT) state as below [88]:

> Central Limit Theorem: If $\bar{X}$ is the mean of a random sample of size n taken from a population with mean μ and finite variance σ^2, then the limiting form of the distribution of
>
> $$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$
>
> as $n \to \infty$, is the standard normal distribution $n(z; 0,1)$

The normal approximation for $\bar{X}$ is generally good when the sample size is larger than 30. If the sample size is smaller than 30, the approximation is only good if the population is symmetric, unimodal, and continuous distribution, i.e., when it is close to the

normal distribution. The classical CLT is limited to when the random variable is independent and identically distributed.

Total cooling load of the campus is the sum of the building's cooling load as expressed in Equation (48).

$$\dot{Q}_{CAMPUS} = \sum_{i=1}^{N} \dot{Q}_{BUILDING_i} = c_P \dot{m}_{CAMPUS} \Delta T_{CAMPUS} \tag{48}$$

If DR is applied to the entire campus building, the modified cooling load can be estimated using $\Delta Load_{DR}$ as a random variable. If $\Delta Load_{DR}$ is the sample from the population having a distribution, say a triangular, then the campus cooling load can be estimated using Equation (49).

$$\dot{Q}_{CAMPUS,DR} = \sum_{i=1}^{N} \dot{Q}_{BUILDING_i,DR} = \sum_{i=1}^{N} (1 - \Delta Load_{DR,i}) \dot{Q}_{BUILDING_i} \tag{49}$$

$$\Delta Load_{DR} \sim triangular(min, mode, max)$$

If all the buildings are identical, it can be reduced as in Equation (50).

$$\begin{aligned} \dot{Q}_{CAMPUS,DR} &= \dot{Q}_{BUILDING} \left(N - \sum_{i=1}^{N} \Delta Load_{DR,i} \right) \\ &= \dot{Q}_{CAMPUS} - \dot{Q}_{BUILDING} \sum_{i=1}^{N} \Delta Load_{DR,i} \end{aligned} \tag{50}$$

Therefore, the load shed of the campus becomes:

$$\Delta Load_{DR,Campus} = \frac{\dot{Q}_{CAMPUS} - \dot{Q}_{CAMPUS,DR}}{\dot{Q}_{CAMPUS}} = \frac{\dot{Q}_{BUILDING} \sum_{i=1}^{N} \Delta Load_{DR,i}}{\dot{Q}_{CAMPUS}} = \frac{\dot{Q}_{BUILDING} \sum_{i=1}^{N} \Delta Load_{DR,i}}{N\dot{Q}_{BUILDING}} = \frac{\sum_{i=1}^{N} \Delta Load_{DR,i}}{\mathrm{N}} \quad (51)$$

According to the Central Limit Theorem, $\Delta Load_{DR,Campus}$ will have a normal distribution regardless of the distribution of $\Delta Load_{DR,i}$.

If there are an infinite number of buildings in the campus, the reduced cooling load by DR when $100 \times p$ % of the campus buildings are activated can be expressed as in Equation (52).

$$\dot{Q}_{CAMPUS,DR} = p\dot{Q}_{CAMPUS}\left(1 - \Delta Load_{DR,i}\right) + (1-p)\dot{Q}_{CAMPUS} \quad (52)$$

Using the statistics presented in Figure 59, the distribution of $\Delta Load_{DR,i}$ can be inferred. If there are enough number of sample buildings to distinguish the building type, different distribution can be inferred for each type of buildings. In this thesis, it will be approximated as a triangular distribution to bound the range of the load shed. Using Equation (52) maybe unrealistic because there are a limited number of buildings in a campus. However, it will provide insight into the estimation of DR benefit for the entire campus using the small sample.

Now that the generalized DR load shed model is developed, the next section will examine the modeling approach for a chiller plant to evaluate the DR benefit on the plant operation.

5.2 Chiller Plant Modeling

This section will develop a predictive chilled water network model. As the plant is the main electricity consumer, the benefit of DR is bigger to the plant than the individual buildings. However, a responsive plant model is required to predict how the system load change affects the plant energy consumption and availability changes. In CHAPTER 4, Modelica Buildings library [46] was selected as the starting point of modeling a district energy system. The basic element models of building HVAC systems such as chillers, cooling tower, and pumps are available from the library. The components can be connected to exchange information and physical variables so that the system model can be built while the physics laws are satisfied. The electrical chiller model is implemented in Building library using DOE-2 model [89], [90]. The model consists of the three second order polynomial curves, which represent the thermal capacity and the electrical efficiency at the full load and part load conditions. The curves and variables are listed in Equation (53)-(56). The plant consists of multiple chillers and is controlled by switch and supply temperature. The chiller model calculates the power consumption corresponding the load using Equation (57). The power consumption of the plant is calculated by summation of the individual chiller's power consumption. The coefficients of the curves can be determined by the standard least square linear regression method using measured data or performance data provided by the manufacturer. More detailed explanation of the model and calibration process will be provided in Appendix A.

$$CAPFT = a_1 + b_1 T_{CHW,S} + c_1 {T_{CHW,S}}^2 + d_1 T_{CW,S} + e_1 {T_{CW,S}}^2 + f_1 T_{CHW,S} T_{CW,S} \quad (53)$$

$$EIRFT = a_2 + b_2 T_{CHW,S} + c_2 {T_{CHW,S}}^2 + d_2 T_{CW,S} + e_2 {T_{CW,S}}^2 + f_2 T_{CHW,S} T_{CW,S} \quad (54)$$

$$EIRFPLR = a_3 + b_3 \times \mathrm{PLR} + c_3 \times \mathrm{PLR}^2 \quad (55)$$

$$PLR = \frac{Q}{Q_{ref} \times CAPFT(T_{CHW,S}, T_{CW,S})} \quad (56)$$

$$P_{predicted,i} = P_{ref,i} \times CAPFT_i \times EIRFT_i \times EIRFPLR_i \quad (57)$$

Where, CAPFT: the capacity function of temperatures

EIRFT: the energy input ratio function of temperatures

EIRFPLR: the energy input ration function of part load ratio

$T_{CHW,S}$: Chilled water supply temperature

$T_{CW,S}$: Condenser water supply temperature

Q: Cooling load

Q_{ref}: Capacity at the reference condition

PLR: Part load ratio

$a_i, b_i, c_i, d_i, e_i, and\ f_i$: Coefficients

$P_{ref,i}$: Reference power of ith chiller

Once the individual chiller model's curves are determined, they can be arranged with the piping network, buildings, and other components to represent a district energy system as presented in Figure 60. The configuration of the system considered in this thesis is the fixed primary flow and the variable secondary flow system. It is a simplified version of the real chilled water system model. The model illustrated here has all the basic elements of a district energy system. Even though the campus load is modeled as one lumped load of thermal volume that emits heat, it drives the pump by calculating the required cooling load from the cooling load and delta T model, which are the functions of weather and time

variables. More information of building the model is available in [91]. The structure of a model is hierarchical because Modelica is an object-oriented modeling language. The component models are treated as black boxes that have the inputs and the outputs. Therefore, it is replaceable. For example, if the plant is expanded, then the plant model can be replaced with the upgraded model.

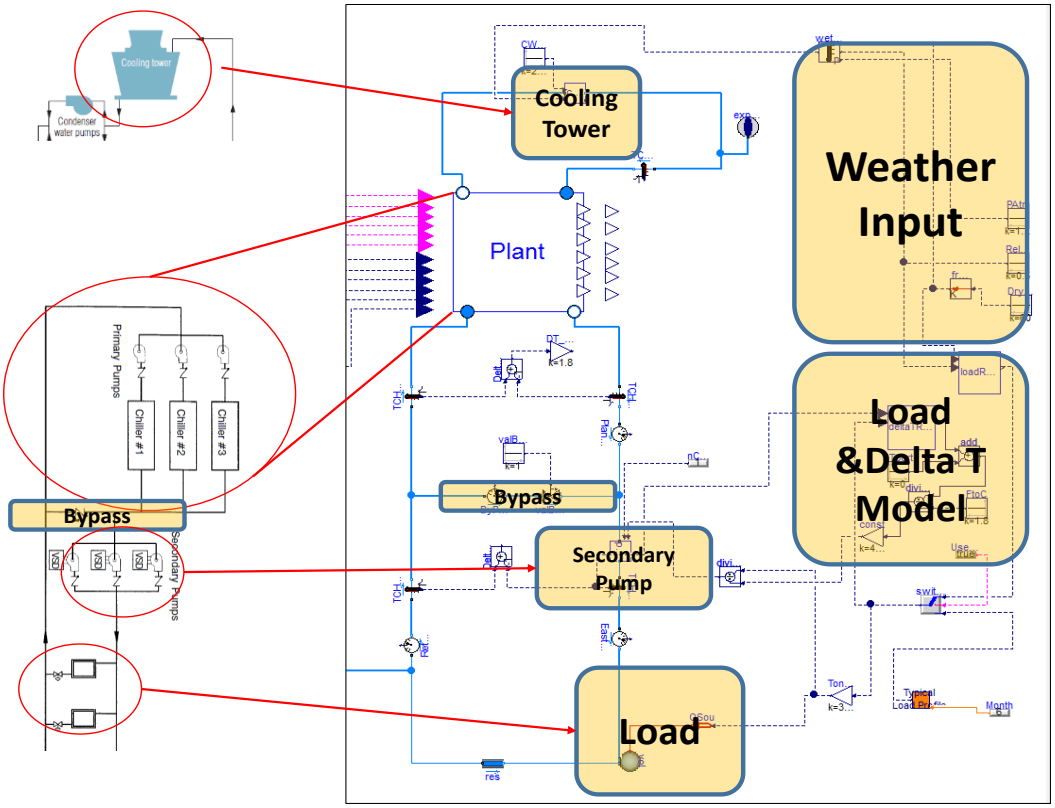


Figure 60 A Chilled Water System Model with One Lumped Load Representing a Campus District Energy [91]

The plant model has seven chiller components as presented in Figure 61, which is shown as a block in Figure 60. In the same way, each block of the chillers in the plant model consists of a chiller, a condenser pump, and a chilled water pump as illustrated in

Figure 62. The specific information of the chillers, the used data, and the validation results are provided in Appendix A.

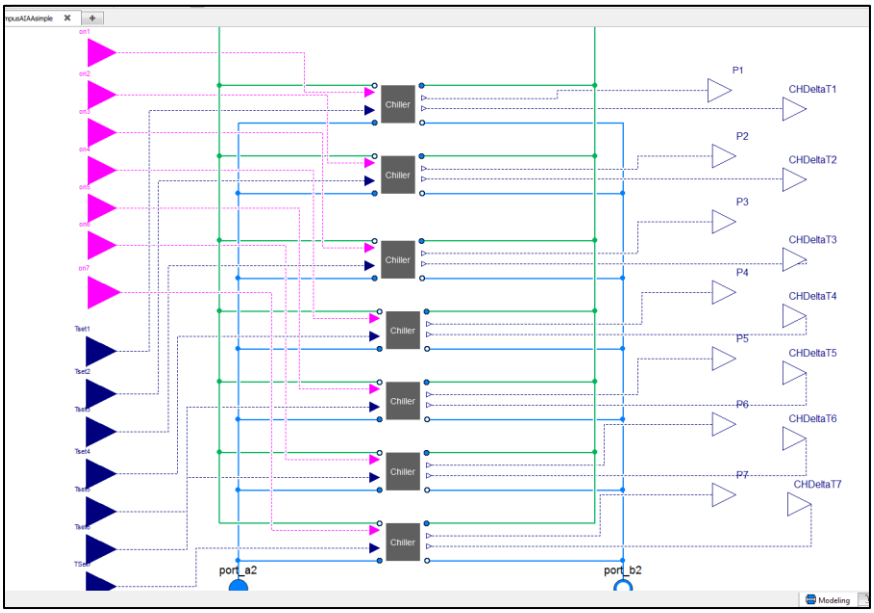


Figure 61 Chiller Plant Component Model

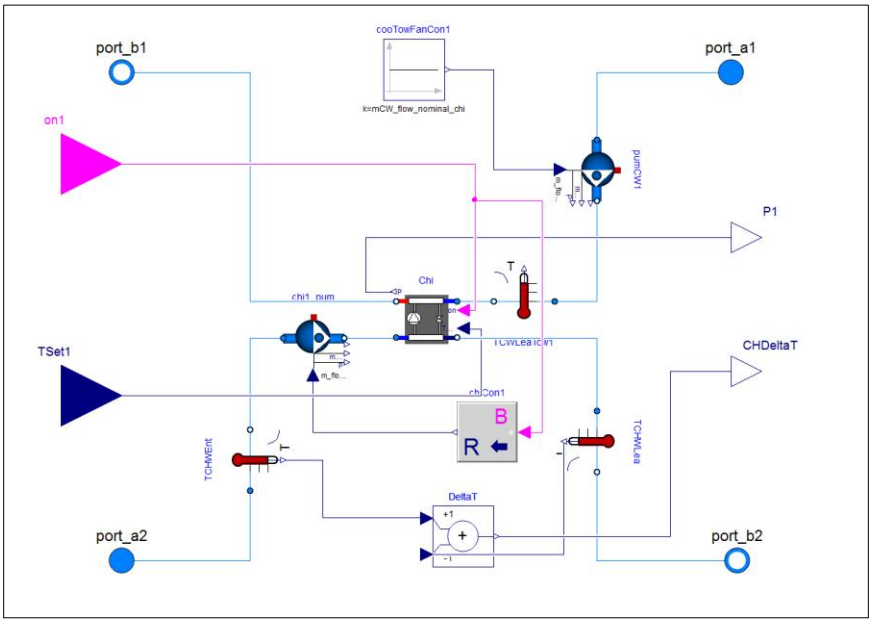


Figure 62 Chiller Component Model

The campus load model can be further divided into multiple buildings. In that case, each building should have its own model. The expanded model, which has two load blocks, is presented in Figure 63. A building is typically represented as a cooling coil, a heat exchanger, in a district energy system. Instead of the cooling coil model, the data-driven cooling load model and delta-T model are used for simplification and avoiding a control system development. Delta-T model is developed as a function of the cooling load, supply temperature, weather, and time variables so that it can calculate the required cooling flow rate to emulate a hydraulic system. ANN model is used to create the campus cooling load model based on the measured data at the plant level. The load component model consists of a mixing volume, which produces heat, and a pump. The cooling load and delta T from the ANN model are the input variables, and the output is the flow rate and the return temperature.

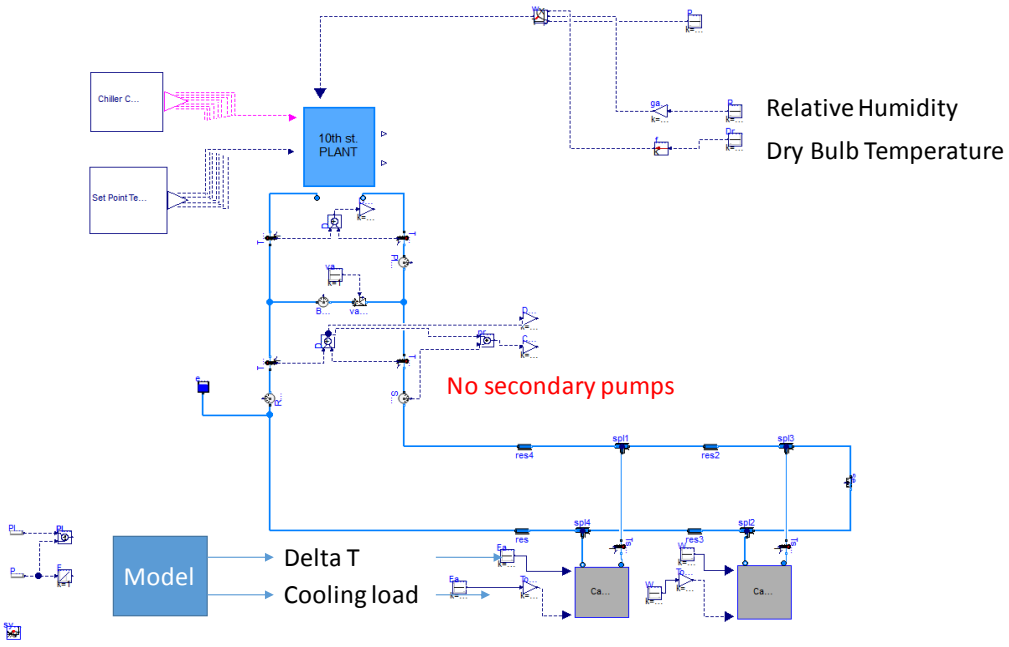


Figure 63 A Chilled Water System Model with Two Load Blocks

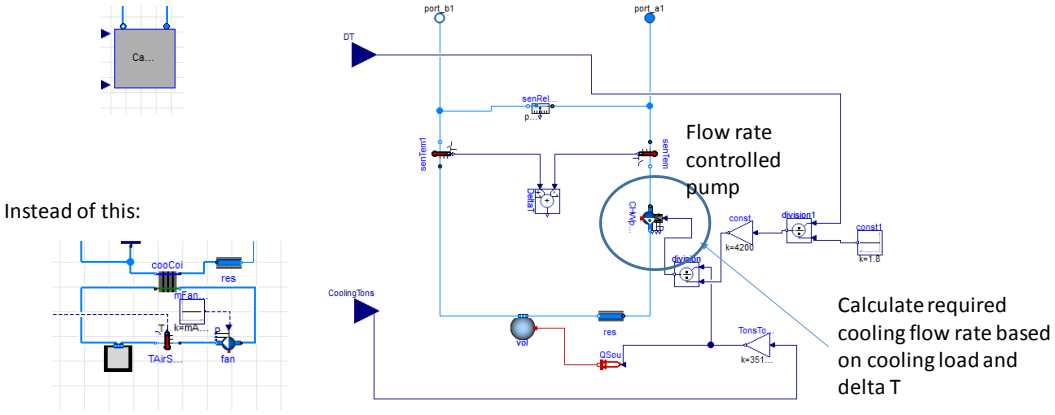


Figure 64 Building Load Model

5.2.1 Plant Model Validation

The chilled water network system of Georgia Tech is implemented using the Buildings library. The plant model presented in Figure 63 is validated using the measured data by simulating for one week in summer 2013. The plant had six chillers and total 12,500 tons of thermal capacity back then. The cooling load profile measured during the week is presented in Figure 65. For the same cooling load and the chiller staging settings (switch and supply temperature setpoint), the simulated result is well following the measured power as shown in Figure 66. Even though some mismatch between the measured value and the simulation result is observed, the overall trend is well captured. The model error metrics listed in Table 7 are well under the acceptance limit. Hence, the model is good enough to use in predicting the performance of the plant under the various conditions. The cooling load and the power consumption show the similar trends.

Table 7 Plant Model Error Statistics

Model Error	Value (%)
MBE	2.44
CV-RMSE	5.51

Figure 65 Cooling Load Profile during Validation Week

Figure 66 Plant Power Consumption Simulation vs. Measured

5.2.2 Experiment 2: Plant Operation Approximation

The chiller staging information was manually entered into the validation experiment. However, actual control is needed to simulate under the new condition which has not been observed and stored. It is necessary to examine the DR capability. As the chiller model has the control inputs and the network configuration can be represented with the pipes and pumps, one can implement a chiller dispatch control on top of this model. However, it is too complicated and requires an accurate representation of the hydraulic system. As discussed earlier, developing a control system is beyond the scope. Instead, Research Question 2 seeks a more efficient way of representing the dynamic response of the chiller plant to the cooling load to evaluate the capability of DR to affect the operation of the plant.

Research Question 2:

How can the benefits of DR for a chiller plant be assessed without implementing a complex control logic?

Investigation of the system and operating characteristics conducted in Section 3.2 leads to Hypothesis 2.

Hypothesis 2:

If the optimal control design can be approximated as a set of the best combinations per loading conditions to minimize power consumptions, then it can predict changes of operational status of a chiller plant without a sophisticated control logic.

To test this hypothesis, an experiment is devised to generate data to create a design space as illustrated in Figure 33. The purpose of this experiment is to explore the plant operational

design space and discover the performance frontier of the system. The experimental process is presented in Figure 67. It demonstrates how plant model uses the data-driven demand model to calculate the corresponding power consumption at the chiller plant. To represent a thermal demand on the building side and hydraulic characteristics, the aggregated campus cooling load ($\dot{Q}, Tons$) and temperature differential ($\Delta T, F$) are modeled as functions of weather and time. The required output of the simulation model is the optimal chiller staging with respect to the cooling load and the resulting power consumption of the chillers. The resulting performance map and the optimal setting lookup table will support the hypothesis.

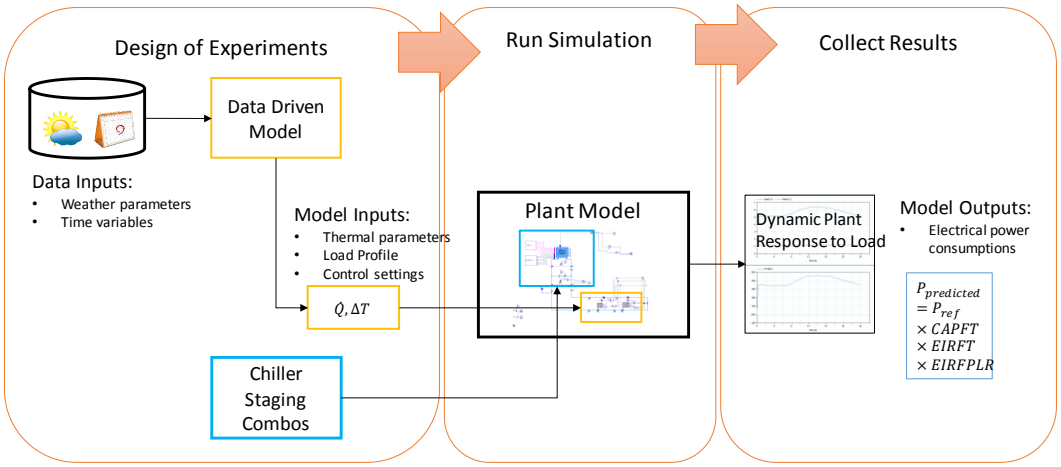


Figure 67 Plant Operating Performance Experiment Process

The model is validated using the measured data so that the prediction of the model is credible. Using the model, the power consumptions regarding all possible operating conditions and the chiller settings can be generated. The chiller plant model has been expanded to have seven chillers reflecting the real system's expansion. The basic information of the chillers is listed in Table 8.

Table 8 Chiller Basic Information

	Thermal Capacity (Tons)	Flow Rate Capacity (GPM)	Reference Power (kW)
Chiller 1	1500	3600	921
Chiller 2	1500	3600	921
Chiller 3	2000	4000	1293
Chiller 4	2000	4000	1293
Chiller 5	2250	4000	1221
Chiller 6	3000	6000	2026
Chiller 7	3000	6000	2026

Design of experiment table is produced as the input variables of the demand model and chiller staging combinations. The demand model has five input parameters of time and weather as explained earlier. The operating condition space is illustrated in Figure 68. It covers the summer weather of Atlanta presented in Figure 69. The time variables are assigned to the points to produce the input array for the ANN demand model. The input DOE table for the plant simulation model is generated by combining the demand model's inputs and outputs, and the chiller combinations. The full factorial number of chiller combination is 128 (2^7). To reduce the number of inputs, the chiller combinations that cannot be the demand are filtered. Only chiller switch option (on/off) is considered, and the supply temperature is fixed as 40 °F in this experiment. More extensive design space exploration is conducted in [91]. Total 5170 combinations of the operating conditions and chiller staging are used in this experiment.

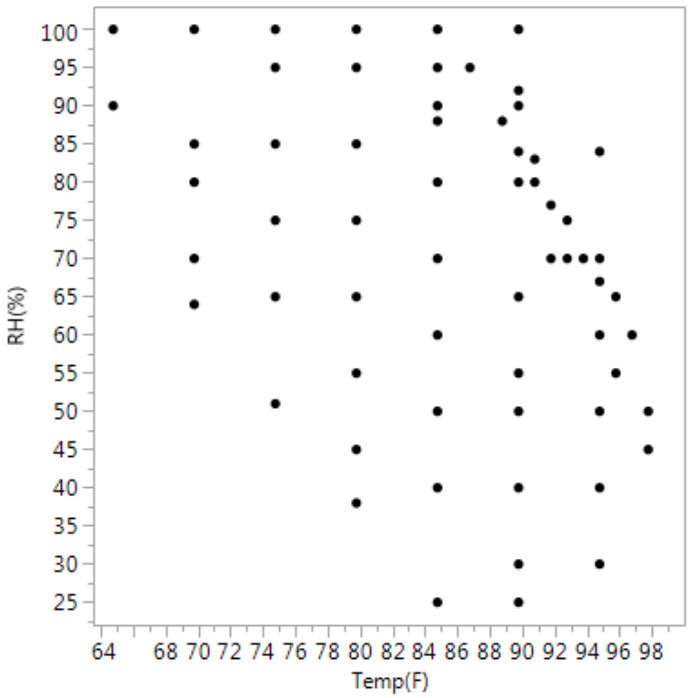


Figure 68 Operating Condition DOE Space

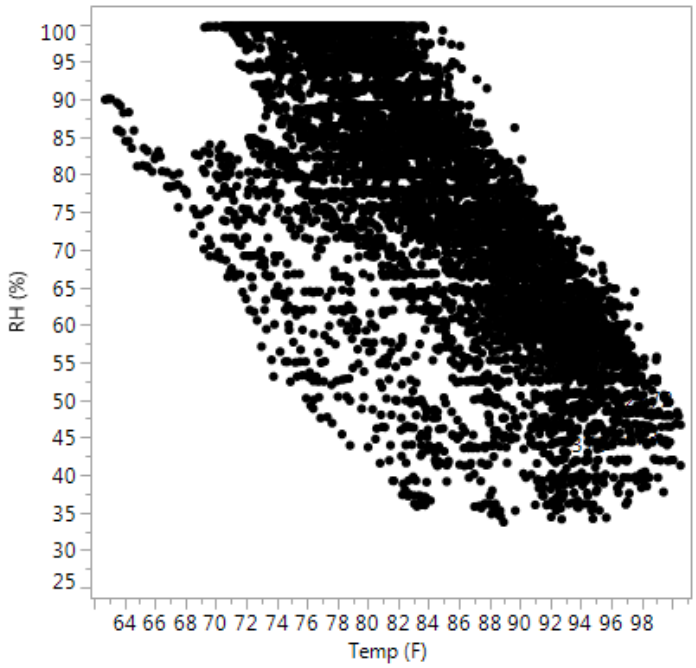


Figure 69 Summer Weather of Atlanta 2016

Once the DOE input table is generated, the automation Python code runs the simulation for all the cases and collects the interested output parameters. For this experiment, the plant power consumption is the output parameter. The resultant performance map is presented in Figure 70. Multiple alternative chiller combinations can serve the same load. Data sets are divided into smaller sets with an increment of 250 tons in order to obtain the optimal chiller staging setting for each load level that requires the minimal power use in meeting the same campus cooling demand.

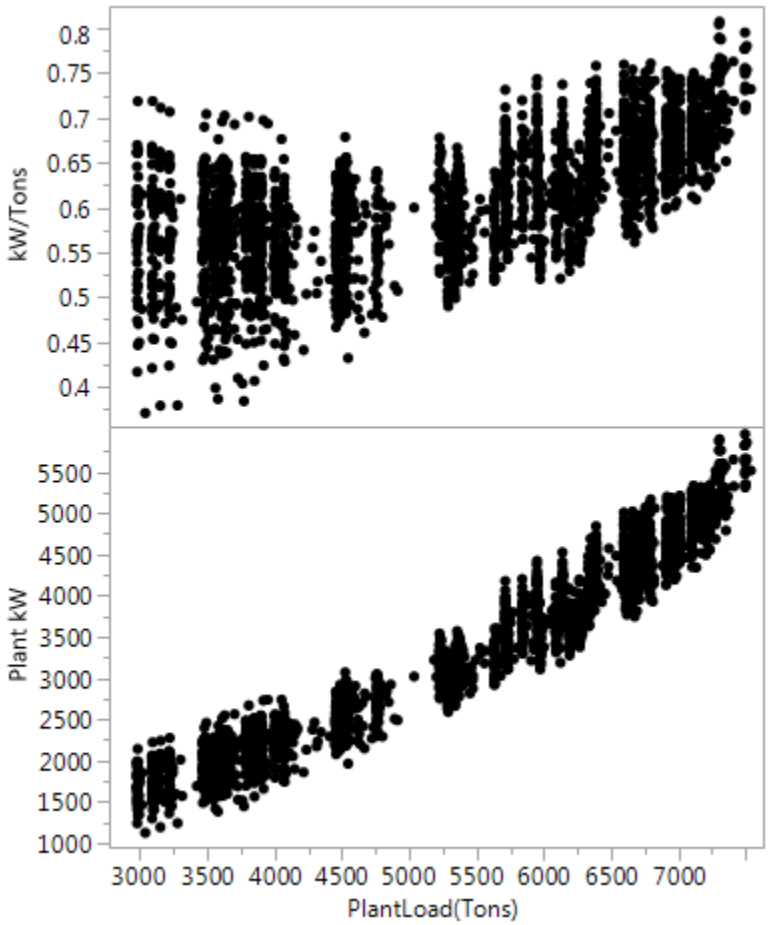


Figure 70 Plant Performance Map Generated by the Simulation

Figure 71 shows the optimal points. Each of the optimal points is linked to the chiller staging option. The optimal chiller settings are illustrated in Figure 72 as a cell plot. The blue colored cells indicate which chillers should be on at the specified cooling load level. It can be stored as a lookup table to emulate an optimal chiller staging control. The optimal sets can be approximated as a polynomial function. Figure 73 shows that either of second order or third order polynomial curve fits well the power consumption vs. the cooling load. The approximation of chiller staging scheme and power consumption now can be treated as a surrogate model of the plant operation. It enables fast Monte Carlo Simulation to evaluate the DR impact without running the plant model for each case.

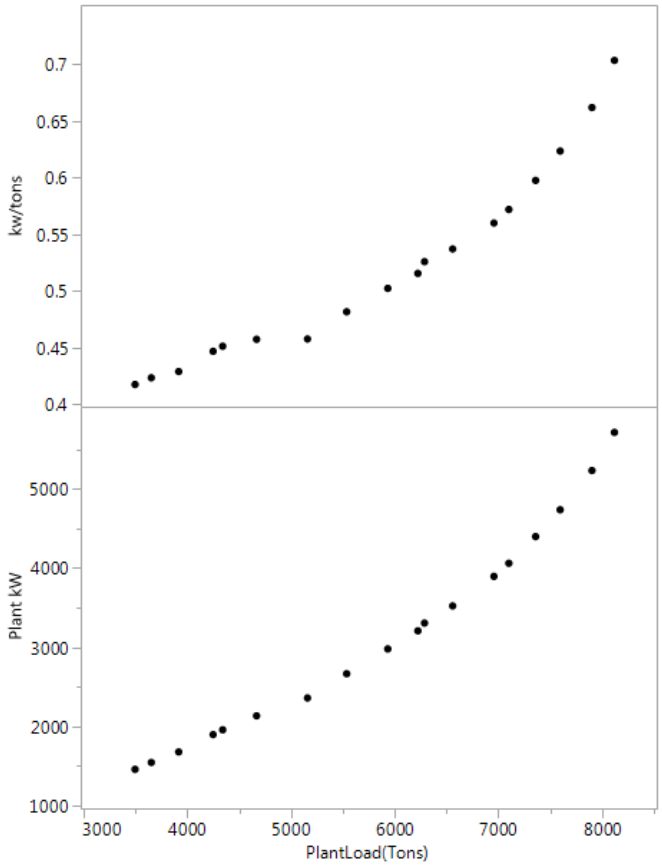


Figure 71 The Optimal Operating Points

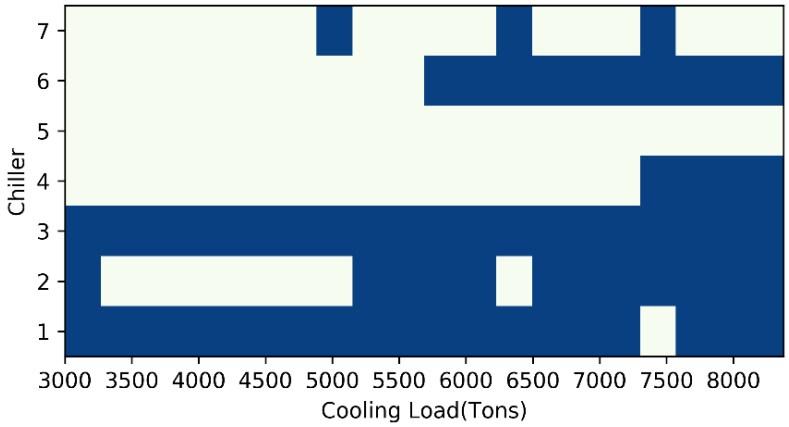


Figure 72 Optimal chiller setting for each cooling load level.

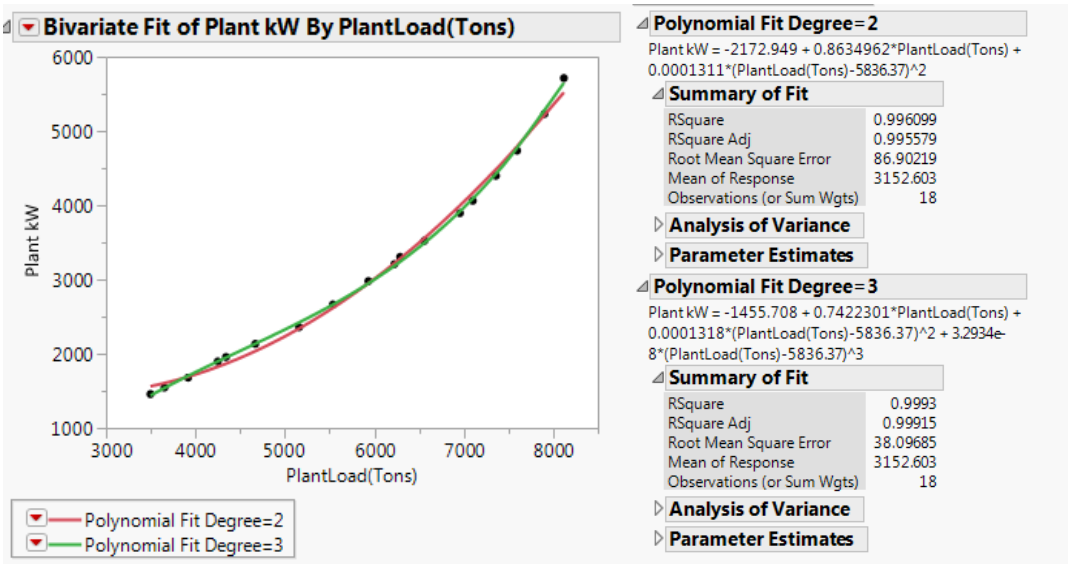


Figure 73 Plant Performance Curve Fit

An experiment is performed to see if the plant surrogate model can capture the dynamics between the load and the plant. Instead of running the original plant model, a computation code is developed to calculate the power consumption and select the optimal chiller setting in response to a dynamic load as presented in Figure 74. A typical summer day load profile is generated to see the dynamics. Figure 75 shows how the surrogate model

responses to the cooling load and the power consumption calculated using the polynomial curve. Both approximation models perform well. The optimal setting has excess capacity than the actual cooling requirement. It is because of the cooling flow rate requirement at low delta T condition as discussed in Section 3.3.2. The experimental results support Hypothesis 2. The surrogate model will be used to evaluate the DR impact on the plant instead of running the Modelica chilled water system model.

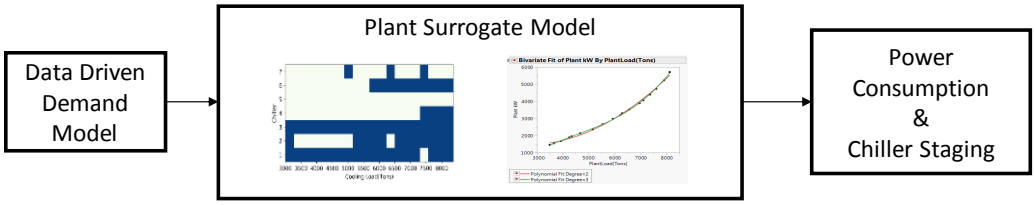


Figure 74 Simulation Procedure Using the Plant Surrogate Model

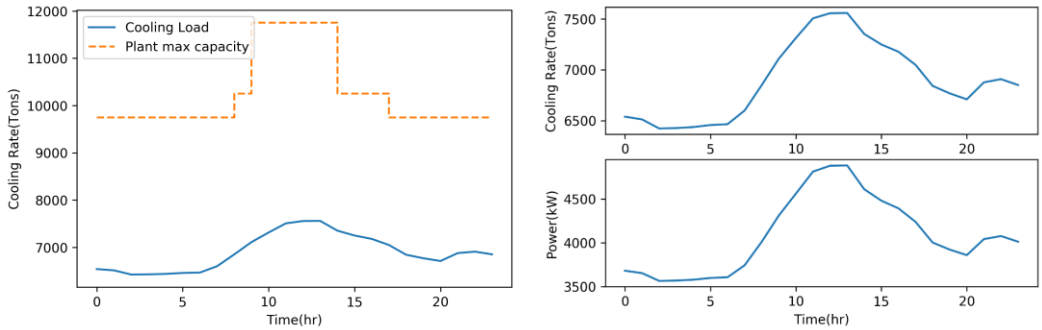


Figure 75 the Optimal Chiller Operation and Power Consumption

5.3 Reliability Assessment

This section will describe how to develop the simulation based reliability assessment model and perform verification of the simulation model. The system reliability analysis requires the demand and the plant models. Research Question 3 prompted to investigate proper reliability metrics to evaluate and compare DR and the redundancy.

Research Question 3:

How can the performance of DR as a contingency plan for a district chilled water system be measured and fairly compared to the deterministic redundancy?

In Section 3.3.1, the probability of failure at the peak and system adequacy metrics are determined as the evaluation metrics. To calculate those metrics, a probabilistic simulation is essential.

5.3.1 Model Development

Reliability assessment module is implemented in Python environment. The chiller's state-duration model is developed using SimPy version 3.0.10 and Python version 2.7. A chiller is modeled as two state repairable system with the constant failure rate (λ) and repair rate (μ). Once a chiller failure occurs, it is put into repair immediately and restored to as good as new condition after repair. The time to repair and the time between failures are random variables following exponential distributions. The probability density function of an exponential distribution is as follows.

$$f(t;\lambda) = \begin{cases} \lambda e^{-\lambda t} & t \geq 0 \\ 0 & t < 0 \end{cases} \tag{58}$$

The mean time between failure (MTBF) and the mean time to repair (MTTR) are as follows for exponentially distributed random variables.

$$\begin{aligned} \text{MTBF} &= \frac{1}{\lambda} \\ \text{MTTR} &= \frac{1}{\mu} \end{aligned} \tag{59}$$

If the failure rate and repair rate (or MTBF and MTTR) are historically known, the repairable system state duration can be simulated using the discrete event simulation model. The chiller model yields the stochastic sequence of state until the specified time as presented in Figure 76.

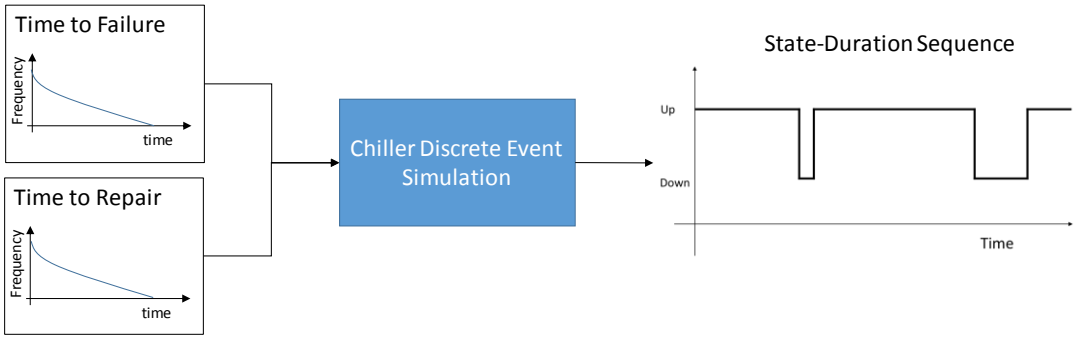


Figure 76 Stochastic Chiller State Model

A typical chiller plant is designed to have multiple chillers in parallel. Assuming every chiller is independent and has the same failure rate and repair rate, the total available capacity of the chiller plant can be obtained through the parallel simulation. Once the simulation is done, all the chillers' state-duration sequences will be multiplied by each chiller's capacity to calculate the available system capacity. As discussed previously, both of the thermal and the mass balance should be satisfied. The thermal capacity is calculated using Equation (60) while the mass capacity is calculated using Equation (61). Given the annual cooling load profile, the reliability metrics will be calculated using those as counting the hours of unfulfilled demand over the year.

$$Available\ Capacity\ \dot{Q}_{plant}\ at\ t = \sum_{i}^{N} Q\_Cap_i \times x_i(t) \quad (60)$$

$$Available\ Capacity\ \dot{m}_{plant}\ at\ t = \sum_{i}^{N} M\_Cap_i \times x_i(t) \quad (61)$$

Q_Cap_i: Design capacity of chiller i (Tons)

M_Cap_i: the Design mass flow rate of chiller i (GPM)

$x_i(t)$: Chiller i's state at t $x_i \in \{0,1\}$

Figure 77 demonstrates how the reliability assessment simulation module is constructed. The reliability metric calculation function requires the demand profile and the plant availability as inputs. Given the annual cooling demand profile, the required chilled water flow rate is estimated using the delta T model, which is presented in Section 3.3.2. Based on the assumption of the failure rate and the repair rate, the individual chiller's state vectors are generated from the simulation. The system capacity is calculated using the state vectors and the chillers' capacity of the thermal and mass availability. Once the system's available capacity vectors are generated, the unsatisfied hours and energy (mass) are calculated by detecting the discrepancy between the demand requirement and the capacity availability. This process is repeated for the specified number of simulations to get the distribution of the reliability metrics. The average values result in the evaluation metrics.

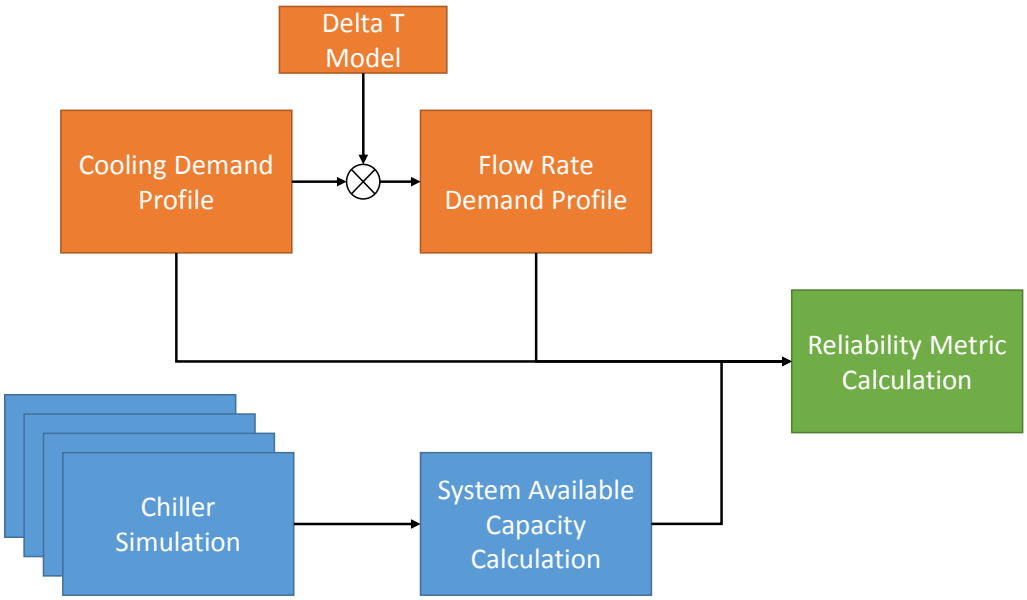


Figure 77 Reliability Metric Calculation Algorithm

5.3.2 Verification of the simulation model

The developed model is verified using data listed in Table 9. The test system has total 7chillers (2 small, 3 medium, and 2 large) so the total system capacity is 15,000 tons. The reliability statistics of the centrifugal chillers are found from the literature [92]. This data tells that MTBF is more than 20 years. It seems to mean the end of life type failure. In this experiment, a much smaller number is used as listed in Table 9 to consider the repairable failure. It can be updated if better information is found. It is assumed that the individual chiller's failure is independents of each other.

Category/Class	*Item Number*	*Reliability*	*Inherent Availability*	*Operational Availability*	*Unit Years*	*Failures*	*Failure Rate (Failures/Year)*	*MTBF*	*MTTR*	*MTTM*	*MDT*	*MTBM*	*Hrdt/ Year*
Charger, Battery	Item: E11-000	0.992621004	0.999999577	0.999986472	270.0	2	0.00741	1182768	0.50	0.0000	0.133	9816	0.1185
Chiller	Item: H10-000	0.888515818	0.999829779	0.997620632	2021.9	239	0.11820	74109.90	12.62	1.0881	1.164	489	20.843
Absorption	Item: H10-100	0.841986658	0.999769437	0.995132437	430.3	74	0.17199	50932.86	11.74	1.0000	0.653	134	42.639
Centrifugal	Item: H10-200	0.955142622	0.999923928	0.997604888	544.7	25	0.04589	190872.1	14.52	5.2247	5.333	2227	20.981
600 -1000 Tons.	Item: H10-220	0.955142622	0.999923928	0.997604888	544.7	25	0.04589	190872.1	14.52	5.0000	5.333	2227	20.981

Figure 78 Chiller Failure Statistics [92]

Table 9 Test System Assumptions

Chiller Option	Cooling Capacity (Tons)	Flow Rate (GPM)	MTBF (Hours)	MTTR (Hours)
Small	1,500	3,600	20,000	24
Medium	2,000	4,000	20,000	24
Large	3,000	6,000	20,000	24

Given the system capacity is 15,000 tons, the annual load profile is generated to have the peak load of 15,000 tons using the normalized load profile proposed in Section 4.4. The annual profiles of the cooling load, the flow rate, and delta T are presented in Figure 79.

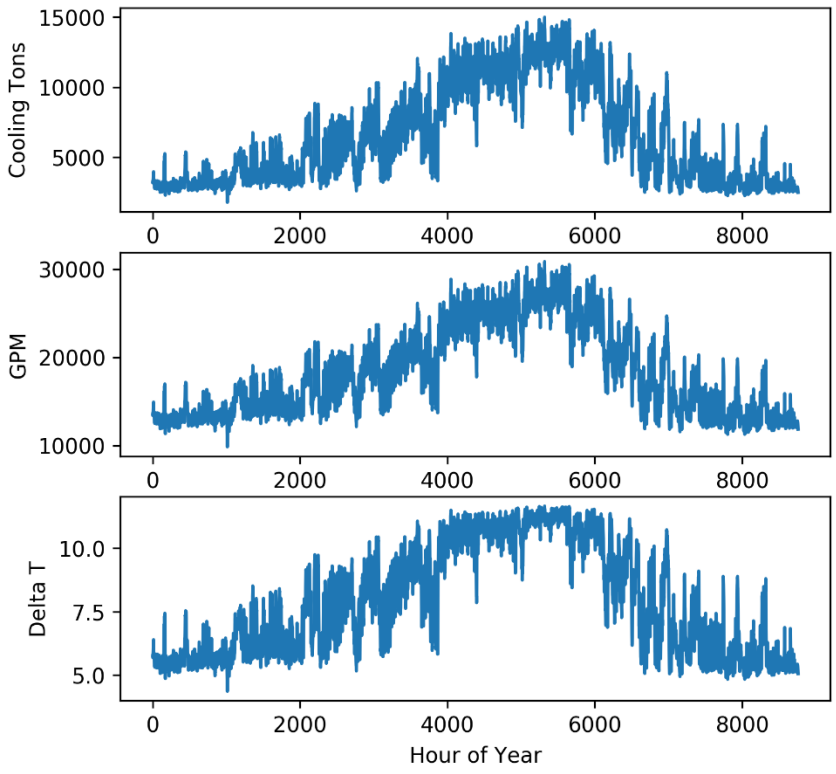


Figure 79 Annual Demand Profile

Figure 80 and Figure 81 show how the system capacity is added up from the individual chillers. Each of the seven chillers has different state sequences and the resulting capacity. Stochasticity of the failure event and the duration of unavailability are well captured from the simulation. One of the simulation cases is presented in Figure 82. In this case, there are times when the available capacity is lower than the demand. Figure 83 shows the zoomed in view of the period of the failure. As expected, the period of the GPM requirement is not met is longer than that of the cooling tons. Monte Carlo Simulation is performed to get the reliability metric with 1000 number of simulations. The LOLE for cooling load is 4.28 hours/year and for the GPM is 6.04 hours/year.

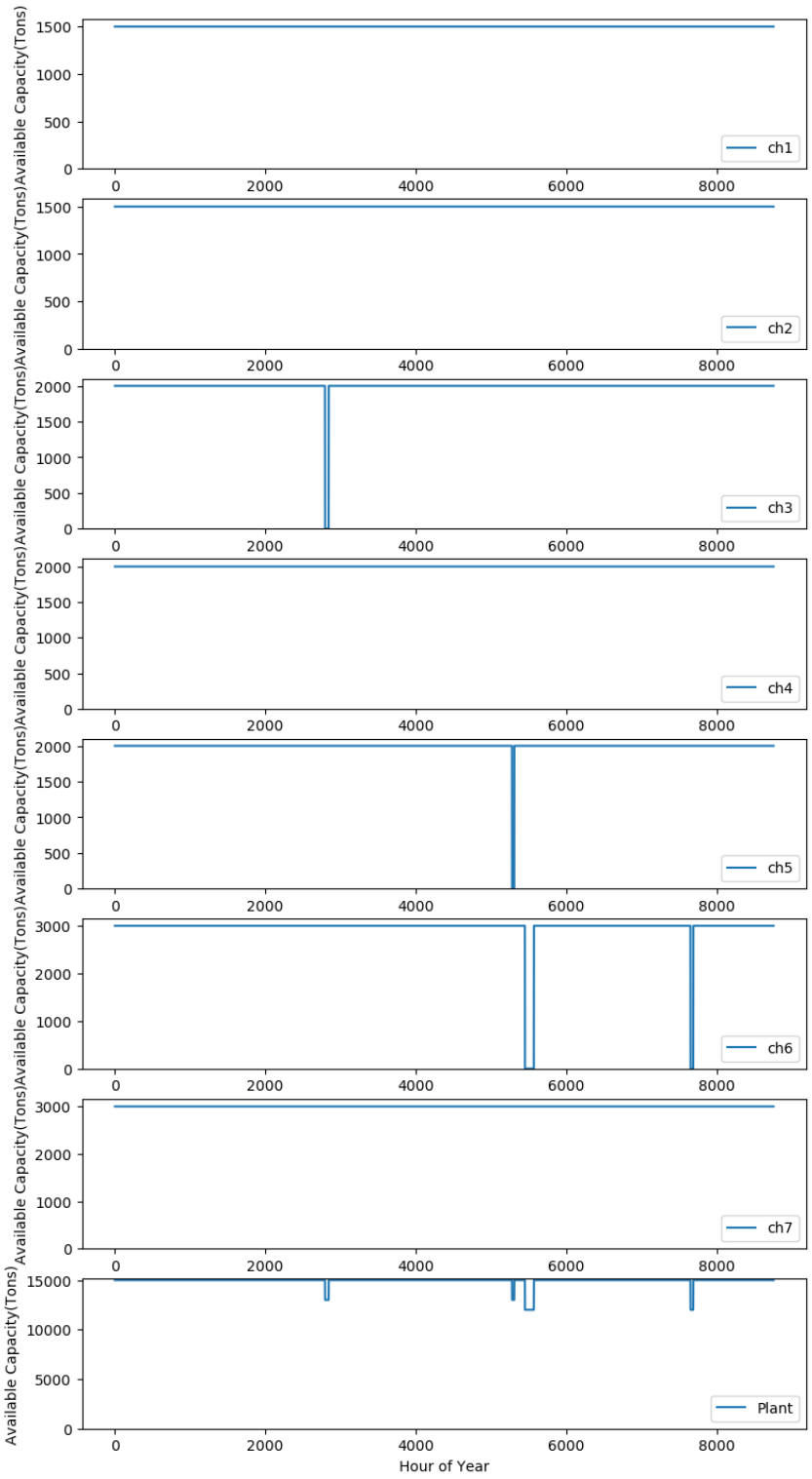


Figure 80 Plant Availability Simulation Sample for Cooling Capacity

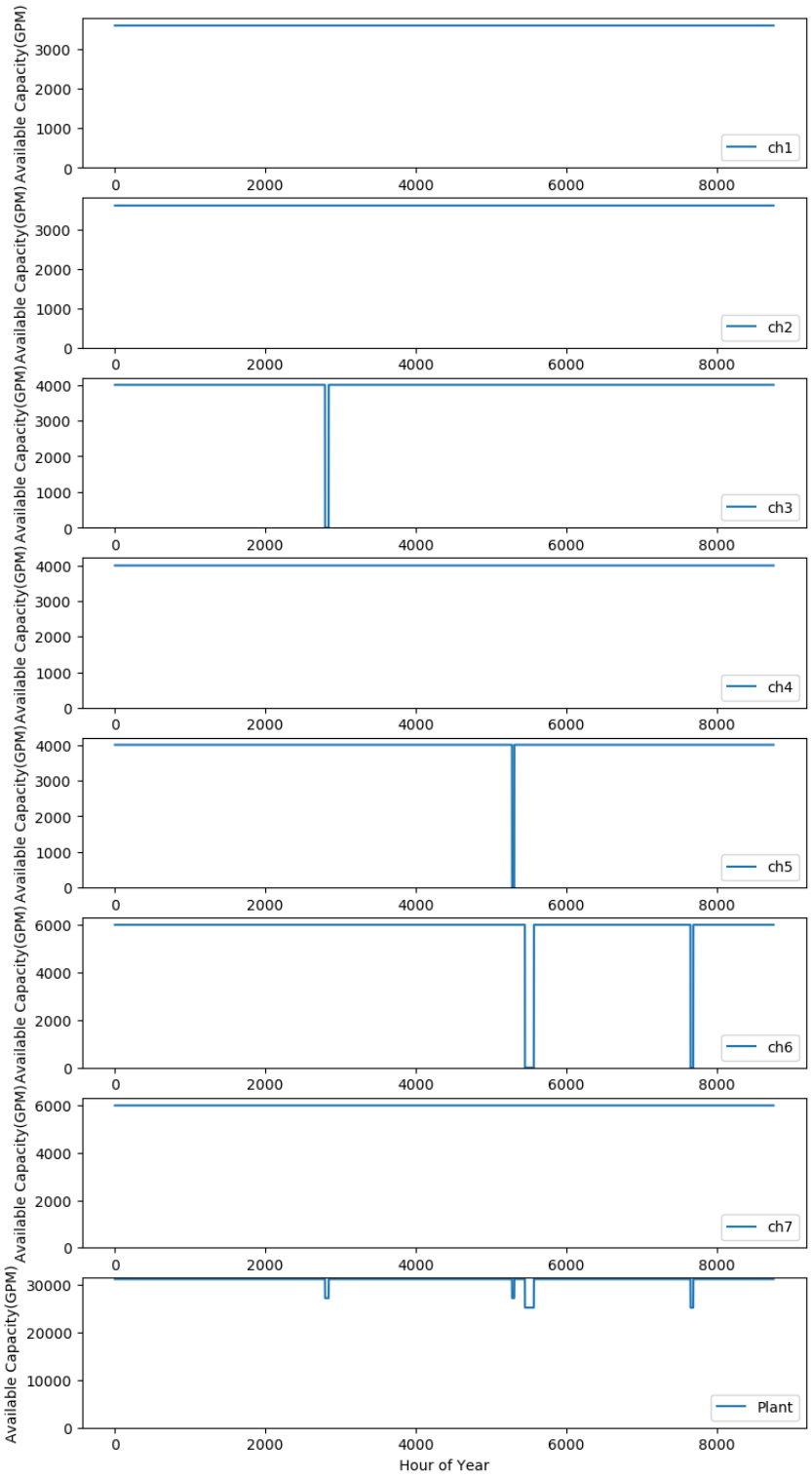


Figure 81 Plant Availability Simulation Sample for Mass Capacity

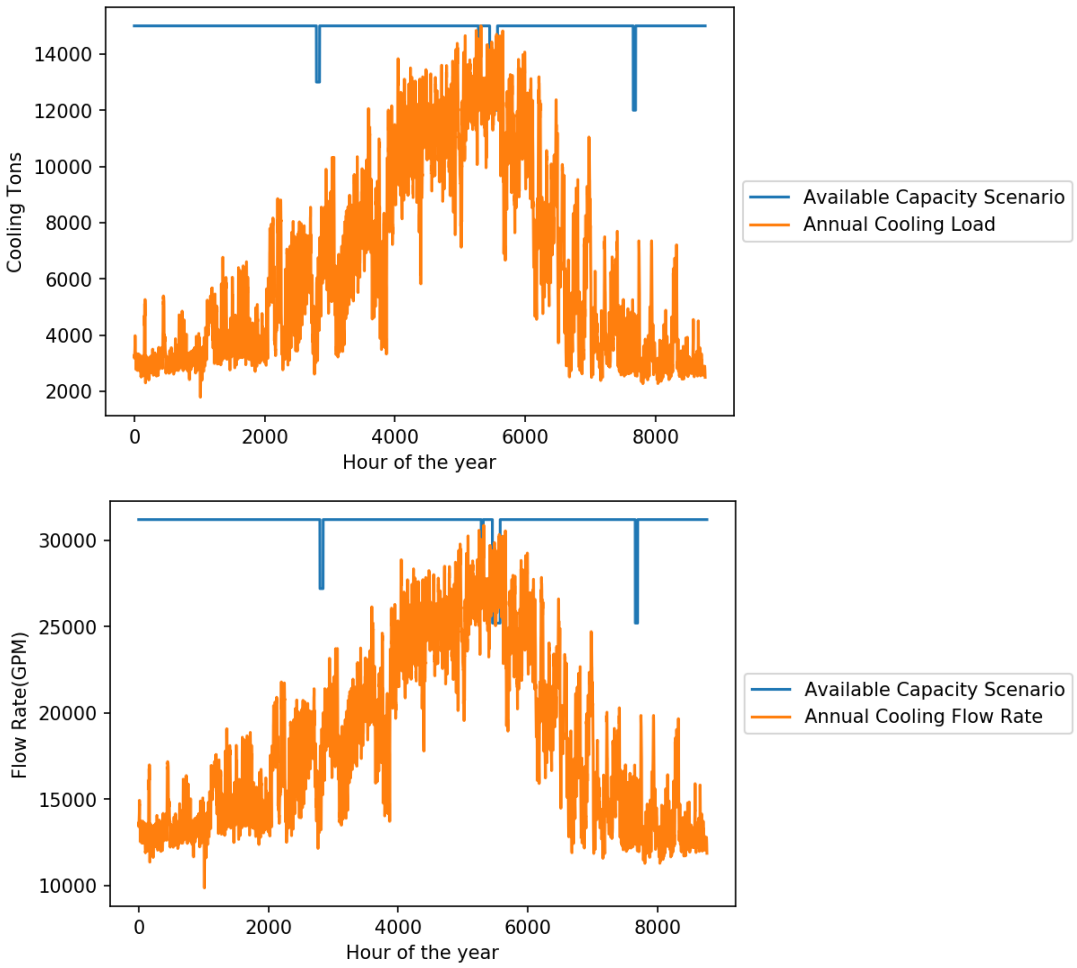


Figure 82 a Simulation Sample that has a Contingency Event

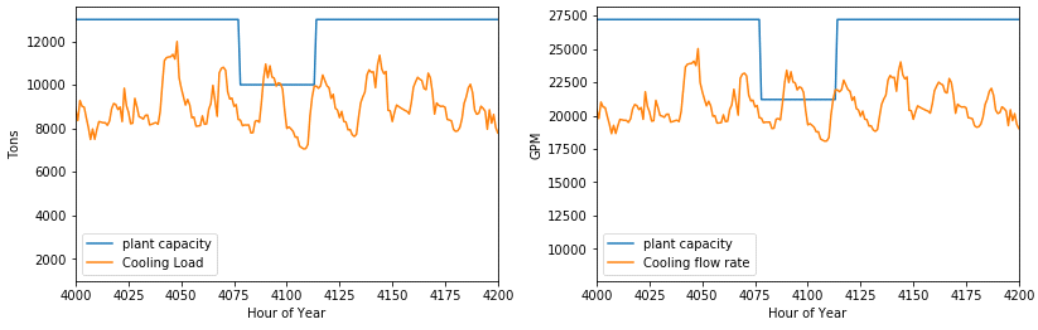


Figure 83 an Example Case When the System Fail to Meet the Demand

5.4 Capacity Expansion Plan Optimization Model

The MILP formulation of the capacity expansion planning problem expressed as in Equation (35)-(39) is implemented using Python 2.7 and Gurobi 7.5.1. This formulation is feasible under the two assumptions addressed in Section 3.4. These assumptions were made to simplify the cost function. Operating cost (assumption1) and demand response cost (assumption2) can be ignored when the problem satisfies specific conditions. If the HVAC systems of the buildings are controlled by automated building management systems, then the assumption 2 is valid because no additional hardware cost is required to enable DR. The assumption 1 will be tested using the plant surrogate model developed in Section 5.2.

5.4.1 Experiment 3: Operating Cost Test

According to the chiller plant design guideline [25], a chiller's life cycle cost consists of the first cost (capital cost for purchase and installation), the utility cost (fuel, water, etc.), and the maintenance cost. The capital cost can be formulated as a linear integer function, but the operating cost is complicated. The operating cost is a nonlinear function of cooling load, the chiller plant configuration, and control logics. This makes it very challenging to formulate as an explicit function. For the power grid system, where the renewable energy and various types of fuel and technologies are alternatives, the operating cost can be an important factor that affects the results. However, different chillers are all in the same conversion technology category that transforms the electrical energy to the thermal energy. The operating cost is mainly dependent on the electricity price and demand profile if the plant is optimally operated with the fixed configuration. Then the operating cost can be simply approximated as a function of demand using the plant surrogate model.

If two different plant configurations do not differ the operating cost, then it can be ignored in capacity expansion planning optimization because the capital cost is much higher than the operating cost. To validate the assumption, two different system models are created by changing the number of chillers. Having the basic information of chillers listed in Table 8, two test systems are made as described in Table 10. The plant model presented in Figure 61 is easy to modify the number of chillers. The plant 1 has 6 chillers (chiller 1 to 6), and the plant 2 has an additional 3,000 tons chiller to the plant 1 (chiller 1 to 7).

Table 10 Test Systems

	Plant 1	Plant 2
Number of chillers	6	7

The performance maps are created for the test systems by the simulation described in Figure 67 to approximate the operating cost function. The optimal chiller combinations extracted from DOE results are presented in Figure 84. These optimal sets are fitted to the second order polynomial curve of the efficiency function. The polynomial fits are presented in Figure 85 and compared in Figure 86. The plant 2 is more efficient at high load while the plant 1 is more efficient at lower load. These curve fits are used to calculate the annual operating cost.

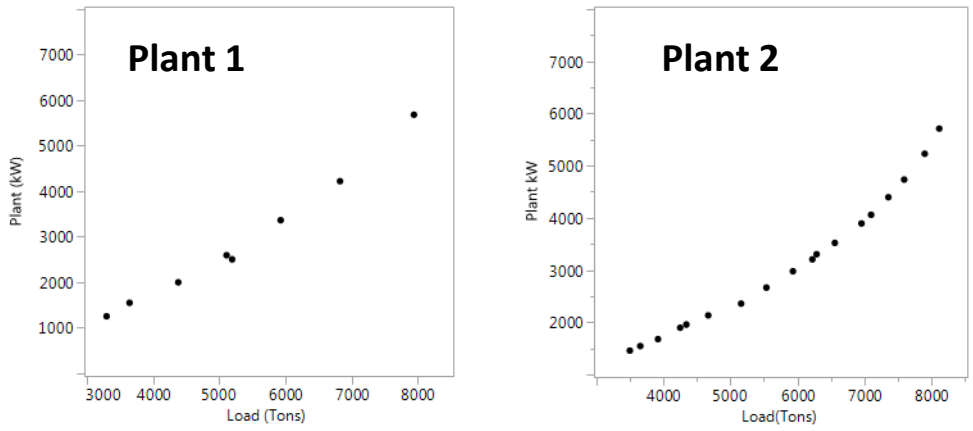


Figure 84 Optimal Chiller Combinations

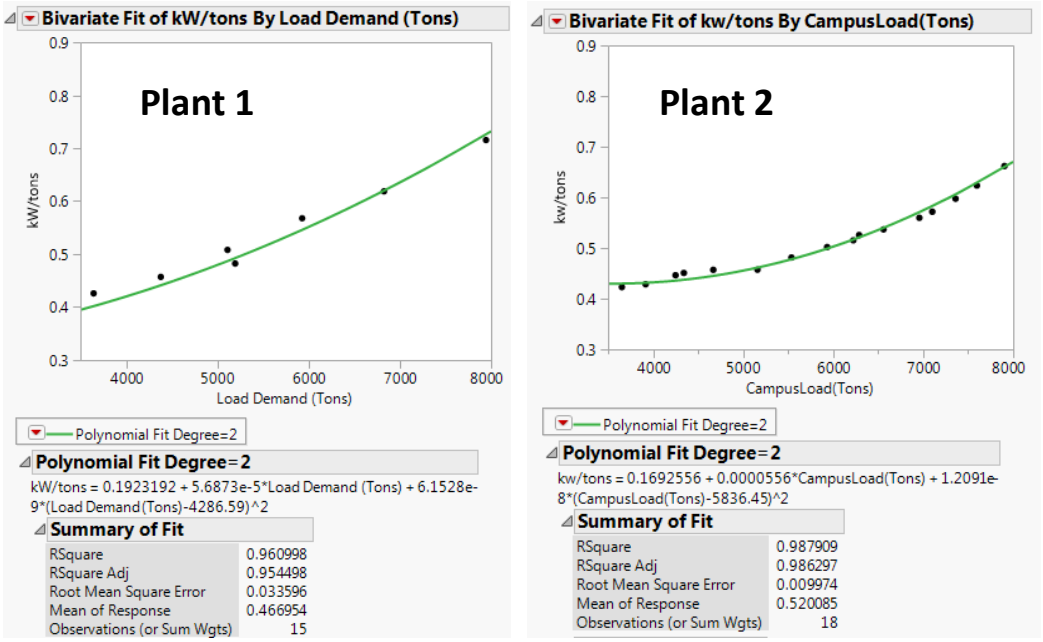


Figure 85 Plant Performance Surrogate Models

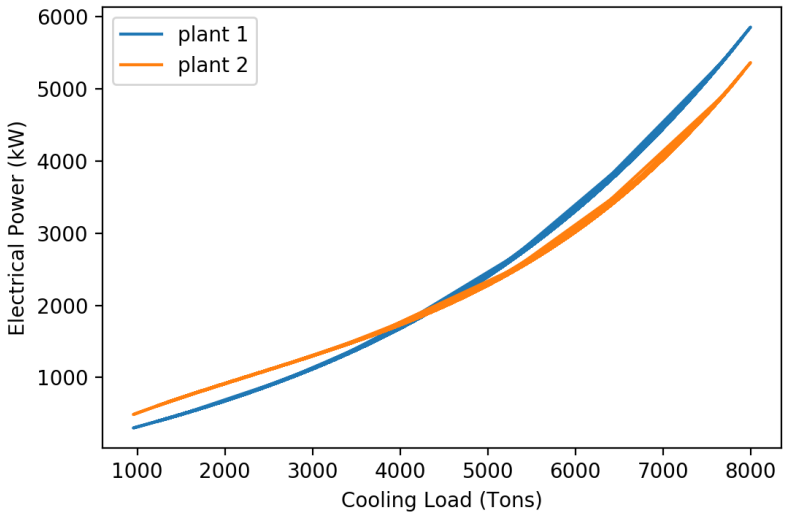


Figure 86 Approximate Performance Curve Comparison

The chiller plant operating cost can be calculated using Equation (63). The efficiency function $kW/tons$ is approximated as the second order polynomial function of cooling load $\dot{Q}$ with the optimal control assumption in this case. Therefore, the annual operating cost can be calculated by integrating the function over the year.

$$\int \dot{Q} \times \frac{kW}{tons}\left(\dot{Q}, chiller\ control\right) \times \frac{\$}{kW}(t)dt \tag{62}$$

The annual cooling load profile for calculation is the scaled-up cooling load, which was described in Section 4.4. Figure 87 shows the resultant annual power consumption profiles of two plants. The difference between the annual operating cost and the price assumption are depicted in the graph. Assuming the average electricity price to be 4 Cents per kW, the cost difference is 18,954 Dollars per year. It is the negligible difference compared to a chiller's capital cost, which is much more than a million Dollars. The

smallest chiller option is about 4M Dollars, so the operating cost difference is less than 1% of the capital cost. Therefore, the assumption 1 is valid for the problem of this research.

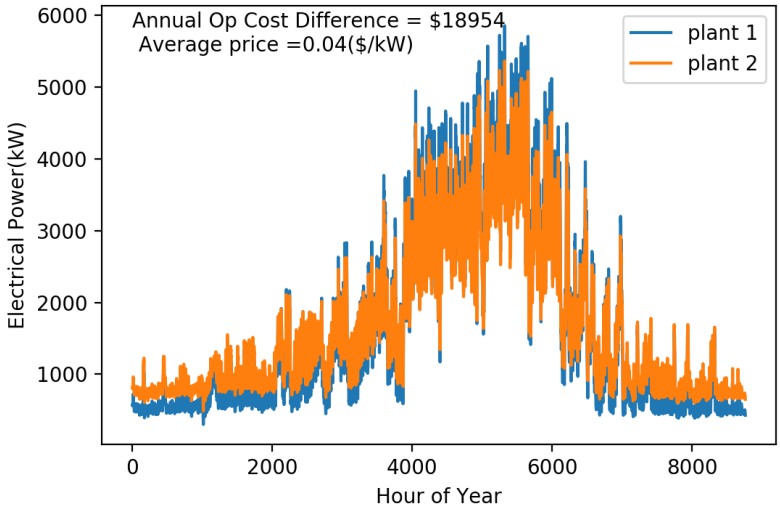


Figure 87 Annual Electrical Power and Operating Cost Comparison

CHAPTER 6 IMPLEMENTATION

Chapter 5 developed the constituent models that are the foundations of the proposed methodology. This chapter will demonstrate how the proposed methodology can be used for decision making support of a capacity expansion planning problem. Some of data and models used for the example problem are already described in the previous chapter. For those overlapped data, the detailed explanation will be skipped. By demonstrating the methodology, Hypothesis 4 will be tested.

6.1 Example Problem

Based on the partial chilled network of Georgia Tech campus, a hypothetical system is created to demonstrate the proposed framework for forecasting the DR impact on the capacity expansion planning. One of the chiller plants in the campus chilled water network is used as the baseline system. Real data and the configurations of the system are used to create the models. The plant currently has seven chillers, and two more spots are available to install new chillers for meeting the future demand growth. The future demand growth is assumed to be driven by new buildings in the next 15 years. The campus master plan specifies the size of new buildings and when they will be built. The peak load value is approximated based on the planned gross square footage of those buildings. Their load shapes are unknown and cannot be simulated with a building model in the planning stage because a design of the building is not created yet. To meet the increasing demand, a new chiller of the right size should be installed before the total load would exceed the capacity limit of the existing plant. With consideration of the lead time and budget planning, the capacity expansion plan for the entire planning horizon should be made. There are a few

buildings participating in demand response program for the local electricity utility company. A system manager would like to know if adopting demand response as contingency plan instead of redundant chillers can save the total cost over the planning horizon while ensuring the desired reliability level.

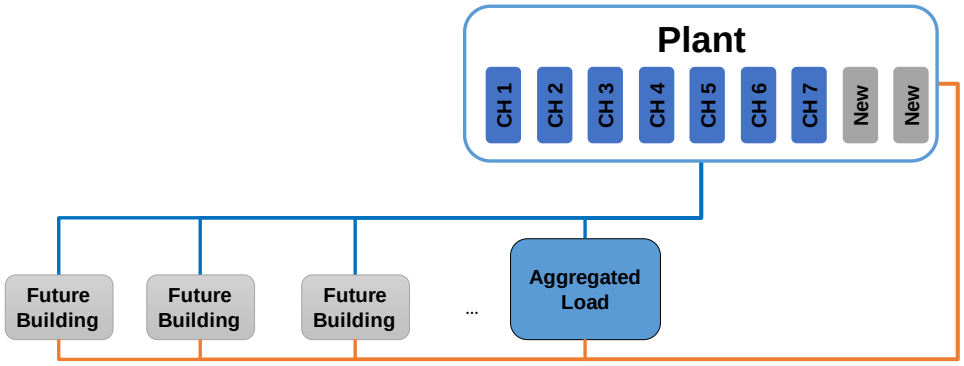


Figure 88 Example Chilled Water System to Be Expanded

To reduce complexity, some assumptions are made.

- Chillers of the same size have the same performance.
- New chiller options are the same as the existing chillers
- Aging of chillers are not considered
- The master plan is fixed. The year of new buildings to be connected to the network is deterministic variable.
- Only new buildings and climate change contribute to demand growth.

6.2 Step 1: Data Preparation

At this step, data described in Table 3 will be collected and prepared to be used for the modeling and analysis tasks in the next steps.

6.2.1 Database Description

The measured data was collected from the GT energy system databases. There are two types of sensing and measurement data. One is from a database for electrical and thermal meter data. Almost every building and energy plants are equipped with a metering system, and the meter data is recorded at the 15 minute intervals. Historical meter data is available up to past 5 years from the database. Modeling for the building or plant level cooling demand and the chiller plant use data from this database. The other type of data is available from the building management system. Most of the modern commercial buildings are controlled and maintained by a Building Management System (BMS) such as Metasys [93]. Floor or room level HVAC control inputs and thousands of sensor measurement data including building level meter data are available from the BMS. Unfortunately, this type of data is not stored in a database for longer period. They are only available for the past day or up to a few weeks depending on parameters. The DR event log was obtained from this system. For detailed data of the DR event log, see Appendix B.

Georgia Tech has its own weather station on campus. From this weather station, basic climate data can be obtained including dry bulb temperature and relative humidity. Historical weather data was recorded at the 15 minute intervals as same as the meter data. The on-site weather station data is used in this research. In case there is no on-site weather station, local historical weather data is available from a publicly accessible database such as National Centers for Environmental Information [94]. Figure 89 shows the simplified chilled water network configuration to explain where data are measured. Temperature and flow rate sensors measure the quantity at the supply and return sides of the buildings, the chillers, and the plant. Cooling load is calculated using the temperature and flow rate measurement.

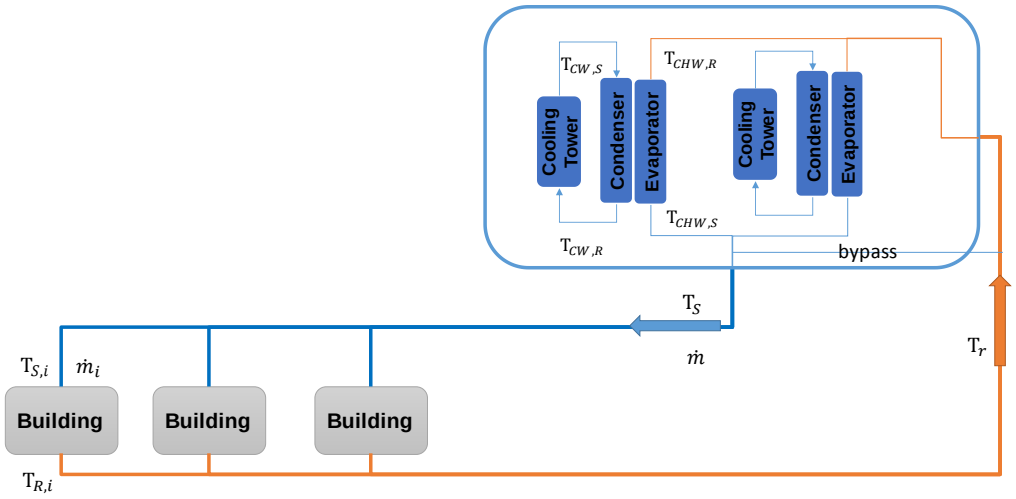


Figure 89 Chilled Water Data Measuring Point

6.2.2 Data Cleaning

Raw data contains lots of corrupted data including missing data, sensor errors, and physical or numerical outliers. Those data should be filtered before use. Figure 90 shows how those types of data were filtered. The multivariate scatter plot is used to inspect data set for the same entity (i.e., a building or a chiller). First, data containing invalid numbers were excluded. If the same values were recorded in a row, those data were considered as sensing or recording errors. Outliers from the cluster of data points were also detected and excluded. This process was repeated for all data to create a clean data set. This is an important step for the data driven modeling. If data used for model development contains outliers, the resultant model does not represent the normal behavior of the system.

All measured data are time-stamped. Care should be taken with the timestamp. The time zone used in the database is UTC (Coordinated Universal Time). It should be converted to the local time to be properly analyzed. Georgia Tech is located in the Eastern Time Zone; therefore, the timestamp was converted to EST (Eastern Standard Time, 5 hours behind UTC) or EDT (Eastern Daylight Time, 4 hours behind UTC).

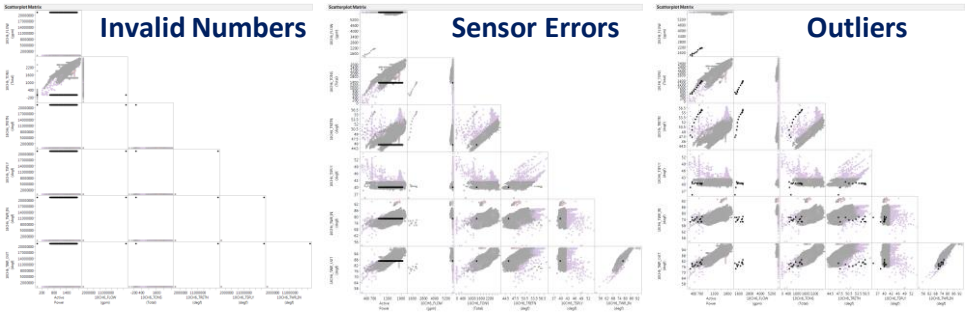


Figure 90 Filtered Data Types

6.2.3 Descriptive Data

The current chiller plant has 7 chillers with 4 different sizes as described in Table 8. The hypothetical campus is assumed to have 7 chillers with 3 different sizes, and the same size chillers are identical. The new chiller options are the same with the current system. The capital cost of purchasing and installation cost is obtained from a credible source. The reliability data is assumed to be applicable to all the chillers. The basic chiller information is listed in Table 11.

Table 11 Chiller Information

	Cooling Capacity (Tons)	Flow Rate (GPM)	Capital Cost	MTBF (Hours)	MTTR (Hours)
Large	3,000	3,600	$5M	20,000	24
Medium	2,000	4,000	$4.5M	20,000	24
Small	1,500	6,000	$4M	20,000	24

The future building construction plan was created based on the Georgia Tech's Master Plan with a slight modification. The planning horizon is 15 years from 2018. The

estimated peak load and the expected year to be built of the future buildings are listed in Table 12.

Table 12 New Building Plan

	Estimated Peak Load (Tons)	Expected Year (years from now)
Building 1	1,000	2019 (1)
Building 2	800	2022 (4)
Building 3	900	2026 (8)
Building 4	800	2028(10)

Data described in this section will be used in the following steps. They will be recalled when it is needed.

6.3 Step 2: Demand Response Modeling

In Section 5.1, the proposed data driven DR model has been developed and validated using data. The generalized DR model is induced as a random variable by hypothesis testing. It will be applied to the campus load to evaluate the peak load reduction capability of DR. The DR method used in this research is the global HVAC setpoint temperature change. When the same DR is applied to the entire campus building, how much of cooling load can be reduced is the question. The real chilled water plant modeled in this example supplies cooling energy to more than 25 buildings including large research buildings and dormitory buildings. Even though they are all different in size and demand characteristics, it is assumed that the campus buildings identical and infinite so the DR application can be increased proportionally for the sake of simplicity in the analysis as in Equation (63). Having the individual building model, it can be decomposed into Equation (64). Consideration of the individual building will be left as the future work.

$$\dot{Q}_{CAMPUS,DR} = p\dot{Q}_{CAMPUS}(1 - \Delta Load_{DR}) + (1 - p)\dot{Q}_{CAMPUS} \quad (63)$$

$$\dot{Q}_{CAMPUS,DR} = \sum_{i=1}^{N} \dot{Q}_{Building,i}\left(1 - x_i \Delta Load_{DR,i}\right) \quad (64)$$

where, x_i: decision variable for building i $\begin{cases} 1 & if\ DR\ is\ applied \\ 0 & if\ DR\ is\ not\ applied \end{cases}$

Using the statistics of past DR data, DR load shed is modeled as a random variable. The distribution functions of the DR parameters are listed in Table 13. The starting and ending hours are considered random because they cannot be controlled by the community. The relative amount of load shed for the campus level ($\Delta Load_{DR}$) can be either a triangular distribution or a normal distribution. It was proven that the total load shed distribution will be close to the normal distribution in Section 5.1.5. Triangular distribution is a reasonable choice when the true distribution is unknown and having small samples. Being bounded and unimodal, it is enough to represent the expected value and uncertainty. The distribution function can be inferred from the total load shed of the three sample buildings. The comparison among the sample data, normal and triangular distribution is presented in Figure 91. The histogram of the approximated distribution was plotted using 100,000 samples from Monte Carlo Simulation.

Table 13 Demand Response Parameters

Parameter Name	Distribution
DR Starting hour (t_{Start})	Uniform(13,16)
DR Ending hour (t_{End})	Uniform(17,20)
Amount of load shed ($\Delta Load_{DR}$)	Triangular(0,0.15,0.26) Normal(0.15,0.05)

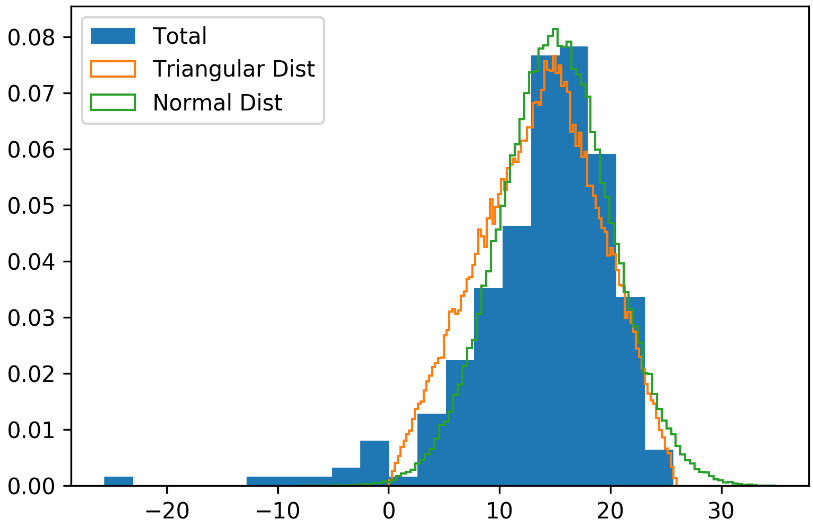


Figure 91 Total Load Shed Distribution vs. Approximation

Having the parametric DR model, it can be tested on the campus cooling load to evaluate how much load can be shed by DR. Using of Equation (63), one can generate a time sequential load profile by simulating for a specified period. To account for the DR activity, an indicator variable a_t is introduced. The modified cooling load at time t is simulated using Equation (65). $\Delta Load_{DR}(t)$ shall be sampled at each time step.

$$\dot{Q}_{CAMPUS,DR}(t) = p\dot{Q}_{CAMPUS}(t)(1 - a_t\Delta Load_{DR}(t)) + (1-p)\dot{Q}_{CAMPUS}(t) \qquad (65)$$

where, a_t: indicator variable at time t $\begin{cases} 1 & if\ DR\ is\ active \\ 0 & if\ DR\ is\ not\ active \end{cases}$

Given the cooling load profile for a day, the DR load can be predicted by implementing Equation (65) with the Monte Carlo Simulation as in the algorithm below. Once the simulation is done, n number of stochastic load profiles are generated. Having this statistical data, it enables to predict the load reduction by DR and capture the stochasticity of demand. Thus, it can provide the probability of achieving the demand reduction goal to evaluate the benefit and the risk of expanding DR to the other buildings.

Algorithm 1: Pseudocode of DR Load Forecasting

Input: the original load profile $\mathrm{Q}_{CAMPUS} = (\dot{\mathrm{Q}}_{CAMPUS}(1), \cdots, \dot{\mathrm{Q}}_{CAMPUS}(24))$

Output: $\mathrm{Q}_{CAMPUS,DR} = (\mathrm{Q}_{CAMPUS,DR,1}, \cdots, \mathrm{Q}_{CAMPUS,DR,n})$: n number of load profiles

Determine n: the number of simulation

Determine p: the fraction of campus buildings to apply DR

Initialize $\mathrm{Q}_{CAMPUS,DR}$

for i = 1 to n do

  Sample t_{Start} and t_{End}

  for t = 1 to 24 do

    Sample $\Delta Load_{DR}$

    If $t_{Start} < \mathrm{t} < t_{End}$ **then**

      $a_t = 1$

      calculate $\dot{\mathrm{Q}}_{CAMPUS,\mathrm{DR,i}}(\mathrm{t})$ using Equation (65)

    Else

      $\dot{\mathrm{Q}}_{CAMPUS,\mathrm{DR,i}}(t) = \dot{\mathrm{Q}}_{CAMPUS}(t)$

  end for

  Update $\mathrm{Q}_{CAMPUS,DR,i} = (\dot{\mathrm{Q}}_{CAMPUS,\mathrm{DR,i}}(1), \cdots, \dot{\mathrm{Q}}_{CAMPUS,\mathrm{DR,i}}(24))$

end for

A typical summer cooling load profile seen by the chilled water plant is generated using the historical weather profile and the ANN model for the campus load and presented in Figure 92. Having this load profile, the reduced load profile is generated to evaluate the peak load reduction capability of DR and the overall energy savings. Normal distribution is used for sampling $\Delta Load_{DR}$ and Monte Carlo Simulation is performed for 10,000 runs. The samples of the reduced load are illustrated in Figure 93. The resultant load shapes exhibit the load reduction during the DR active period. The different amount of load shed and the duration of DR for each sample prove the stochasticity of DR model. As discussed in the preliminary data analysis in Section 3.1, DR cannot reduce the peak load effectively

because the peak time of the campus cooling load and the utilities' price signal are not well synchronized. Figure 94 shows the expected load shape when DR is expanded to other buildings. Increasing the percentage of DR buildings, the significant amount of load reduction can be achieved with the slight peak load reduction. Note that this is the averaged result from the 10,000 cases. Now that the reduced cooling load is characterized by the parametric model, it will be transferred to the next step to evaluate the impact of DR on the plant operation.

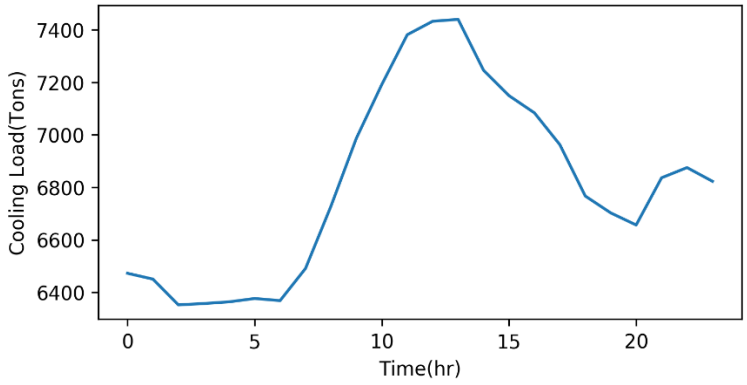


Figure 92 a Typical Summer Day Cooling Load

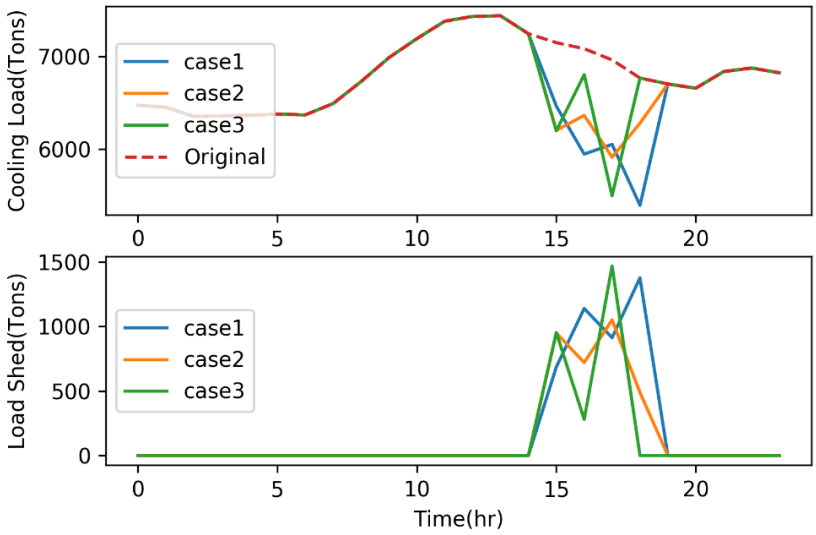


Figure 93 Samples of Cooling Load Reduction by DR (p=100%)

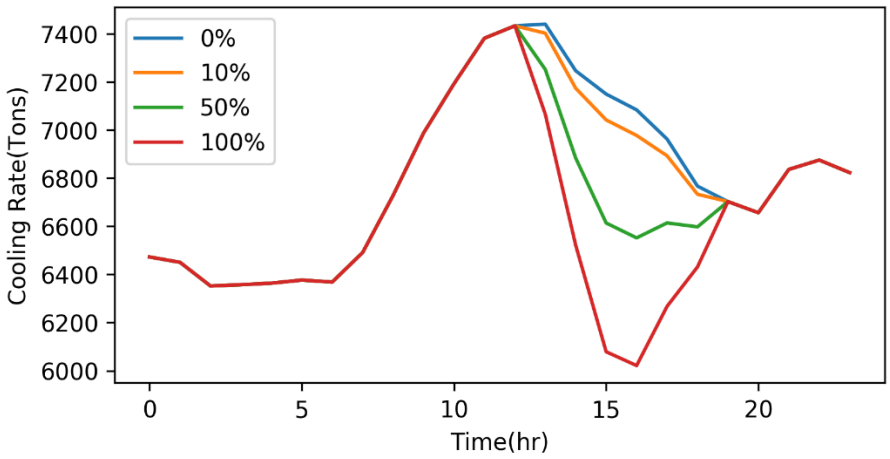


Figure 94 Expected DR Load Shape by Percentage of DR Buildings

6.4 Step 3: Plant Side Impact Analysis

In this step, the DR benefits are evaluated from the two perspectives mentioned in Section 3.1. The plant performance approximation curve and the lookup table developed in Section 5.2 are utilized to evaluate the impact of DR on the plant operation. The evaluation metric is the operating cost savings at the plant during active DR period. Although the amount of power consumption can be converted into the monetary value, the fluctuation of electricity price is highly unpredictable, especially for the DR periods. Therefore, electrical energy savings will be provided to estimate the operating cost savings. The plant availability change will be examined by looking at the optimal chiller combination sequences. The implemented assessment framework presented in Figure 95. The DR load profile generated in Step 2 and the original load profile are the inputs of the plant surrogate model developed in Section 5.2. Once the corresponding power consumption profiles and the optimal chiller sequences for the inputs, the energy savings are calculated and the change in the operation is observed.

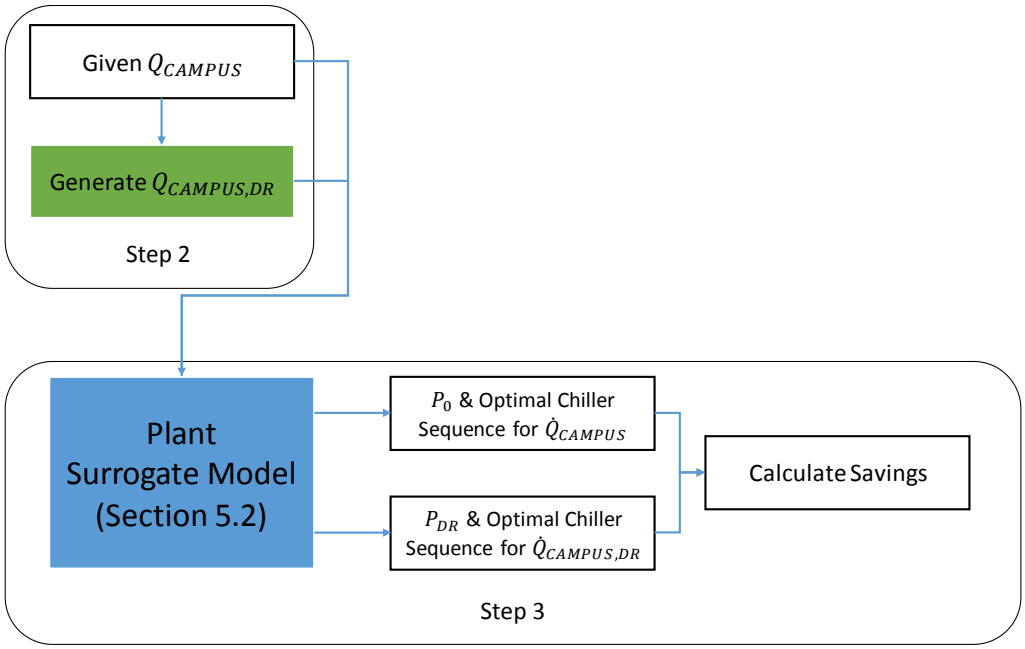


Figure 95 DR Impact Assessment Framework

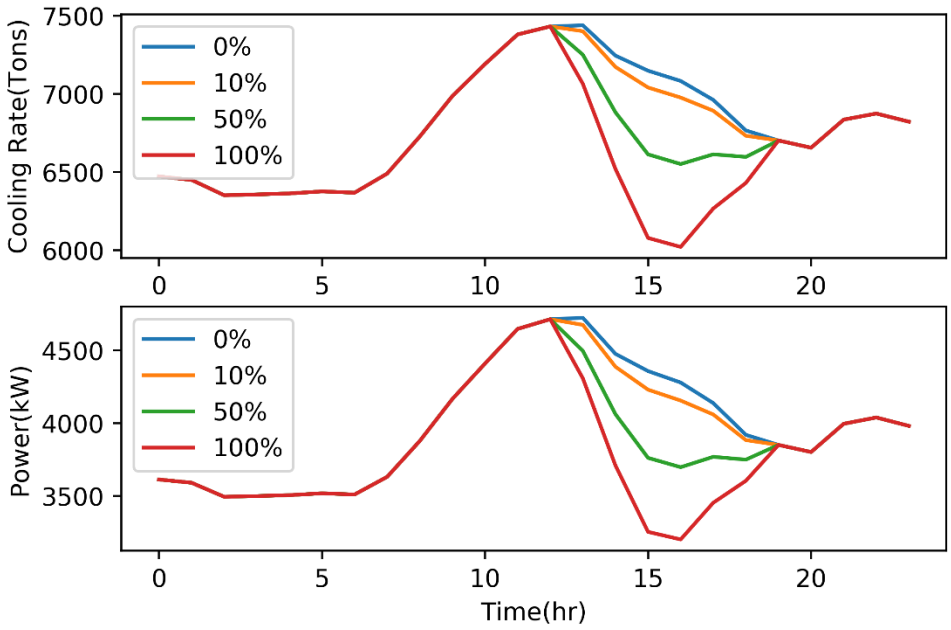


Figure 96 Expected DR Load Shape and Plant Power Consumption

Plant power consumption is expected to decrease as in the same way of cooling load. It is the obvious trend because the power consumption of the plant is monotonically increasing with the cooling load as can be seen in Section 5.2 when the plant is operated with the optimal chiller settings. By increasing the number of DR buildings, energy savings per day can be achieved up to 4,368 kWh in average. If the electricity price is $0.04, then the operating cost saving is about $175 per day. During DR period, the price increases by the multiple times of ordinary days' price. If the price is doubled, and 10 times of DR event occurs during a summer, then the cost saving is estimated to be $3,494. Therefore, quite big savings are expected when the entire campus buildings participate in DR program. However, the uncertainty in the prediction is large and increasing with the percentage of DR buildings as presented in Figure 97. It shows the range of achievable energy savings. If a decision maker considers increasing DR participation to get rewards, the uncertainty should be kept in mind to estimate the benefit. To this end, the inverse CDF plot is provided to help a decision maker to gauge the probability of achieving the target savings as presented in Figure 98. The intersection with the vertical line and the inverse CDF plot for each case indicates the probability of getting the desired energy savings. In this example, 2,000 kWh savings can be achieved with almost certainty if the entire campus participates in DR while it is about 70% success when the half of the campus participate in DR. Using this plot, one can plan to increase the DR participating buildings in the campus.

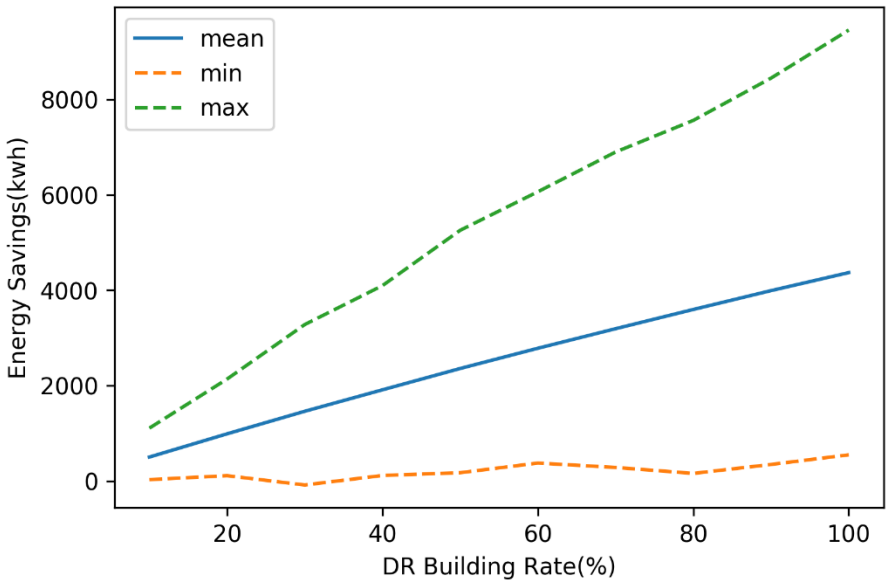


Figure 97 Energy Savings Envelop for Varying DR Building Percentage

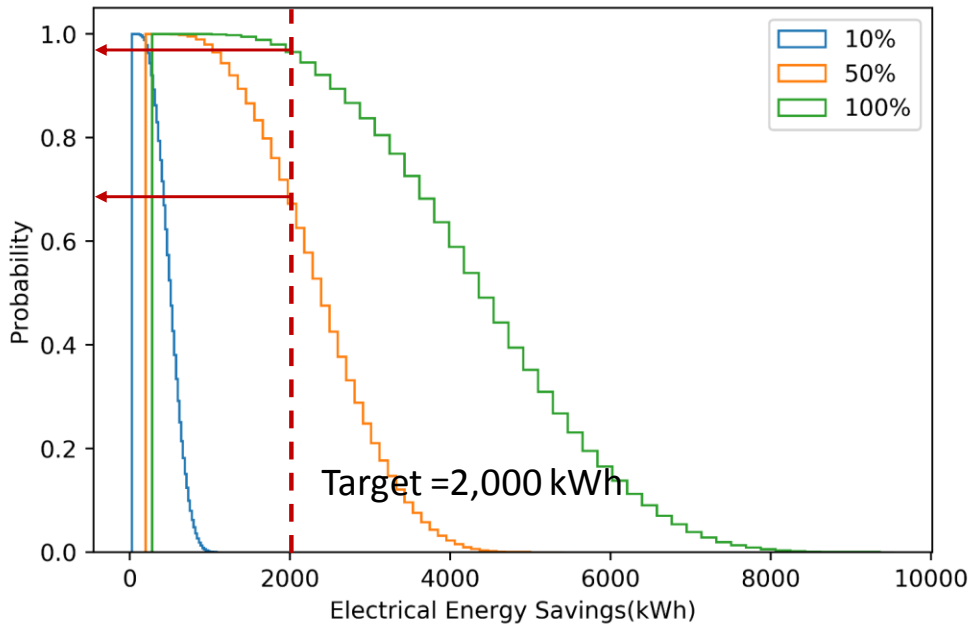


Figure 98 Inverse CDF Plot to Estimate the Probability of Achieving the Target Energy Savings.

Up to this point, the DR benefit is evaluated from the perspective of the energy customer. Now it will be evaluated from the second perspective, an energy producer and the system operator. Figure 99 demonstrates how the reduced load can affect the plant operation. The plant maximum capacity profile is calculated from the optimal chiller sequences by multiplying the chiller's capacity. It is observed that the number of required chillers are reduced during the DR period, which means that the plant availability increases due to DR. Because DR cannot reduce the peak load, the peak load availability is not affected. However, it can be done by the facility manager of Georgia Tech when it is needed. To confirm this statement, a simulation is performed by fixing the DR period to be at the peak time. It is assumed that DR starts at 11 am and ends at 3 pm. Figure 100 shows the resulting cooling demand and chiller sequencing curves. By reducing the peak load, it removes the necessity of staging another chiller at the peak load.

In this step, the DR benefits are examined for the cost savings and the improvement of the operational availability for the current system. The capability of DR to alter the operations of the chiller plant is confirmed by simulations. Having the DR model developed in Step 2, the following steps will explore the potential of DR to replace a redundant chiller in the long term capacity expansion planning.

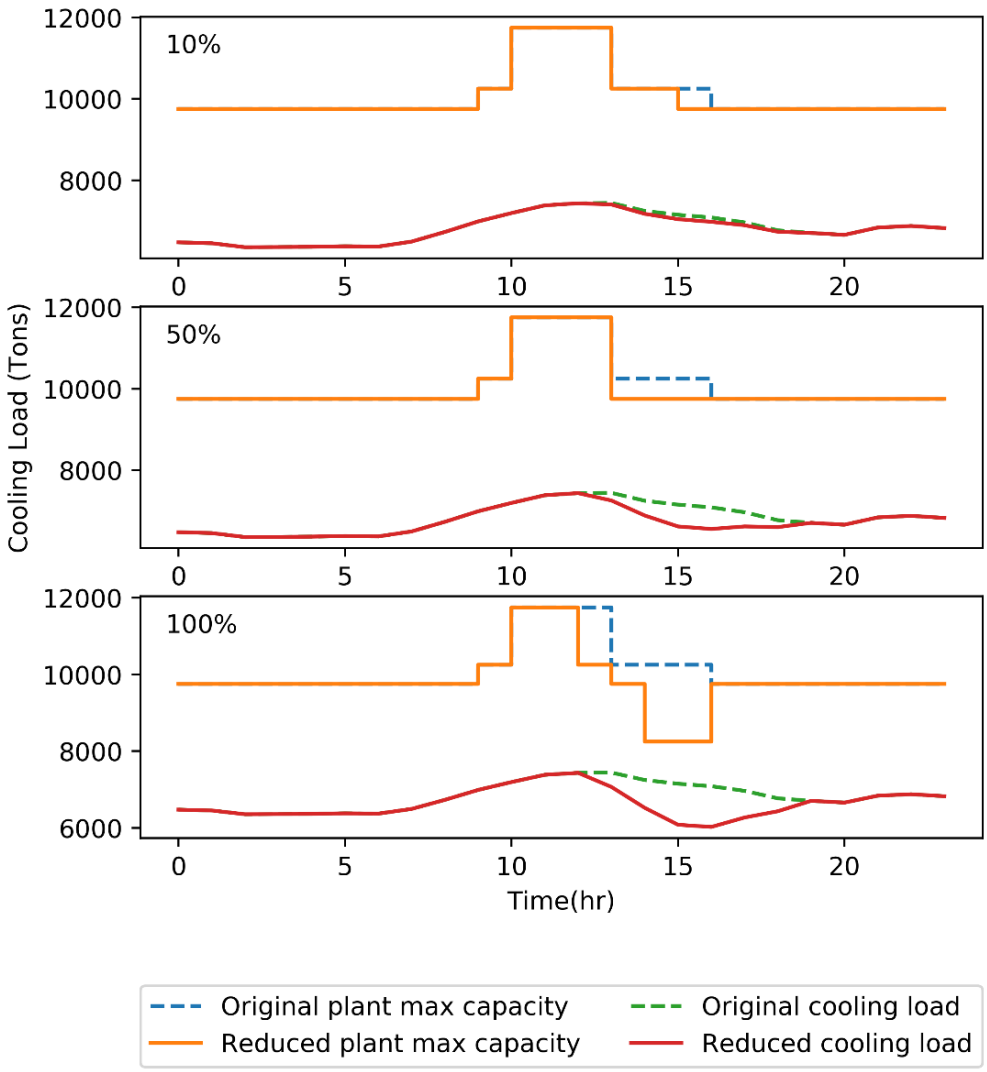


Figure 99 Expected Plant Optimal Chiller Sequencing by Varying DR Building Percentage

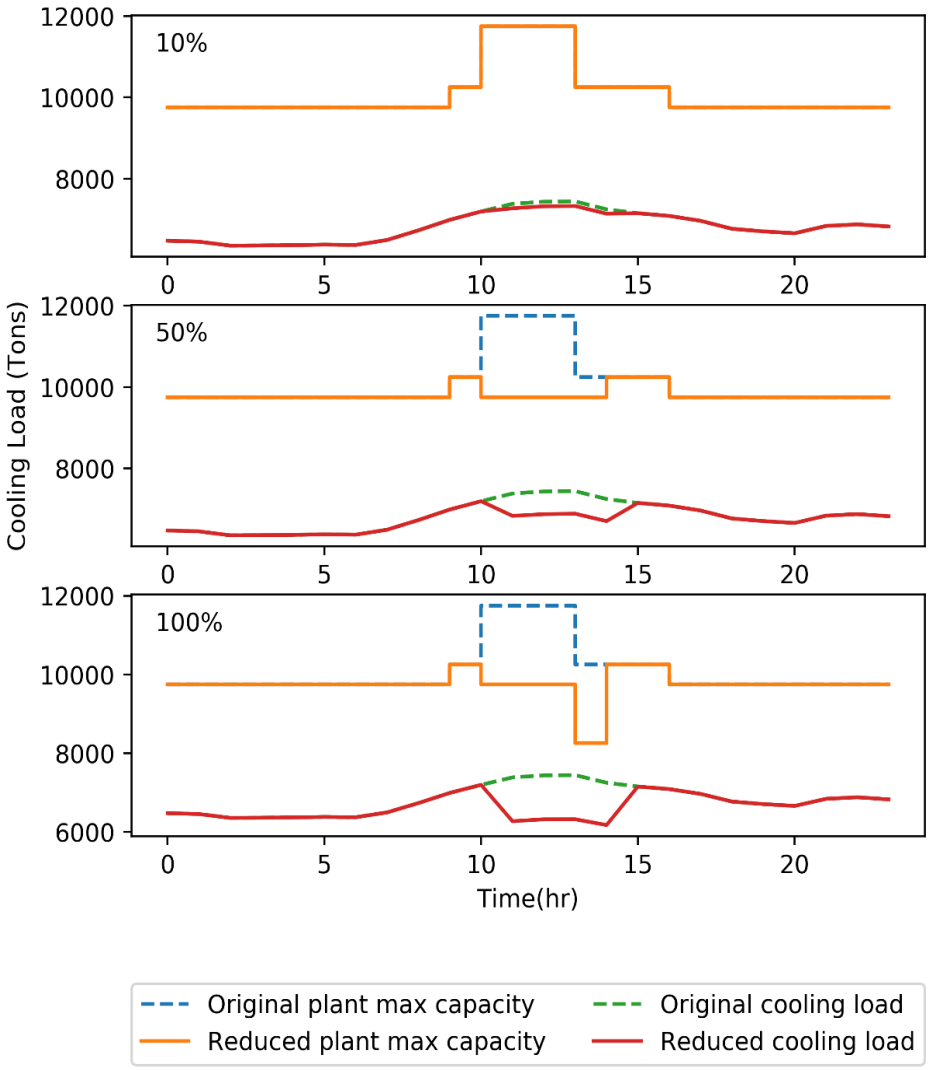


Figure 100 Expected Plant Optimal Chiller Sequencing When DR Event Occurs at Peak Time

6.5 Step 4: Demand Growth Forecast

This step prepares the input for the next steps. The capacity expansion plan is driven by the load growth. The future demand is determined by the peak load and the annual load shape. The peak load is used to determine the capacity expansion plan, and the annual load profile is required to evaluate the system reliability given the plant configurations.

6.5.1 Baseline Load Profile

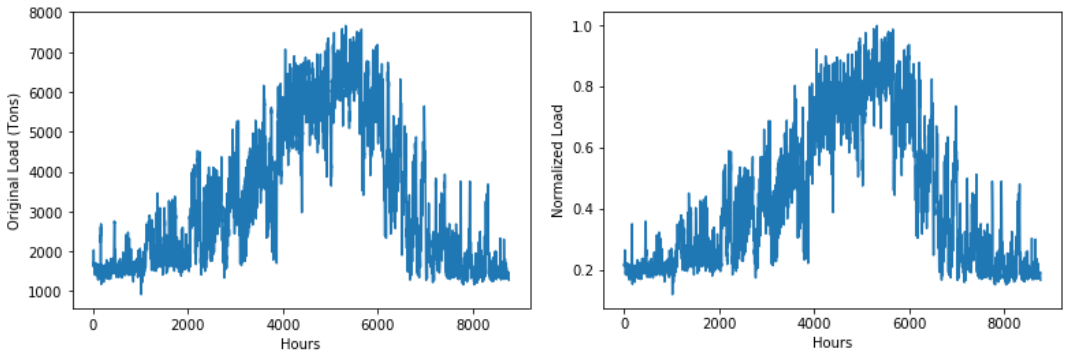


Figure 101 Baseline Annual Load and Normalization

The baseline annual load profile is created from data of the recent four years (2012-2016). The ANN model is used to generate the load profile for 8760 hours at one hour intervals (1 year). The model predicted load profile is normalized to be scaled up for the future load growth. The original load and the normalized profiles are presented in Figure 101. They have the same shape. The current plant capacity is 15,000 tons while the current peak load is well below 10,000 tons. It shows how the system is oversized. For the hypothetical campus, the current peak load is assumed to be 10,000 tons. So, the normalized load is scaled up to10,000 tons of peak as presented in Figure 102.

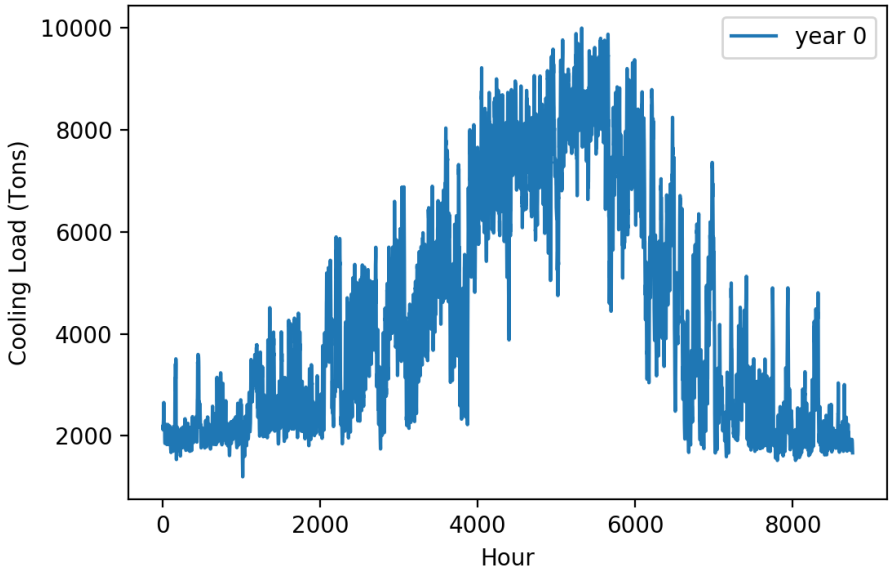


Figure 102 the Annual Cooling Load in Year 0

6.5.1 Demand Growth Scenario

The peak load is increased by the natural growth and the new building connection. The factors affecting the natural growth can be population growth or climate change. The constant growth rate is assumed here. The new buildings listed in Table 12 are modeled as random variables to generate the peak load growth scenario. If the peak load can be deviated from the estimation up to 20%, then the peak load of a new building is sampled from the uniform distribution between 80% and 120% of the estimation. Algorithm 2 describes the peak load growth scenario generation process. The load growth scenarios generated in this step are transferred to Step 5 and Step 6. Step 5 picks a deterministic load growth path to optimize the capacity expansion plan, and Step 6 takes all the scenarios and generate a set of annual load profiles by multiplying the normalized baseload for Monte Carlo Simulation of the reliability metric calculation. How each of steps is connected is illustrated in Figure 103.

Algorithm 2: Pseudocode of Peak Load Growth Scenario Generation

Input: the new building plan given by 2-D array $[T_{NB}; \dot{Q}_{NB}]$, $\dot{Q}_{peak}(0)$: the current peak load in year 0

Output: the peak load growth projection $\dot{Q}_{peak} = \left(\dot{Q}_{peak}(1), \cdots, \dot{Q}_{peak}(T)\right)$

Determine gr: the annual growth rate

Determine ur: the uncertainty range of peak load estimation

Initialize $\dot{Q}_{peak}$

for t = 1 to T do

If t ∈ T_{NB} **then**

$\dot{Q}_{nb}(t) = \text{uniform}(1 - \text{ur}, 1 + \text{ur}) \times \dot{Q}_{NB}(T_{NB}(i) = t)$

Else

$\dot{Q}_{nb}(t) = 0$

$\dot{Q}_{peak}(t) = (1 + gr)\dot{Q}_{peak}(t - 1) + \dot{Q}_{nb}(t)$

end for

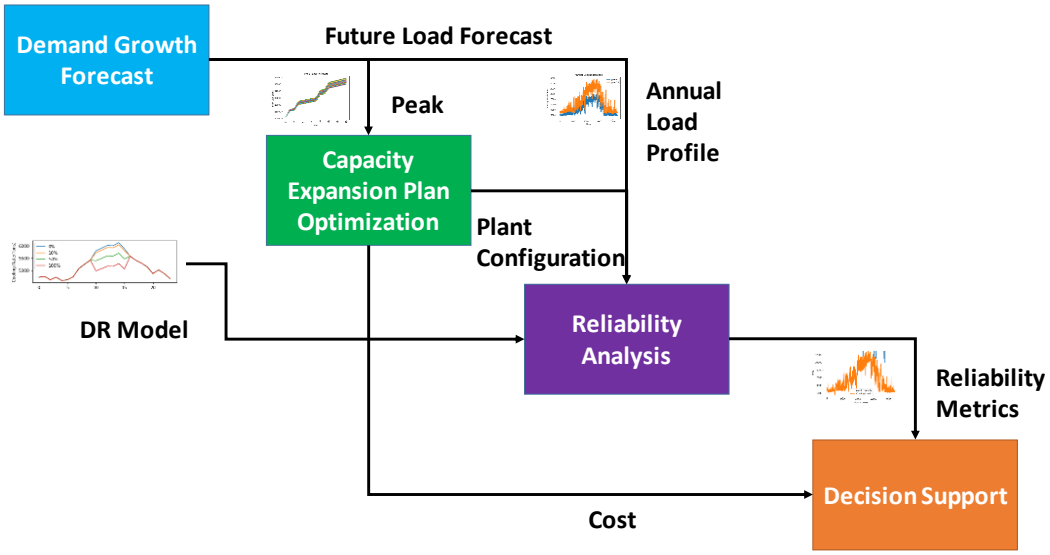


Figure 103 Information Flow between Steps

Having the building construction plan, Monte Carlo Simulation is performed to generate a set of load growth scenarios under the assumptions described in Table 14. Figure 104 shows the generated scenarios from 10,000 runs of MCS. The deterministic forecast represents the traditional sizing method which multiplies a safety factor to the estimation. In this case, 20% of safety factor is used. The uncertainty bound is increasing over time because the number of new buildings represented as random variables is added up over time. Once the peak load growth path is determined, then the annual load profile is generated as displayed in Figure 105. The annual load profile is shifted up while maintaining the shape.

Table 14 Scenario Assumptions

Parameter	Value
Initial Peak Load	10,000 tons
Planning Horizon	15 years
Growth Rate	1 % /year
Uncertainty Range	20%

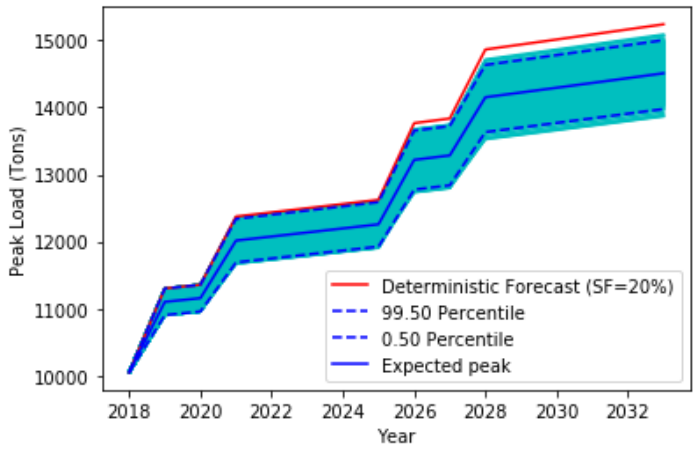


Figure 104 Peak Load Growth Scenarios

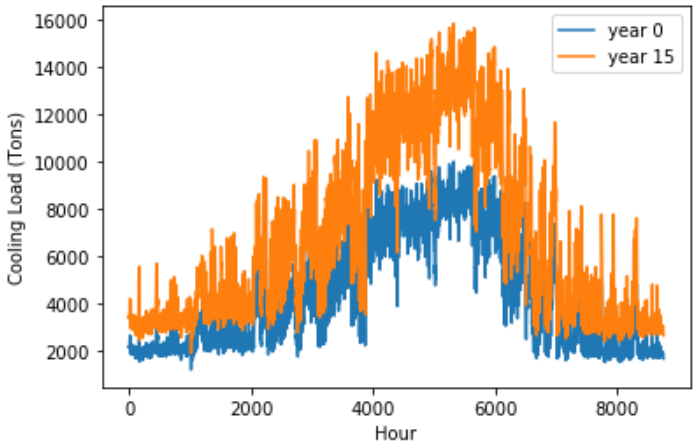


Figure 105 Annual Load Profile After 15 Years

6.6 Step 5: Capacity Expansion Planning Optimization

The purpose of formulating an optimization problem is to establish a baseline for the traditional approach and evaluate how DR would change the original plan. It allows generating a consistent optimal capacity expansion plan under the same condition. Given the load growth path and the planning horizon, the MILP optimization module yields the optimal capacity expansion schedule as a set of year and size, and the objective function calculate NPV (Net Present Value) of the optimal plan based on the assumption of 3% discount factor. To explore the planning design space, the safety margin was introduced in Equation (40). It alleviates the strict redundancy constraint while offering a smaller buffer. By changing the safety margin, one can produce a set of solutions between two extreme points as illustrated in Figure 41. In this example, the safety margin is increased by an increment of 500 tons from zero to 2,000 tons. The optimal solutions are transformed to the corresponding plant configuration array of each year over the planning horizon and transmitted to the reliability simulation module.

Starting from the year of 2018, the expected future demand growth path for the next 15 years is created as presented in Figure 104. Among those scenarios, two sample paths are selected to demonstrate the problem solving approach for the capacity expansion planning optimization. The expected growth path and the extreme path, which has 20% safety factor added to the deterministic estimation, are used as inputs for the MILP formulation.

Table 15 Optimal Capacity Expansion Plans by Margin

	Expected Growth Path		Safety Factor 20%	
Margin	NPV ($MM)	Plan (Year, Size)	NPV ($MM)	Plan (Year, Size)
0	2.80552	(12, 1500)	2.97638	(10, 1500)
500	2.97638	(10, 1500)	3.34842	(10, 2000)
1000	3.34842	(10, 2000)	3.94705	(8, 3000)
1500	3.94705	(8, 3000)	3.94705	(8, 3000)
2000	3.94705	(8, 3000)	6.60079	(7, 1500) (10, 2000)
3000 (N+1)	7.24795	(4, 3000) (12, 1500)	7.41881	(4, 3000) (10, 1500)

The solutions by the margin and the demand projection are listed in Table 15. Obviously, smaller margin yields cheaper solutions. The planning time line is pushed back as reducing the margin from the N+1 plan as presented in Figure 106 and Figure 107. However, the change in the optimal plan is irregular. This is because the integer variables create discrete solution space. The original N+1 plan requires adding 3,000 tons of chiller in year 4 and 1,500 tons of chiller in year 12 based on the expected peak growth projection. The safety factor on the peak load estimation enforces earlier expansion schedule. Either

the safety factor or the margin does the same function: creating a buffer. As can be seen in Table 15, adding safety factor pushes forward the plan. For example, the solution for the margin of 1500 and the expected growth path is the same as the solution for the margin of 1000, and the safety factor added projection. In the traditional capacity planning, both of them are used to result in overly sized systems.

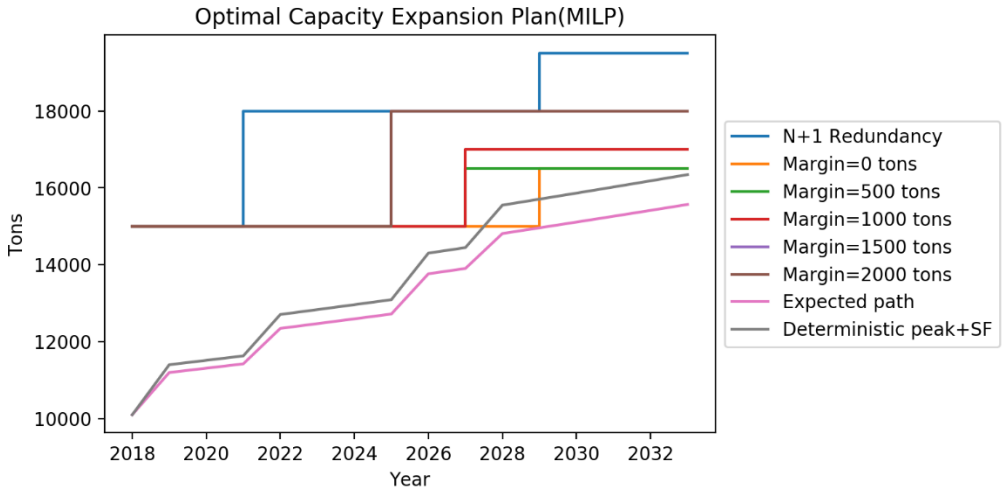


Figure 106 Optimal Plans for Expected Peak Growth Path

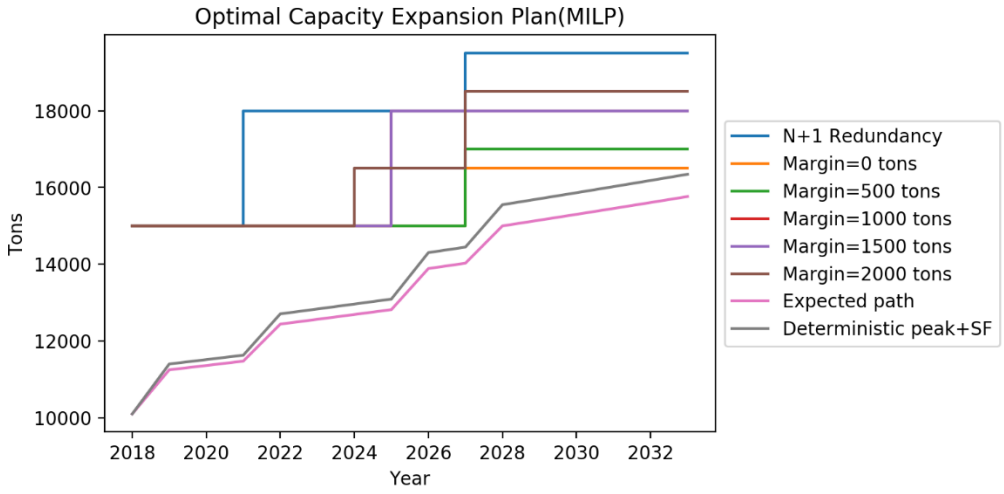
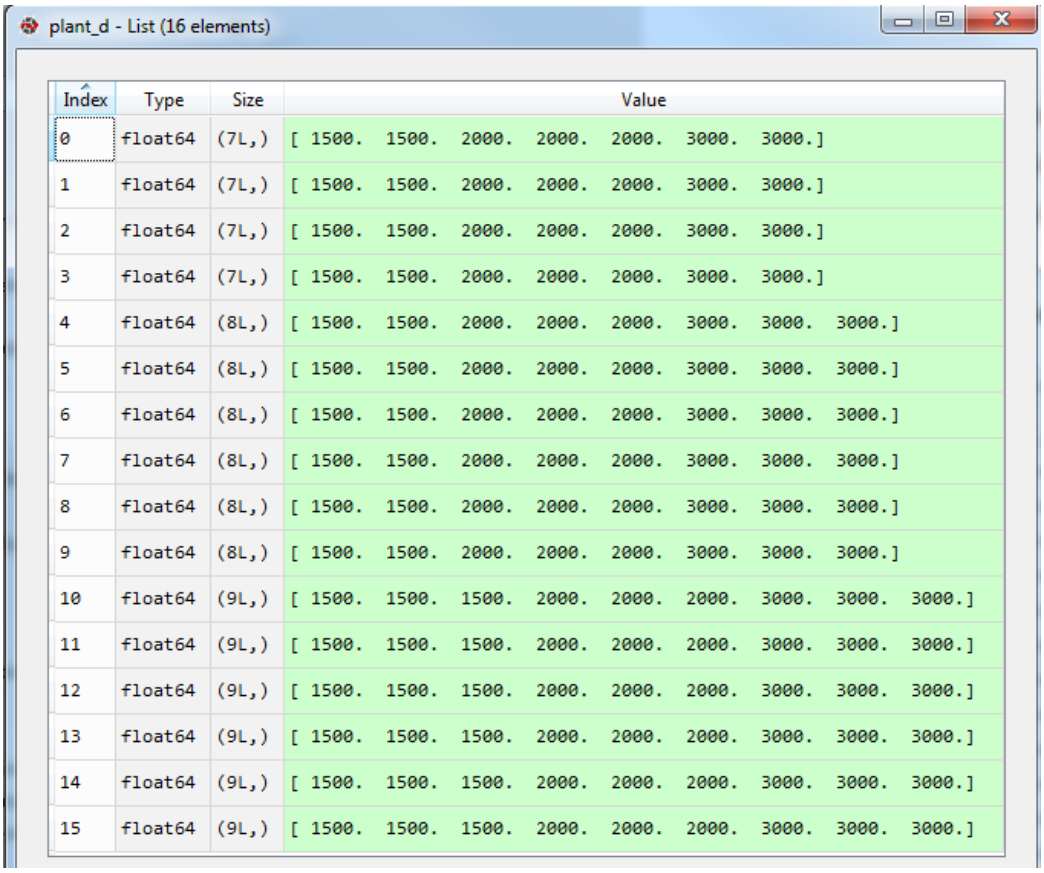


Figure 107 Optimal Plant for Cautious Projection (Safety Factor 20%)

Now that the optimal solutions are generated for different margins, how reliable or resilient the resultant system is in a contingency event should be quantified. The next step will run the stochastic simulations to quantitatively evaluate the plans. The optimal plans should be converted to the array of chillers for each year to run the plant availability simulation. The plant array represents the evolutionary path of the plant capacity expansion as presented in Figure 108. Index represents year in the planning horizon and the array value shows the combination of chillers. It is passed to the next step for assessment of the system reliability with DR application.



Index	Type	Size	Value
0	float64	(7L,)	[1500. 1500. 2000. 2000. 2000. 3000. 3000.]
1	float64	(7L,)	[1500. 1500. 2000. 2000. 2000. 3000. 3000.]
2	float64	(7L,)	[1500. 1500. 2000. 2000. 2000. 3000. 3000.]
3	float64	(7L,)	[1500. 1500. 2000. 2000. 2000. 3000. 3000.]
4	float64	(8L,)	[1500. 1500. 2000. 2000. 2000. 3000. 3000. 3000.]
5	float64	(8L,)	[1500. 1500. 2000. 2000. 2000. 3000. 3000. 3000.]
6	float64	(8L,)	[1500. 1500. 2000. 2000. 2000. 3000. 3000. 3000.]
7	float64	(8L,)	[1500. 1500. 2000. 2000. 2000. 3000. 3000. 3000.]
8	float64	(8L,)	[1500. 1500. 2000. 2000. 2000. 3000. 3000. 3000.]
9	float64	(8L,)	[1500. 1500. 2000. 2000. 2000. 3000. 3000. 3000.]
10	float64	(9L,)	[1500. 1500. 1500. 2000. 2000. 2000. 3000. 3000. 3000.]
11	float64	(9L,)	[1500. 1500. 1500. 2000. 2000. 2000. 3000. 3000. 3000.]
12	float64	(9L,)	[1500. 1500. 1500. 2000. 2000. 2000. 3000. 3000. 3000.]
13	float64	(9L,)	[1500. 1500. 1500. 2000. 2000. 2000. 3000. 3000. 3000.]
14	float64	(9L,)	[1500. 1500. 1500. 2000. 2000. 2000. 3000. 3000. 3000.]
15	float64	(9L,)	[1500. 1500. 1500. 2000. 2000. 2000. 3000. 3000. 3000.]

Figure 108 Plant Array of Optimal Capacity Expansion Plan

6.7 Step 6: Reliability Analysis

The purpose of this step is to evaluate how much DR can improve the reliability metric. Having the optimal plans, the yearly plant configurations are created as the array vector as in Figure 108 and used as the input to the reliability metric calculation algorithm presented in Figure 77. Loss of Load Expectation is adopted from the power grid reliability metrics. Even though the failure is defined as the deficiency of capacity to serve the required load, a district energy system will not experience black out type failure. Therefore, the metric is renamed as "Average unsatisfactory hours." Calculation of the average unsatisfactory hours needs demand and plant models. In Step 4, demand growth scenarios were generated to account for the demand side uncertainty. In this step, the plant side uncertainty is quantified by running the stochastic plant availability simulation model, which is developed in Section 5.3.

The unsatisfactory hours are calculated by detecting whether the demand exceeds the available capacity of the plant. An algorithm is devised to integrate the DR model into the calculation. As discussed in Section 3.3.2, both of thermal and mass balance should be satisfied. Thus, the metrics are separately calculated for the cooling load and the mass flow rate capacity. The mass flow requirement is calculated from the approximation model of delta T. Throughout the loop for a year, DR is activated when a contingency event occurs, and the reliability metrics are calculated again for the DR modified load. Therefore, total four metrics will be the outputs of this calculation algorithm. Having this, MCS is performed to get statistics of the metrics with the stochastic plant availability and demand growth scenarios.

Algorithm 3: Pseudocode of Reliability Metric Calculation

Input: annual cooling load $(\dot{\mathrm{Q}}_t(1), \cdots, (\dot{\mathrm{Q}}_t(8760))$, plant array, DR model, planning horizon T

Output: unsatisfactory hours

Determine p: fraction of DR buildings

Initialize uhr_q: unsatisfactory hours of cooling, uhr_q_DR: with DR

Initialize uhr_m: unsatisfactory hours of mass flow, uhr_m_DR: with DR

for t = 1 to T do

  Calculate $\dot{\mathrm{m}}_t$ using $\dot{\mathrm{Q}}_t$ and ΔT model

  for h = 1 to 8760 do

    If $\dot{\mathrm{Q}}_t(\mathrm{h}) > \dot{Q}_{plant,t}(h)$

      uhr_q = uhr_q + 1

      $\dot{\mathrm{Q}}_{t,DR}(h) = p\dot{Q}_t(h)(1 - \Delta Load_{DR}) + (1-p)\dot{Q}_t(h)$

      If $\dot{\mathrm{Q}}_{t,DR}(\mathrm{h}) > \dot{Q}_{plant,t}(h)$

        uhr_q_DR = uhr_q_DR + 1

    Calculate $\dot{\mathrm{m}}_{t,DR}$ using $\dot{\mathrm{Q}}_{t,DR}$ and ΔT model

    If $\dot{\mathrm{m}}_t(\mathrm{h}) > \dot{m}_{plant,t}(h)$

      uhr_m = uhr_m + 1

      If $\dot{\mathrm{m}}_{tmDR}(\mathrm{h}) > \dot{m}_{plant,t}(h)$

        uhr_m_DR = uhr_m_DR + 1

  end for

end for

Total 3000 runs of MCS is performed to evaluate the system reliability. 300 samples of the plant availability and 10 samples of the demand growth scenarios are combined to the 3000 samples. The calculated reliability metrics for the optimal capacity expansion plans are shown in Figure 109 and Figure 110, for the expected load growth path and the safety factor path respectively. Upper four figures show unsatisfactory cooling load

for different margins, and the bottom four figures show the unsatisfactory flow rate. It is assumed that DR is applied to every campus buildings in this example. Blue bars depict the average unsatisfactory hours per year without DR, and the orange bars represent the average unsatisfactory hours per year when DR is applied to the system, and the gaps between two show that how much of the unsatisfactory hours are reduced by DR. When no margin is enforced to capacity sizing, it is expected to not satisfy the cooling load more than 4 hours per year in 2029. However, DR can alleviate this by reducing the unsatisfactory time to less than 1 hour. Having small margins, the system has much more improved reliability metrics. Even with the margin of 500 tons and DR, the reliability metrics become close to zero. The cautious demand projection using the safety factor results in more reliable systems. It advances the expansion schedule by 2 years earlier, and the reliability metric at the same year (2029) reduces to less than 1 hour.

Graphs in Figure 109 and Figure 110 can present how much DR can improve the system reliability in terms of the unsatisfactory hours. Each graph comes from one optimal plan. It can help a decision maker to predict the optimal capacity plans' reliability. Cost of each plan is given in Step 5. Having two performance metrics of each plan, cost and reliability, the next step will integrate step 5 and 6 into a trade-off space to help decision making.

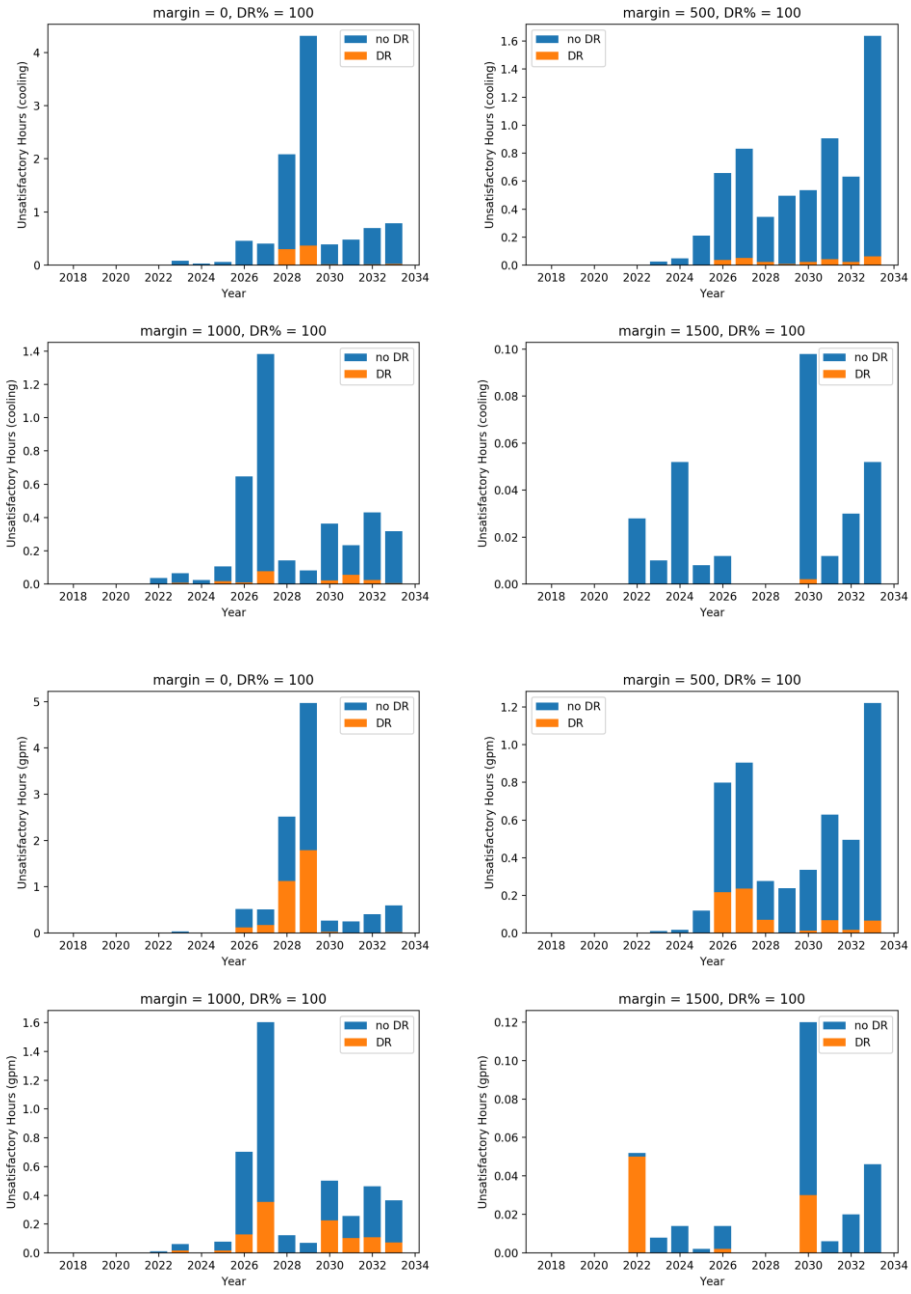


Figure 109 Improvement of Unsatisfactory Hours by DR: Expected Growth Path

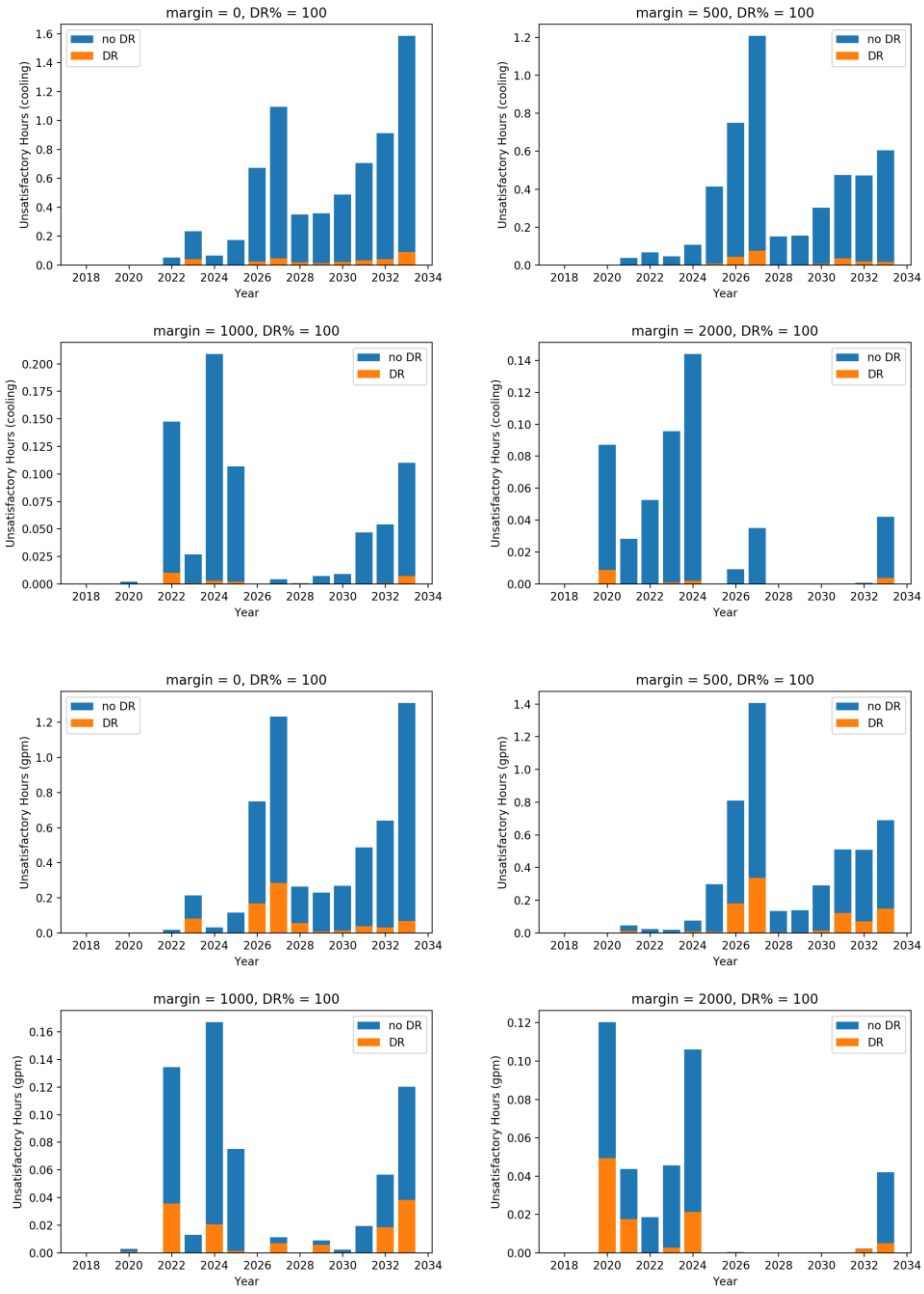


Figure 110 Improvement of Unsatisfactory Hours by DR: Safety Factor

6.8 Step 7: Decision Support

The final step of the methodology is to visualize the trade-off space to support decision making. By creating the Pareto frontier, it proves the validity of the methodology. Recall the last research question.

Research Question 4:

How can we predict if DR is capable of avoiding or deferring capacity expansion without sacrificing reliability?

Research Question 4 aims to develop a prediction method of the DR capability and its effectiveness on the capacity expansion plan. This question cannot be answered by a simple method. The methodology has been developed to use measured data and convert them into the predictive models, and connect them into an integrated framework of optimization and simulations for capacity expansion planning and the reliability assessment. This formulation approach raised another research question.

Research Question 4.1:

How can we know the optimality of the solution for the capacity expansion plan with the proposed decision making framework?

A thought experiment led to Hypothesis 4.1.

Hypothesis 4.1:

With the two extreme points of the N+1 constraint and no safety margin, the Pareto frontier can be generated by adjusting safety margin on the demand satisfaction constraint between the two points and the fraction of DR buildings.

The charts in Figure 109 and Figure 110 provide an idea of how the planned system would perform in contingency events to ensure the system reliability. Which plan is better is not a question that can be simply answered. As discussed earlier, it is a multi-objective problem that requires a compromise between capital cost and reliability. Different levels of margins and load growth scenarios can be interpreted into the preferences of a decision maker about risk taking and frugality. Comfort cooling is usually considered as non-critical load, which is why the cooling load is a great resource for DR. There are, however, non-negotiable cooling load such as process cooling of a data center. In this case, a decision maker rather spends money on the large enough capacity than taking the risk of frequent load shed. A university or corporate campus has mixed purposes of cooling. To account for this, a condensed, visualized trade-off space is created by taking all information from step 5 and 6.

6.8.1 Trade-off Study

Figure 111 describes how the constituent optimization and analysis results are integrated into the visualized trade-off space. The Pareto frontier is a collection of multiple optimal solutions that cannot be improved in any directions. The assumption behind this plot is that an optimal capacity expansion plan for the minimal capital cost can be further improved in terms of reliability using DR without additional cost.

Once the optimal plans are generated based on a load growth scenario and the safety margin (Step 5), the resultant plans are fed into the stochastic simulation environment (Step 6). The NPV calculated in Step 5 will be the x value of the Pareto frontier. The reliability metrics are computed for each of the plans with and without DR, which are y values of the unsatisfactory hours vs. NPV plot. The fraction of DR can be varied, but it was fixed to 100% in this example study to maximize the effectiveness of DR.

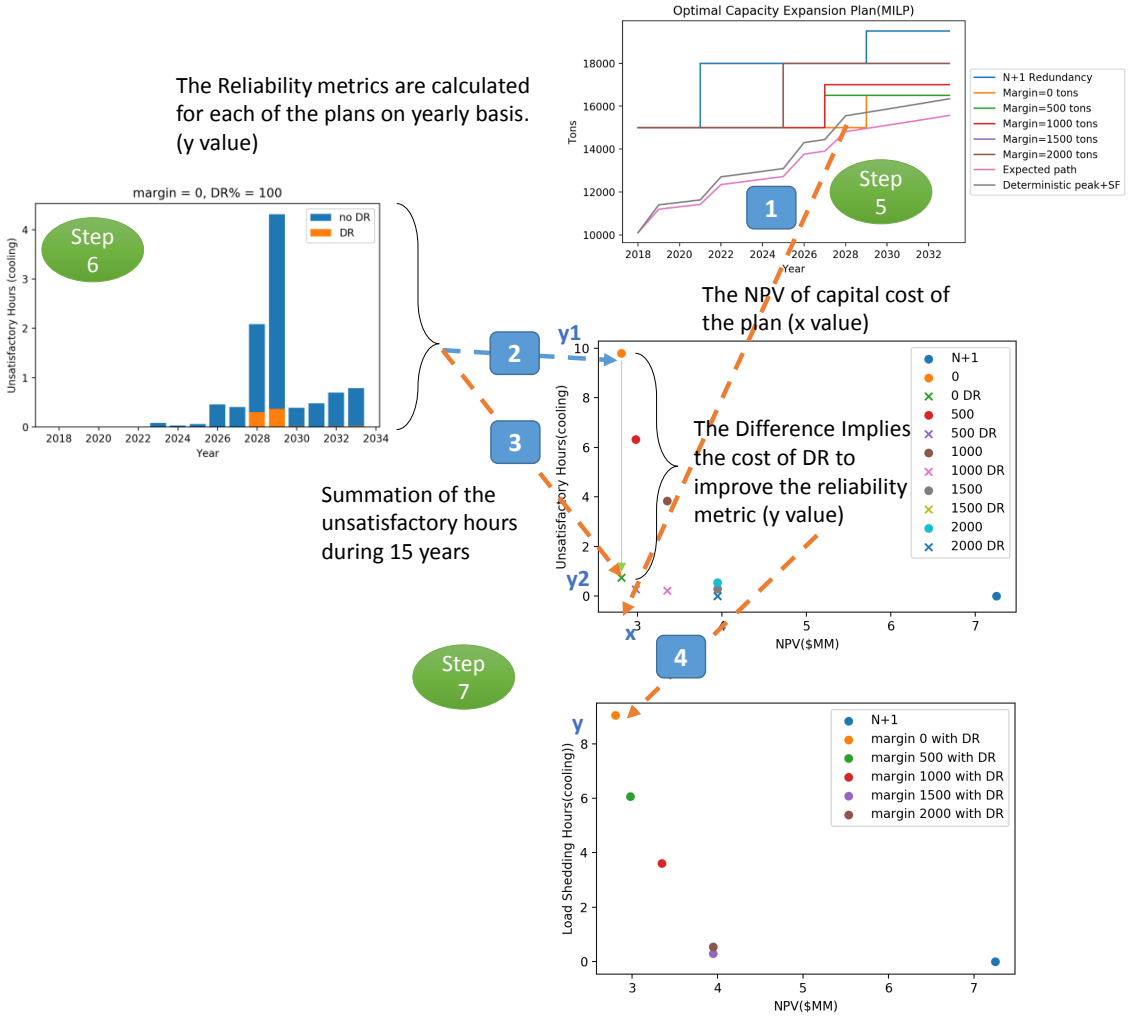


Figure 111 Information Flow of the Analysis Result of Each of the Steps into the Pareto Frontier Plot

The unsatisfactory hours for the entire planning horizon is collected for each of the plans and joined to the data set of the NPV of the optimal plans. The collected data set is plotted in the NPV vs. the reliability metric space as presented in Figure 112 for the thermal load and Figure 114 for the mass flow rate. The gap between the original (denoted by circles) and the DR activated systems (denoted by cross signs) for the same plan can be interpreted as the cost of DR because it is required to shed load for the specified amount of

time by the difference, which is the y value of another Pareto frontier plot, e.g., seen at the bottom of Figure 111. Even though there is no penalty cost in a district energy system for load shedding, it can cause discomfort to occupancies and should be avoided as far as possible. To that end, another Pareto frontier plot is generated as presented in Figure 113 for thermal and Figure 115 for mass. The same plots are generated for the extreme demand growth scenario with 20% safety factor are presented in Figure 116-Figure 119. Numbers for the metrics are listed in Table 16.

In summary, the Pareto frontier plots are generated through the steps as below.

- Determine a load growth scenario and a safety margin
- Step 5: Get the optimal NPV (x) for a specific scenario from the MILP optimization module
- Step 6: Calculate the reliability metrics for the optimal plan without DR (y1) and with DR (y2) using the stochastic simulation module
    - Pareto frontier: NPV vs. Unsatisfactory hours
- Calculate the difference between y1 and y2 (y)
    - Pareto frontier: NPV vs. Load shedding hours

The decision maker can now use these plots to explore design space of a capacity expansion planning problem. The NPV vs. unsatisfactory hours plots, which are presented in Figure 112, Figure 114, Figure 116, and Figure 118, show how each of plans will perform in contingency events. Instead of a deterministic redundancy or safety margin, the expected hours that a plan cannot satisfy demand over the planning period provide more precise prediction of the system reliability. She can make a decision based on her

preference for reliability or capital cost. The most economical plan, denoted by the orange circle in Figure 112, is expected to fail to meet cooling load for about 10 hours while DR can reduce it to less than 2 hours as denoted by the green cross sign. This difference is interpreted as cost of DR in terms of load shedding hours, and cost of the Pareto frontiers is mapped into another Pareto frontier plot as presented in Figure 113. Having those Pareto frontiers, the decision maker has multiple optimal options for capacity expansion plan. If she does not want to enforce load shed for about eight hours, then she can consider spending one million dollars more on capacity to ensure the system have a larger safety margin and perform similar to the N+1 system.

An interesting trend can be found in the trade-off space. The N+1 redundancy system is highly reliable as expected, yet very expensive. Compared to other plans, it requires more than three million dollars. The plans with a small margin need to shed load to achieve the comparable reliability to the N+1 system. In between those, the plans with a fairly large margin (1500 and 2000, both yield the same plan) can maintain the quite high reliability with less spending than the N+1 system. The same trend is found in the other scenario. In the extreme demand growth scenario, the margin of 1000 and 1500 yield the same plan, and the margin of 2000 is close the N+1. It is not a surprising result because having the safety factor on the peak load estimation and adding a margin on the capacity requirement are combined into the larger margin.

The practical question led to this research was to ask if DR can replace the redundant chiller while maintaining the same level of reliability. That question can be answered at this point. If the system capacity is determined exactly to meet the estimated peak load, the system will more likely fail. DR can mitigate the risk of insufficiency in capacity; however, it is cannot be recovered to the same level of the N+1 system. Adding a small margin to enforce the system capacity to a higher level, the system reliability can

be improved significantly. The effectiveness of the margin is not linearly increasing. There is a point where adding more margin is ineffective in terms of the improvement of reliability. Exploration of the trade-off space provides an insight into the relationship between the system reliability and margin. Because the relationship is nonlinear, it is not easy to predict without the simulations. The final decision is dependent on the preference of a decision maker. Having DR as another option, the capacity expansion plan can be deferred up to 8 years later and save several million dollars. The resulting plots and this statement support the Hypothesis 4 and prove the validity of the proposed methodology.

Hypothesis 4:

The proposed two-step decision making framework can yield multiple Pareto solutions and the corresponding optimal capacity expansion planning schedule. Using those, one can predict the impact of DR on the capacity expansion planning and the system reliability.

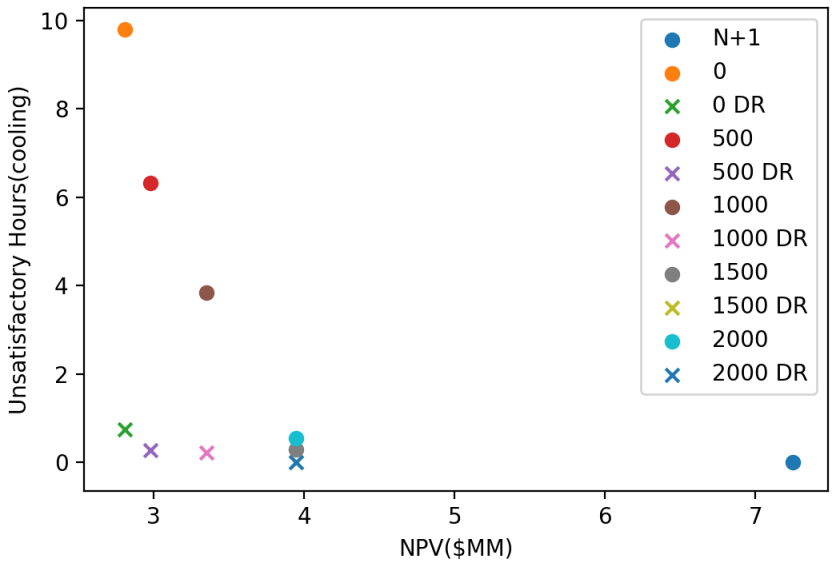


Figure 112 Pareto Frontier of Expected Demand Growth Scenario (Thermal Balance Reliability vs. NPV)

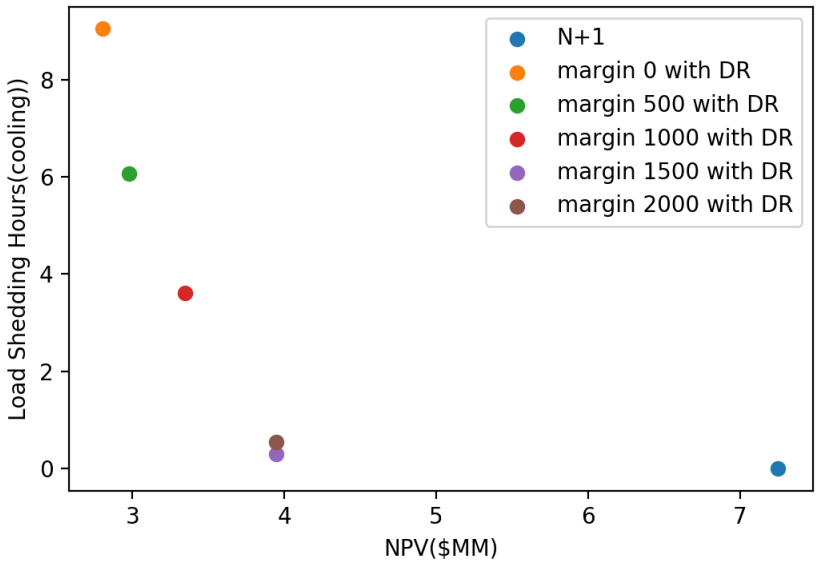


Figure 113 Pareto Frontier of Expected Demand Growth Scenario (DR Cost vs. NPV)

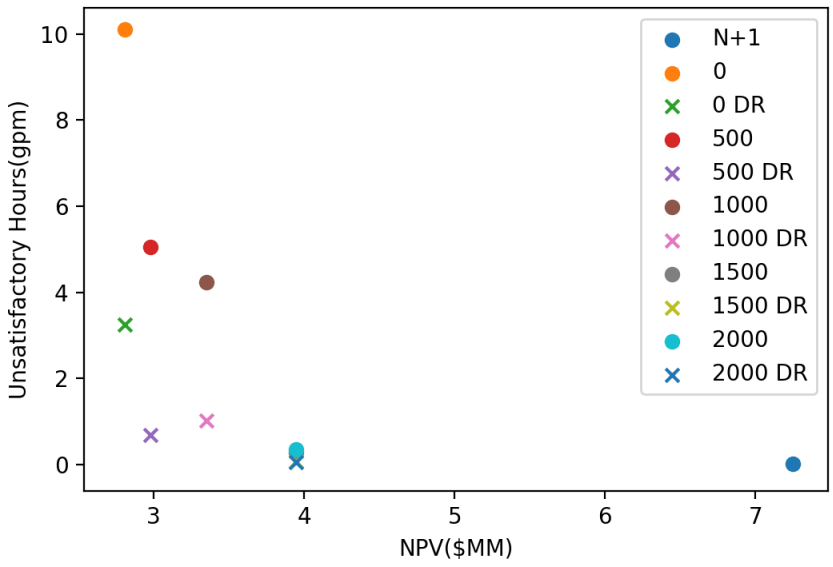


Figure 114 Pareto Frontier of Expected Demand Growth Scenario (Mass Balance Reliability vs. NPV)

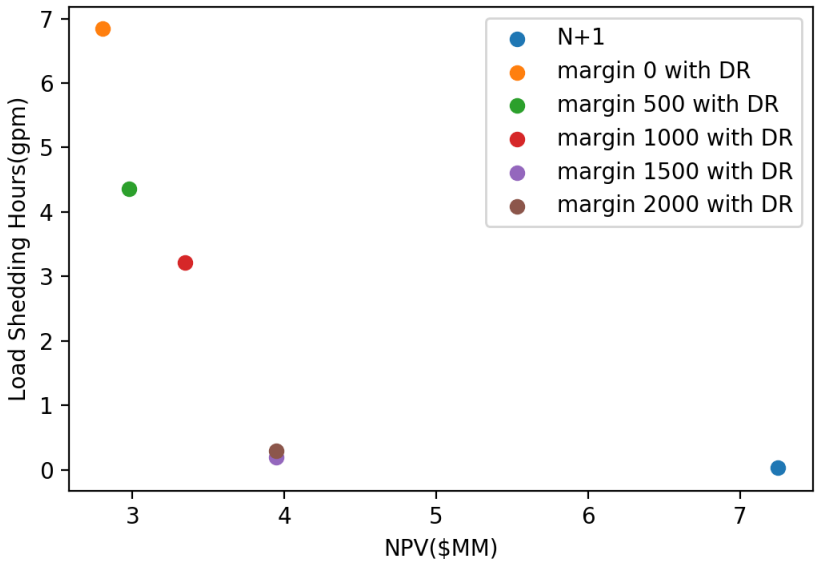


Figure 115 Pareto Frontier of Expected Demand Growth Scenario (DR Cost vs. NPV)

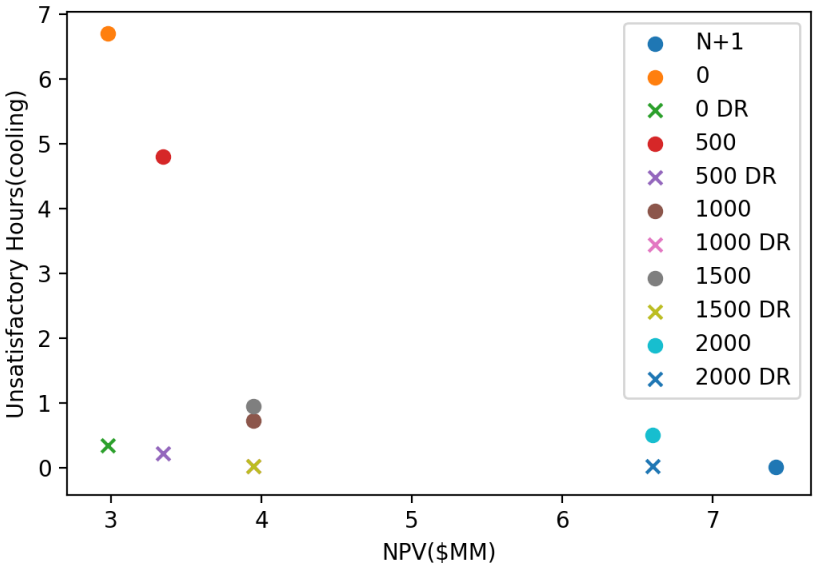


Figure 116 Pareto Frontier of Extreme Demand Growth Scenario (Thermal Balance Reliability vs. NPV)

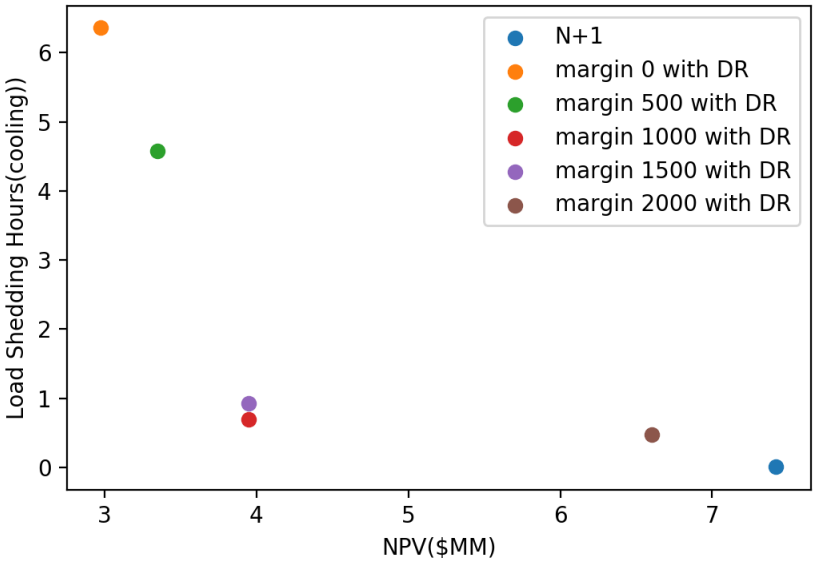


Figure 117 Pareto Frontier of Extreme Demand Growth Scenario (DR Cost vs. NPV)

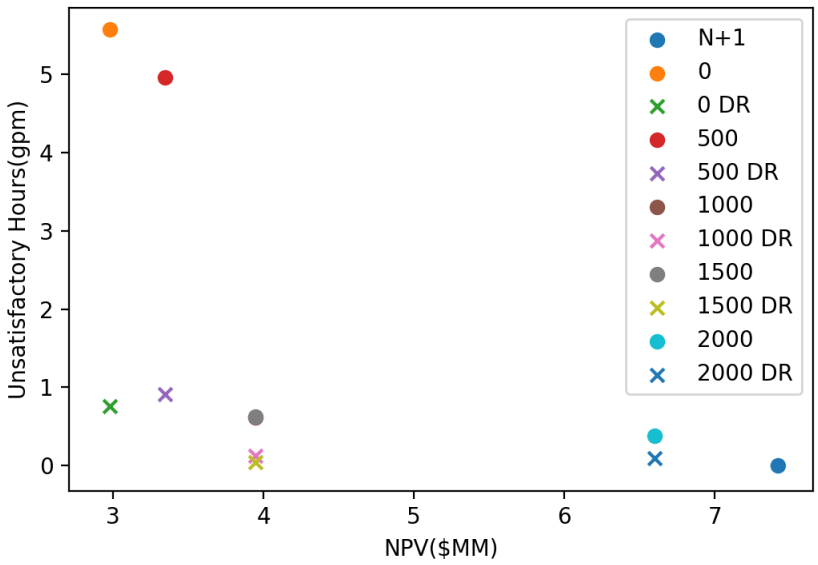


Figure 118 Pareto Frontier of Extreme Demand Growth Scenario (Mass Balance Reliability vs. NPV)

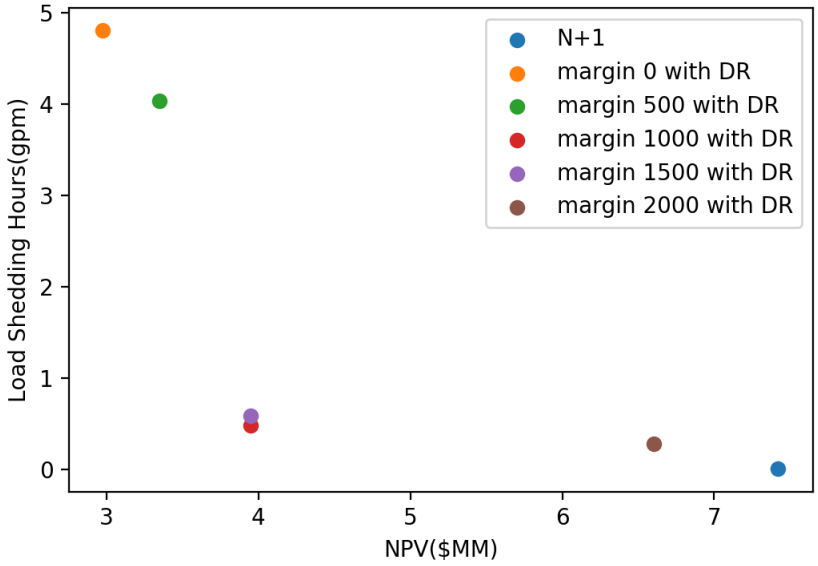


Figure 119 Pareto Frontier of Extreme Demand Growth Scenario (DR Cost vs. NPV)

Table 16 Cost and Reliability Metrics

	Plan	Cost	Unsatisfactory Hours (Thermal Balance)		Unsatisfactory Hours (Mass Balance)	
	Margin	NPV	no DR	DR	no DR	DR
Expected Load Growth	0	2.806	9.786	0.746	10.094	3.254
	500	2.976	6.328	0.272	5.046	0.686
	1000	3.348	3.834	0.222	4.238	1.028
	1500	3.947	0.302	0.002	0.282	0.082
	2000	3.947	0.556	0.010	0.350	0.054
	3000	7.248	0.002	0.002	0.028	0.028
Extreme Load Growth	0	2.976	6.822	0.353	5.494	0.764
	500	3.348	5.342	0.297	5.878	1.313
	1000	3.947	0.898	0.040	0.752	0.145
	1500	3.947	0.818	0.028	0.680	0.143
	2000	6.601	0.312	0.004	0.168	0.023
	3000	7.419	0.026	0.026	0.014	0.014

CHAPTER 7 CONCLUSION

7.1 Summary of Findings

This research was initiated by examining data of an existing system, where a problem is found that a chilled water plant has the excessive capacity compared to the observed peak load. A literature survey supported that the oversizing is a common problem for HVAC system sizing. Lack of understanding of uncertainty in the demand estimation and the failure mechanism forces to use the safety factor (demand side) and the redundancy (supply side). The combination of two deterministic risk mitigation methods creates a huge gap between the actual load and the designed system capacity. Efforts have been made to close the gap by improving the accuracy of the building's energy use estimation with the HVAC simulation models [95][32][96][97]. None of them questioned the necessity of the redundancy in a district level chiller plant.

A campus level district energy system is in the mixed position of an energy producer and consumer. It consumes the electrical energy to produce the thermal energy to multiple buildings. Advancement in building energy technologies and communication systems enable the concept of Demand Side Management to become a reality. Demand Response is now the established practice for the power system management. The established DR can be utilized by the owner of the district energy system to alleviate the oversizing problem if it can replace the redundancy.

The objective of this thesis is to quantitatively evaluate the impact of DR on a district energy system operation and planning. The primary research question has been posed to establish the research area.

What is the appropriate approach to analyze DR resources and evaluate the impact on the capacity planning of a district energy system?

Existing DR models found from the literature survey are grouped into two approaches. The first approach is for the high-level policy making problem. The price elasticity model is most commonly used one to optimize the price by predicting how customers act on the price. It is not applicable to a district energy system planning because it does not have a market structure. Even though the power grid system and the district energy system have the same structure, the scale and the physical aspect of the systems make a difference. Another approach is the HVAC system simulation to evaluate DR potential of a building and develop a control method. Because a district energy system has hundreds of buildings, the high fidelity HVAC system model is not a feasible solution. Having a set of DR sample data, Research Question 1 was developed to initiate analysis.

Research Question 1: *How can the benefits of DR for the entire campus buildings be assessed with a small number of sample buildings?*

To answer this question, historical data has been thoroughly examined, and two sub research questions have been formulated. The first one seeks an efficient modeling technique using data to predict not-observed phenomena.

Research Question 1.1: *How can we create a credible enough model to quantify the cooling load curtailment by DR where small sets of data for DR events exist?*

An observation has been made from data that there are similar environmental conditions in DR and non-DR days. This led to Hypothesis 1.1.

Hypothesis 1.1: *If a parametric model is built with non-DR data from similar environmental conditions as a function of weather and time parameters, then the fitted*

model will predict the original load during the DR active period. DR load shed can be discerned by the difference between the model predicted and measured load.

The second question asks if the observation from the sample can be generalized.

Research Question 1.2: *Is it possible to build a generalized DR model to predict the DR effectiveness of other buildings with sample data?*

Inconsistent claims about the relationship between the amount of load shed and the environmental parameters are found from literature to explain the DR potential. A hypothesis is made that the reason for the disagreement is that the uncertainty is too high to reveal a true trend if it exists. This statement can be tested using statistical analysis.

Hypothesis 1.2: *No apparent correlation will be observed regardless of the true relationship due to high uncertainty.*

To test the hypotheses and answer the research questions, a test methodology is developed in Section 5.1 and conducted two experiments. The ANN model is used to create the baseline load, and it can successfully discern the load shed by DR. Using this model, data has been generated to test Hypothesis 1.2. As claimed, no correlation or very weak correlations were observed from the generated data set. Having validated hypotheses, a parametric model of DR is proposed as a random variable.

The benefits of DR have been investigated from two different perspectives, and they are electricity bill savings and the operational availability improvement. To quantify those, the chiller plant model is needed. To overcome the complexity of the system and computational burden, Research Question 2 seeks a simplified version of control logic to evaluate the chilled water system's operating cost and availability.

Research Question 2: *How can the benefits of DR for a chiller plant be assessed without implementing a complex control logic?*

Hypothesis 2 states about the modeling approach and the approximation approach.

Hypothesis 2: *If the optimal control design can be approximated as a set of the best combinations per loading conditions to minimize power consumptions, then it can predict changes of operational status of a chiller plant without a sophisticated control logic*.

To test this hypothesis, a chilled water network system model is developed and validated in Section 5.2. It enables to generate data for any operating conditions. An experiment has been performed to support the hypothesis.

The plant approximation model can qualitatively evaluate the DR impact on the plant operational availability. The need for a quantitative assessment of the system reliability led to Research Question 3.

Research Question 3: *How can the performance of DR as a contingency plan for a district chilled water system be measured and fairly compared to the deterministic redundancy?*

Rather than formulating a hypothesis, this research question was answered by a literature survey to evaluate and compare the redundancy and the DR. Capacity adequacy evaluation metrics for the power grid generation systems have been borrowed and slightly modified to evaluate the reliability of a district energy system. Because of the characteristics of thermal energy generation and distribution, there is no blackout type failure in district energy systems. Therefore, the failure is defined by an insufficiency in thermal and mass capacity. The plant availability simulation model is developed and verified in Section 5.3.

The challenge of the problem formulation is to articulate a cost function because the operating cost is very complex and requires a control logic and simulation. It is time consuming and may be not worthwhile. The capacity expansion planning problem is formulated as a MILP formulation based on some assumptions addressed in Section 3.4. To test this assumption, the plant approximation model developed in Section 5.2 is utilized. The experiment conducted in Section 5.4 proved the assumption.

The constituent models developed in Chapter 5 are integrated into a methodology.

Research Question 4: *How can we predict if DR is capable of avoiding or deferring capacity expansion without sacrificing reliability?*

The optimal capacity expansion planning problem was formulated as MILP to answer Research Question 4. This research question is further developed into the sub research question. The capacity expansion planning problem is formulated as a single objective function while the true attribute of the problem is a multi-objective decision making problem. The proposed solution approach is developed into a two-step analysis. This led to Research Question 4.1.

Research Question 4.1: *How can we know the optimality of the solution for the capacity expansion plan with the proposed decision making framework?*

Section 3.4.2 developed the thought on constructing Pareto Frontier of a multi-objective decision making problem and how the proposed method improves the plan in two directions. This thought was summarized into Hypothesis 4.1.

Hypothesis 4.1: *With the two extreme points of the N+1 constraint and no safety margin, the Pareto frontier can be generated by adjusting safety margin on the demand satisfaction constraint between the two points and the fraction of DR buildings.*

Finally, the proposed methodology is implemented with an example problem based on the real system to test the last Hypothesis in Chapter 6.

Hypothesis 4: *The proposed two-step decision making framework can yield multiple Pareto solutions and the corresponding optimal capacity expansion planning schedule. Using those, one can predict the impact of DR on the capacity expansion planning and the system reliability.*

The resultant Pareto Frontier plots aid to discover a new trend in the improvement of the reliability by DR and the safety margin. Findings from demonstrating the methodology are listed below.

- The price based Demand Response activation time is not synchronized with the campus' peak cooling load time
- The amount of load shed by DR is enough to affect the plant operational state
- Confirmed that the N+1 system is highly reliable by simulation
- There is a point where no more improvement of the reliability can be made by increasing the safety margin, which is less than the capacity of a redundant chiller
- This supports the claim that the N+1 is an oversized system
- Having DR as a contingency operation option, the capacity expansion plan can be deferred to several years and save million dollars while maintaining a comparable level of reliability of the N+1 system

7.2 Contributions

The aim of this research was to examine the value of DR in operation and management of district energy systems. The main contribution of this work is the development of a structured and transparent framework for forecasting of the effectiveness of DR. The proposed methodology utilizes measured data, converts them into actionable information by developing models to capture the interaction between demand and supply sides, and provides an insight into the planning design space by connecting the planning optimization and the reliability analysis modules. The flow of data and information is provided throughout the steps in the methodology. Even though DR is not a new concept, it is the first attempt to evaluate DR as a contingency plan for a district energy system in place of a redundant chiller.

The second contribution is made from characterizing the DR load shed as a random variable using a data driven cooling demand modeling approach so that it allows to predict the effectiveness of DR in a probabilistic manner. Uncertainty in demand prediction is too large to reveal any trend of DR load shed correlated with environmental parameters. Previous research overlooked the uncertainty in the effectiveness of DR so that their model failed to predict a trend. A statistical analysis of transformed data confirmed the claim. The generalized DR model is flexible enough to be expanded to other buildings. It fills the gap between the high fidelity HVAC system simulation modeling approach and the high level conceptual model in terms of accuracy and scalability.

Exploration of the effectiveness of DR for a district level chiller plant operation is another contribution. Approximation of the chiller staging sequence enabled the calculation of cost savings and the prediction of capacity availability improvement without developing a sophisticated control system and running a computationally expensive

simulation. The overall analysis framework helps a decision maker to estimate the benefits of DR for the community in two perspectives: the electrical energy consumer and the thermal energy producer.

7.3 Future Work

A few recommendations have been made for potential follow up research. This research is based on the measured data. However, only a small set of data was available for DR events. It is worthwhile to examine further the correlation of data when more samples are available. A new trend may be discovered with more accurate and ample data set. Due to the small sample size, it was assumed that every building has the same effectiveness. Differentiation of building types in modeling and analysis can be investigated, and it enables to optimize the selection of buildings to apply DR. Other than a specific DR method, the same data driven modeling and analysis framework is applicable to other established building energy technologies to evaluate and forecast the effectiveness of the technology.

More research can be done on the capacity expansion planning problem formulation. Assumptions are made to simplify the cost function. It limits the applicability of the proposed MILP formulation. It cannot be used where the operating cost is significant such as designing a new plant. Another formulation is needed to account for the eliminated complexity. Co-optimization of capacity expansion planning and DR controls can extend the MILP formulation to stochastic programming. The nature of DR, which is a corrective action in contingency events, fits well for two stage programming philosophy. However, it requires enumeration of scenarios. This would be challenging.

Uncertainty in the timeline of the new buildings can be incorporated. Currently, only the peak load estimation uncertainty is considered to generate demand growth scenarios. Investigation of the sensitivity to the horizontal variability will be interesting.

APPENDIX A

CHILLER PLANT MODEL VALIDATION

Modelica *Buildings* library developed by LBL [46] is used to model components of the 10th St. plant and chilled water network system. Chillers are modeled as DOE-2 electric chiller model [90] that has three curves as functions of chilled water supply temperature and condenser water entering temperature to calculate electric power consumptions at the full or part load conditions. Coefficients of the model are calibrated using measured data of GT campus chilled water system. Calibration process will be explained further below. Parameters of other components are also calibrated to represent the real GT campus as much as possible based on information availability. The cooling demands are modeled as two lumped loads representing east and west campus using the data driven NN model.

The DOE-2 chiller model uses chilled water supply temperature and condenser water entering temperature as inputs to calculate electric power consumption at specific operating set points. It consists of three curves.

- CAPFT: available capacity as a function of temperatures
- EIRFT: energy input ratio as a function of temperatures
- EIRFTPLR: energy input ratio as a function of part load ratio

Those three curves are fitted as second order functions as follows:

$$\text{CAPFT} = a_1 + b_1 \times T_{CHW,S} + c_1 \times {T_{CHW,S}}^2 + d_1 \times T_{CW,S} + e_1 \times {T_{CW,S}}^2 + f_1 \times T_{CHW,S} T_{CW,S}$$

$$\text{EIRFT} = a_2 + b_2 \times T_{CHW,S} + c_2 \times {T_{CHW,S}}^2 + d_2 \times T_{CW,S} + e_2 \times {T_{CW,S}}^2 + f_2 \times T_{CHW,S} T_{CW,S}$$

$$\text{EIRFTPLR} = a_3 + b_3 \times \text{PLR} + c_3 \times \text{PLR}^2$$

$$\text{PLR} = \frac{Q}{Q_{ref} \times CAPFT\left(T_{CHW,S}, T_{CW,S}\right)}$$

Where,

$T_{CHW,S}$: Chilled water supply temperature

$T_{CW,S}$: Condenser water supply temperature (entering to condenser)

Q: Capacity (cooling tons)

Q_{ref}: Capacity at the reference condition

PLR: Part load ratio

$a_i, b_i, c_i, d_i, e_i, and\ f_i$: Coefficients

The standard least squares linear regression method suggested by [89] is used to determine coefficients of the individual chiller model. This method requires relatively large data sets at full load and part load conditions separately. Monitored data are available for the 6 chillers in wide range of capacity. In case full load monitored data is not available, manufacture's performance data is used. Data pre-processing and curve fitting are performed using JMP.

Calibration method is following these steps:

1. Data collection
2. Data pre-processing
    a. Remove outliers
    b. Choose sample to train and test
3. Calculate intermediate values of CAPFT and EIRFT

$$\text{CAPFT}_i = \frac{Q_i}{Q_{ref}}$$

$$\text{EIRFT}_i = \frac{\frac{P_i}{Q_i}}{\frac{P_{ref}}{Q_{ref}}} = \frac{P_i}{P_{ref} \times CAPFT_i}$$

4. Divide data into full load & part load
    a. Full load: CAPFT>0.9, part load: CAPFT<0.9
    b. When there is no available full load in monitored data then use manufacture's data
5. Least square fit to determine the coefficients of CAPFT and EIRFT with only full load data set
6. Calculate PLR & EIRFT for the combined data set

$$\text{PLR}_i = \frac{Q_i}{Q_{ref} \times CAPFT_i(T)}$$

$$\text{EIRFPLR}_i = \frac{P_i}{P_{ref} \times CAPFT_i(T) \times EIRFT_i(T)}$$

7. Least square fit for EIRFTPLR
8. Validate the fitted model by calculating an error metric for the predicted power

$$\text{CVRMSE} = \frac{\sqrt{\frac{\sum_{i=1}^{N}\left(P_{predicted} - P_i\right)^2}{N}}}{\overline{P_i}}$$

$$P_{predicted} = P_{ref} \times CAPFT_i \times EIRFT_i \times EIRFPLR_i$$

$$\overline{P_i} = \frac{\sum_{i=1}^{N} P_i}{N}$$

Technical information of chillers is listed in Table 17. Specific manufacture's information is not available for Chiller 5. Data is used to guess the reference power of Chiller 5. The Chiller 5 model is not used for this thesis.

Table 17 10th St. Plant Chiller Information

Chiller Name	Manuf.	Type	Capacity (Tons)	Reference power (kW)	Unit Perf (kw/Ton)	COP	Nominal Flow Rate (GPM)
Chiller 1	York	Centrifugal	1500	921	0.614	5.728	3600
Chiller 2	York	Centrifugal	1500	921	0.614	5.728	3600
Chiller 3	York	Centrifugal	1978	1293	0.654	5.378	4000
Chiller 4	York	Centrifugal	1978	1293	0.654	5.378	4000
Chiller 5	McQuay	Dual Centrifugal	2250	Unknown	unknown	unknown	4000(?)
Chiller 6	York	Dual Centrifugal	3000	2026	0.675	5.210	6000

Data used for training and test are explained in Table 18. Data is randomly sampled from the specified date range for the training. The fitted model is tested using data from different data range to validate the predictive capability of the model. The error metric is presented in Table 19. The models fit well as CV-RMSE is smaller than 30%, which is the acceptance limit of the ASHRAE guideline.

Table 18 Data Sample Used for Model Calibration

	Training				Test	
	Date Range	Full Load	Part Load	Total	Date Range	No.
Chiller 1	Manufactures(full) 2011 June-Oct (part)	25	100	125	2013 June-Oct	5777
Chiller 2	Manufactures + 2013(full) 2011 June	36	101	137	2013 June	2106
Chiller 3	2011 Apr - Oct	85	90	175	2013 June	1463
Chiller 4	2011 Apr - Sep	100	150	250	2013 June	2408
Chiller 5 Load	2011 Jan-Nov	100	150	250	2013 June-July	5182
Chiller 5 Tons	2011 Jan- Nov	86	98	184	2013 June-July	5007
Chiller 6	2012 Mar - Nov	100	180	280	2013 June - July	4206

Table 19 Model Errors

	Training		Test	
	N	CVRMSE(%)	N	CVRMSE(%)
Chiller 1	125	5.34	5777	7.50
	125	4.72	5777	8.46
Chiller 2	134	3.43	2106	4.43
	137	5.31	2106	6.16
Chiller 3	175	3.96	1463	3.58
Chiller 4	250	3.40	2408	2.63
Chiller 5	250	9.88	5182	23.84
	184	13.83	5007	21.91
Chiller 6	280	3.67	4206	14.65

APPENDIX B

DEMAND RESPONSE EVENT LOG

DR event log recorded in 2016 summer is presented in Table 20. The start and end time of each event are provided in EDT (Eastern Daylight Time), which is four hours behind UTC (Coordinated Universal Time).

Table 20 DR Event Log

DR Event	Start time	End time
1	2016-07-21 14:40:46-04:00	2016-06-13 18:55:31-04:00
2	2016-06-17 14:40:20-04:00	2016-06-17 16:55:25-04:00
3	2016-06-24 14:40:19-04:00	2016-06-24 18:55:30-04:00
4	2016-06-25 16:40:23-04:00	2016-06-25 18:55:25-04:00
5	2016-06-27 16:40:20-04:00	2016-06-27 18:55:25-04:00
6	2016-07-07 16:40:37-04:00	2016-07-07 18:24:58-04:00
7	2016-07-08 14:40:42-04:00	2016-07-08 19:55:37-04:00
8	2016-07-14 14:40:28-04:00	2016-07-14 17:55:33-04:00
9	2016-07-18 15:40:41-04:00	2016-07-18 18:55:32-04:00
10	2016-07-19 14:40:35-04:00	2016-07-19 18:55:35-04:00
11	2016-07-20 14:40:38-04:00	2016-07-20 19:55:29-04:00
12	2016-07-21 14:40:46-04:00	2016-07-21 17:55:52-04:00
13	2016-07-25 13:40:45-04:00	2016-07-25 19:55:51-04:00
14	2016-07-26 14:40:49-04:00	2016-07-26 18:55:40-04:00
15	2016-08-16 14:40:44-04:00	2016-08-16 17:55:44-04:00
16	2016-08-18 14:41:13-04:00	2016-08-18 17:57:14-04:00
17	2016-08-23 15:40:36-04:00	2016-08-23 16:55:41-04:00
18	2016-08-31 16:40:45-04:00	2016-08-31 17:55:30-04:00

APPENDIX C

CLIMATE CHANGE MODEL

For those who are interested, a climate model is presented here. Annual climate data is essential for the building energy simulation because weather is a main driver for the dynamics of energy usage. Typical Meteorological Year (TMY) has been used as the climate model for year. It consists of data sets that represent climate of specific location and contain hourly weather parameters such as temperature, humidity, wind speed, solar radiation, etc. These data are not statistics from previous years. It selects actual time series data from the existing pool. TMY data was initially constructed in 1978 from historical weather database and periodically updated [98]. The original TMY data were collected from 1952-1975 weather data, the second version (TMY2) is updated with data from 1961-1990, and the current version (TMY3) uses 1976-2005 data.

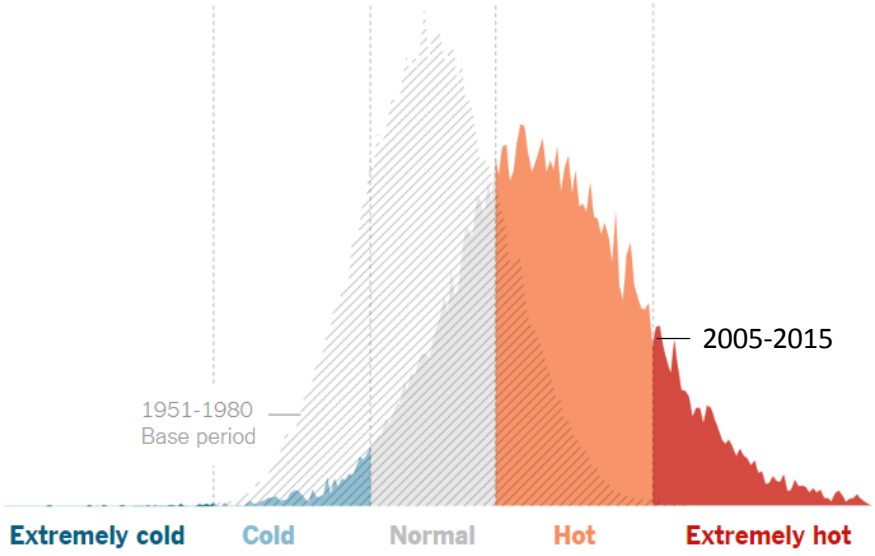


Figure 120 Summer Temperature in Northern Hemisphere (June to August)

[99]

Concern is raising about using TMY data for the simulation of heating and cooling building energy use due to climate change because even the most recent TMY3 is outdated [100]. Recent data shows how recent summer weather is hotter than before. In recent news article [99], how summer has become extremely hot is compared to a few decades ago based on an academic research [101]. As can be seen in Figure 120, the frequency of occurrence plot of summer temperatures is shifted to the right and more dispersed. Recent studies emphases on the risk of using historical data to estimate the future building energy usage [102]–[104].

There are some tools available to public to create new climate models. They use different emission scenarios for the future to create modified climate models based on historical data. CC world weather file generator is an open source tool based on research [105], [106]. This uses TMY data to create a morphed future weather file in consideration of the climate change. Resulting file has the same format of TMY data file to be directly used for the building simulation programs such as EnergyPlus. There is another similar web based tool called WeatherShift™ [107], [108]. This provides a capability to choose location and emission scenarios by user and plots for the future weather data. It is useful to see how annual weather will be changed, but data should be purchased.

How the climate change affect the cooling load is shown here. The CC world weather file generator is selected to create the future climate data because it is publicly available. Figure 121 shows that the simulated future climate data is shifted towards the right side from the current climate model. The temperature is higher at the same day while the relative humidity is slightly reduced. The ANN model for campus load is to simulate the future cooling load. The future cooling load in 2032 is compared to the current year's cooling load profile in Figure 122. Their profile are the same, but the cooling load in 2032 slightly shift up. The peak cooling load projection using the ANN and the climate model

is presented in Figure 123. Each year's temperature and relative humidity are generated by interpolation between the current year and the available future climate model because the CC world weather file generator provides only 2020 and 2050 weather profiles. The peak load keeps increasing due to the climate change. Irregularity in the projection is created because the ANN cooling load model takes time variables as inputs. Day of week and day of year would affect the cooling load besides weather variables.

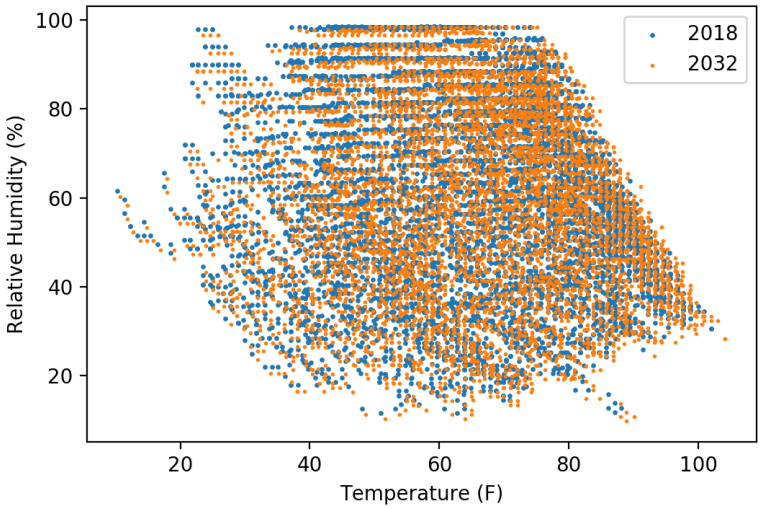


Figure 121 Future Climate Data

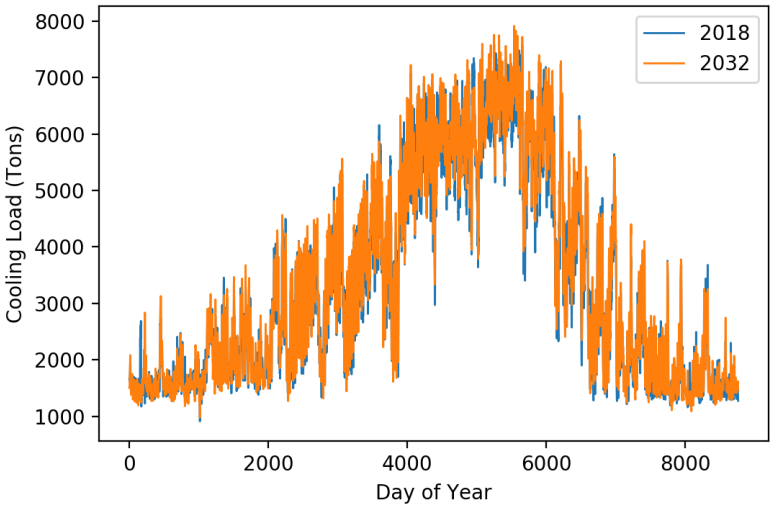


Figure 122 Future Cooling Load

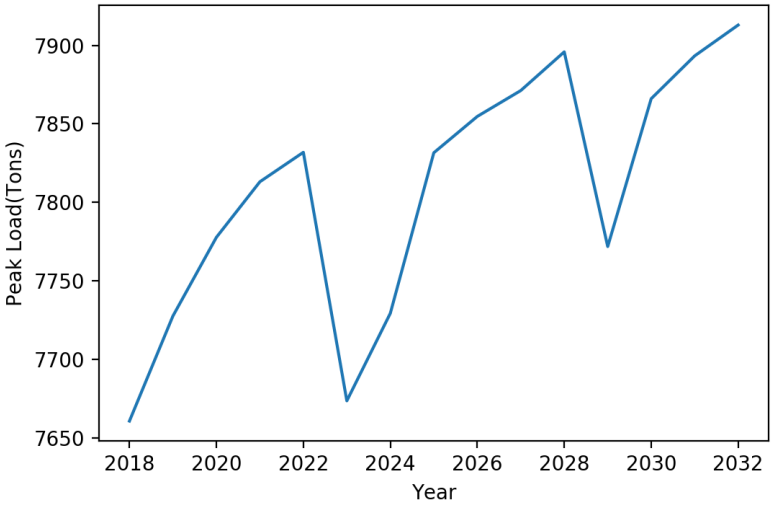


Figure 123 Peak Cooling Load Projection by Climate Model

REFERENCES

[1] "How much energy is consumed in residential and commercial building in the United States?" [Online]. Available: http://www.eia.gov/tools/faqs/faq.cfm?id=86&t=1. [Accessed: 09-Aug-2015].

[2] U.S. Energy Information Agency, "International Energy Outlook 2013," *Outlook 2013*, p. 312, 2013.

[3] W. Goetzler, M. Guernsey, and J. Young, "Research and Development Roadmap for Emerging HVAC Technologies," 2014.

[4] C. W. Gellings, "The concept of demand-side management for electric utilities," *Proc. IEEE*, vol. 73, no. 10, pp. 1468–1470, 1985.

[5] P. Chen, B. Bak-Jensen, and Z. Chen, "Probabilistic load models for simulating the impact of load management," in *2009 IEEE Power & Energy Society General Meeting*, 2009, pp. 1–8.

[6] P. Palensky and D. Dietrich, "Demand Side Management: Demand Response, Intelligent Energy Systems, and Smart Loads," *IEEE Trans. Ind. Informatics*, vol. 7, no. 3, pp. 381–388, Aug. 2011.

[7] U S Department of Energy, "Benefits of Demand Response in Electricity Markets and Recommendations for Achieving Them," 2006.

[8] N. Motegi, M. A. Piette, D. S. D. Watson, S. Kiliccote, and P. Xu, "Introduction to Commercial Building Control Strategies and Techniques for Demand Response," 2007.

[9] W. Gang, S. Wang, F. Xiao, and D. Gao, "District cooling systems: Technology integration, system optimization, challenges and opportunities for applications," *Renew. Sustain. Energy Rev.*, vol. 53, pp. 253–264, Jan. 2016.

[10] "How District Energy Works." [Online]. Available: http://enwaveusa.com/content/how_district_energy_works-38768.html. [Accessed: 01-Jan-2015].

[11] F. Wernstedt, P. Davidsson, and C. Johansson, "Demand side management in district heating systems," in *Proceedings of the 6th international joint conference on Autonomous agents and multiagent systems - AAMAS '07*, 2007, vol. 5, p. 1.

[12] M. E. H. Dyson, S. D. Borgeson, M. D. Tabone, and D. S. Callaway, "Using smart meter data to estimate demand response potential, with application to solar energy integration," *Energy Policy*, vol. 73, pp. 607–619, Oct. 2014.

[13] "12. District heating and cooling," in *2012 ASHRAE handbook heating, ventilating, and air-conditioning systems and equipment*, American Society of Heating, Refrigerating and Air-Conditioning Engineers, Inc., 2012.

[14] S. Doty, "Designing for Reliability: Identify Weak Spots by Asking 'What Can Go Wrong?,'" *ASHRAE J.*, vol. 50, no. 7, pp. 1–6, 2008.

[15] P. Huang, G. Huang, and Y. Wang, "HVAC system design under peak load prediction uncertainty using multiple-criterion decision making technique," *Energy Build.*, vol. 91, pp. 26–36, Mar. 2015.

[16] J. B. Belcher, "SIZING OPTIMIZATION OF A CENTRAL CHILLED WATER SYSTEM." [Online]. Available: http://www.thermalegi.com/images/JBB-IDEA-CWSizing.pdf. [Accessed: 12-Jun-2017].

[17] K. Orehounig, G. Mavromatidis, R. Evins, V. Dorer, and J. Carmeliet, "Towards an energy sustainable community: An energy system analysis for a village in Switzerland," *Energy Build.*, vol. 84, pp. 277–286, Dec. 2014.

[18] S. Pfenninger, A. Hawkes, and J. Keirstead, "Energy systems modeling for twenty-first century energy challenges," *Renew. Sustain. Energy Rev.*, vol. 33, pp. 74–86, 2014.

[19] "EnergyPlus." [Online]. Available: http://apps1.eere.energy.gov/buildings/energyplus/. [Accessed: 30-Nov-2015].

[20] M. Hydeman and G. Zhou, "Optimizing chilled water plant control," *ASHRAE J.*, vol. 49, no. 6, pp. 44–54, 2007.

[21] Y. Sun, S. Wang, and G. Huang, "Chiller sequencing control with enhanced robustness for energy efficient operation," *Energy Build.*, vol. 41, no. 11, pp. 1246–1255, Nov. 2009.

[22] S. T. Taylor, "Optimizing Design & Control Of Chilled Water Plants Part 1: Chilled Water Distribution System Selection," *ASHRAE J.*, 2011.

[23] S. T. Taylor, "Primary-only vs. primary-secondary variable flow systems," *ASHRAE J.*, vol. 44, no. 2, pp. 25–29, 2002.

[24] S. T. Taylor, "Degrading Chilled Water Plant Delta-T : Causes and Mitigation," 2002.

[25] "CoolTools™ Chilled Water Plant Design Guide," no. December. Taylor Engineering, 2009.

[26] *2013 ASHRAE Handbook - Fundamentals (I-P Edition)*. American Society of Heating, Refrigerating and Air-Conditioning Engineers, Inc., 2013.

[27] M. Wetter and C. Haugstetter, "Modelica versus TRNSYS–A comparison between an equation-based and a procedural modeling language for building energy simulation," in *SimBuild IBPSA-USA*, 2006, vol. 11, no. 4, pp. 262–269.

[28] "DOE2." [Online]. Available: http://www.doe2.com/. [Accessed: 30-Nov-2015].

[29] D. Coakley, P. Raftery, and M. Keane, "A review of methods to match building energy simulation models to measured data," *Renew. Sustain. Energy Rev.*, vol. 37, pp. 123–141, Sep. 2014.

[30] Y. Zhang, Z. O'Neill, B. Dong, and G. Augenbroe, "Comparisons of inverse modeling approaches for predicting building energy performance," *Build. Environ.*, vol. 86, pp. 177–190, Apr. 2015.

[31] P. Raftery, M. Keane, and A. Costa, "Calibrating whole building energy models: Detailed case study using hourly measured data," *Energy Build.*, vol. 43, no. 12, pp. 3666–3679, 2011.

[32] Y. Sun, L. Gu, C. F. F. J. Wu, and G. Augenbroe, "Exploring HVAC system sizing under uncertainty," *Energy Build.*, vol. 81, pp. 243–252, Oct. 2014.

[33] R.-R. Schmidt and D. Basciotti, "Demand Side Management in District Heating Systems." 2013.

[34] N. O'Connell, P. Pinson, H. Madsen, and M. O'Malley, "Benefits and challenges of electrical demand response: A critical review," *Renew. Sustain. Energy Rev.*, vol. 39, pp. 686–699, Nov. 2014.

[35] R. Earle, E. P. Kahn, and E. Macan, "Measuring the Capacity Impacts of Demand Response," *Electr. J.*, vol. 22, no. 6, pp. 47–58, 2009.

[36] C. Johansson, F. Wernstedt, and P. Davidsson, "Deployment of Agent Based Load Control in District Heating Systems," *First Int. Work. Agent Technol. Energy Syst. (ATES 2010)*, pp. 75–82, 2010.

[37] D. Basciotti and R. R. Schmidt, "Peak Reduction in District Heating Networks : a Comparison Study and Practical Considerations," in *The 14th International Symposium on District Heating and Cooling*, 2014.

[38] H. Li and S. J. Wang, "Load Management in District Heating Operation," *Energy Procedia*, vol. 75, pp. 1202–1207, 2015.

[39] H. a. Aalami, M. P. Moghaddam, and G. R. Yousefi, "Modeling and prioritizing demand response programs in power markets," *Electr. Power Syst. Res.*, vol. 80, no. 4, pp. 426–435, Apr. 2010.

[40] C. De Jonghe, B. F. Hobbs, and R. Belmans, "Integrating short-term demand response into long-term investment planning," 2011.

[41] N. Zhang, Z. Hu, C. Springer, Y. Li, and B. Shen, "A bi-level integrated generation-transmission planning model incorporating the impacts of demand response by operation simulation," *Energy Convers. Manag.*, vol. 123, pp. 84–94, Sep. 2016.

[42] D. Huang and R. Billinton, "Effects of Load Sector Demand Side Management Applications in Generating Capacity Adequacy Assessment," *IEEE Trans. Power Syst.*, vol. 27, no. 1, pp. 335–343, Feb. 2012.

[43] J. H. Dudley, D. Black, M. Apte, M. A. Piette, and P. Berkeley, "Comparison of Demand Response Performance with an EnergyPlus Model in a Low Energy Campus Building,"

2010.

[44] G. Bianchini, M. Casini, A. Vicino, and D. Zarrilli, "Demand-response in building heating systems: A Model Predictive Control approach," *Appl. Energy*, vol. 168, pp. 159–170, 2016.

[45] F. Sehar, M. Pipattanasomporn, and S. Rahman, "A peak-load reduction computing tool sensitive to commercial building environmental preferences," *Appl. Energy*, vol. 161, pp. 279–289, Jan. 2016.

[46] M. Wetter, W. Zuo, T. S. Nouidui, and X. Pang, "Modelica Buildings library," *J. Build. Perform. Simul.*, vol. 7, no. 4, pp. 253–270, Jul. 2014.

[47] D. H. Blum and L. K. Norford, "DYNAMIC SIMULATION OF REGULATION DEMAND RESPONSE BY VAV HVAC SYSTEMS," in *ASHRAE/IBPSA-USA Building Simulation Conference*, 2014, pp. 402–409.

[48] M. A. Piette, J. Granderson, M. Wetter, S. Kiliccote, and L. Berkeley, "Responsive and Intelligent Building Information and Control for Low-Energy and Optimized Grid Integration Energy Issues in Design , Operations , and Grid Integration," in *ACEEE Summer Study on Energy Efficiency in Buildings*, 2012, pp. 236–249.

[49] R. Yin, E. C. Kara, Y. Li, N. DeForest, K. Wang, T. Yong, and M. Stadler, "Quantifying flexibility of commercial and residential loads for demand response using setpoint changes," *Appl. Energy*, vol. 177, pp. 149–164, Sep. 2016.

[50] S. Kiliccote, M. a Piette, J. Mathieu, and K. Parrish, "Findings from Seven Years of Field Performance Data for Automated Demand Response in Commercial Buildings," 2010.

[51] J. L. Mathieu, D. S. Callaway, and S. Kiliccote, "Examining uncertainty in demand response baseline models and variability in automated responses to dynamic pricing," in *IEEE Conference on Decision and Control and European Control Conference*, 2011, pp. 4332–4339.

[52] P. N. Price, N. Addy, and S. Kiliccote, "Predictability and Persistance of Demand Response Load Shed in Buildings," 2015.

[53] A. Birolini, *Reliability Engineering*, vol. 10, no. 4. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014.

[54] IAEA, *Expansion Planning for Electrical Generating Systems: A Guidebook*, no. 241. 1984.

[55] R. Billinton and W. Li, *Reliability Assessment of Electric Power Systems Using Monte Carlo Methods*. Boston, MA: Springer US, 1994.

[56] M. Milligan and K. Porter, "Determining the Capacity Value of Wind : An Updated Survey of Methods and Implementation Preprint," in *Wind Power*, 2008.

[57] W. Gang, S. Wang, G. Augenbroe, and F. Xiao, "Robust optimal design of district cooling systems and the impacts of uncertainty and reliability," *Energy Build.*, vol. 122, pp. 11–22,

2016.

[58] R. D. Prasad, R. C. Bansal, and A. Raturi, “Multi-faceted energy planning: A review,” *Renew. Sustain. Energy Rev.*, vol. 38, pp. 686–699, Oct. 2014.

[59] R. Thery and P. Zarate, “Energy planning: A multi-level and multicriteria decision making structure proposal,” *Cent. Eur. J. Oper. Res.*, vol. 17, no. 3, pp. 265–274, Jul. 2009.

[60] H. Luss, “Operations Research and Capacity Expansion Problems: A Survey,” *Oper. Res.*, vol. 30, no. 5, pp. 907–947, Oct. 1982.

[61] V. Oree, S. Z. Sayed Hassen, and P. J. Fleming, “Generation expansion planning optimisation with renewable energy integration: A review,” *Renew. Sustain. Energy Rev.*, vol. 69, pp. 790–803, Mar. 2017.

[62] S. D. Pohekar and M. Ramachandran, “Application of multi-criteria decision making to sustainable energy planning—A review,” *Renew. Sustain. Energy Rev.*, vol. 8, no. 4, pp. 365–381, Aug. 2004.

[63] E. Løken, “Use of multicriteria decision analysis methods for energy planning problems,” *Renew. Sustain. Energy Rev.*, vol. 11, no. 7, pp. 1584–1595, Sep. 2007.

[64] M. Bojesen, L. Boerboom, and H. Skov-Petersen, “Towards a sustainable capacity expansion of the Danish biogas sector,” *Land use policy*, vol. 42, pp. 264–277, Jan. 2015.

[65] Y. T. Leong, R. R. Tan, K. B. Aviso, and I. M. L. Chew, “Fuzzy analytic hierarchy process and targeting for inter-plant chilled and cooling water network synthesis,” *J. Clean. Prod.*, vol. 110, pp. 40–53, Jan. 2016.

[66] D. Chinese, G. Nardin, and O. Saro, “Multi-criteria analysis for the selection of space heating systems in an industrial building,” *Energy*, vol. 36, no. 1, pp. 556–565, Jan. 2011.

[67] S. Ghafghazi, T. Sowlati, S. Sokhansanj, and S. Melin, “A multicriteria approach to evaluate district heating system options,” *Appl. Energy*, vol. 87, no. 4, pp. 1134–1140, Apr. 2010.

[68] C. Martínez-Costa, M. Mas-Machuca, E. Benedito, and A. Corominas, “A review of mathematical programming models for strategic capacity planning in manufacturing,” *Int. J. Prod. Econ.*, vol. 153, pp. 66–85, Jul. 2014.

[69] G. A. Bakirtzis, P. N. Biskas, and V. Chatziathanasiou, “Generation Expansion Planning by MILP considering mid-term scheduling decisions,” *Electr. Power Syst. Res.*, vol. 86, pp. 98–112, May 2012.

[70] N. E. Koltsaklis and M. C. Georgiadis, “A multi-period, multi-regional generation expansion planning model incorporating unit commitment constraints,” *Appl. Energy*, vol. 158, pp. 310–331, Nov. 2015.

[71] B. F. Hobbs, “Optimization methods for electric utility resource planning,” *Eur. J. Oper. Res.*, vol. 83, no. 1, pp. 1–20, May 1995.

[72] C. H. Antunes, A. G. Martins, and I. S. Brito, “A multiple objective mixed integer linear programming model for power generation expansion planning,” *Energy*, vol. 29, no. 4, pp. 613–627, Mar. 2004.

[73] T. Lohmann and S. Rebennack, “Tailored Benders Decomposition for a Long-Term Power Expansion Model with Short-Term Demand Response,” *Manage. Sci.*, p. mnsc.2015.2420, Jun. 2016.

[74] N. E. Koltsaklis, P. Liu, and M. C. Georgiadis, “An integrated stochastic multi-regional long-term energy planning model incorporating autonomous power systems and demand response,” *Energy*, vol. 82, pp. 865–888, Mar. 2015.

[75] S. Jin, A. Botterud, and S. M. Ryan, “Impact of Demand Response on Thermal Generation Investment With High Wind Penetration,” *IEEE Trans. Smart Grid*, vol. 4, no. 4, pp. 2374–2383, Dec. 2013.

[76] Long Zhao and Bo Zeng, “Robust unit commitment problem with demand response and wind energy,” in *2012 IEEE Power and Energy Society General Meeting*, 2012, pp. 1–8.

[77] C. Wang, B. Jiao, L. Guo, Z. Tian, J. Niu, and S. Li, “Robust scheduling of building energy system under uncertainty,” *Appl. Energy*, vol. 167, pp. 366–376, Apr. 2016.

[78] M. Milligan and M. O. Malley, “Stochastic Methods for Planning and Operating Power System with Large Amounts of Wind and Solar Power Preprint,” in *11th Annual International Workshop on Large-Scale Integration of Wind Power into Power Systems as well as on Transmission Networks for Offshore Wind Power Plants Conference*, 2012.

[79] W. Gandulfo, E. Gil, and I. Aravena, “Generation Capacity Expansion Planning under demand uncertainty using stochastic mixed-integer programming,” *2014 IEEE PES Gen. Meet. | Conf. Expo.*, vol. 2014–Octob, no. October, pp. 1–5, 2014.

[80] E. Hajipour, M. Bozorg, and M. Fotuhi-Firuzabad, “Stochastic Capacity Expansion Planning of Remote Microgrids With Wind Farms and Energy Storage,” *IEEE Trans. Sustain. Energy*, vol. 6, no. 2, pp. 491–498, Apr. 2015.

[81] A. Liu, B. Hobbs, J. Ho, J. McCalley, V. Krishnan, M. Shahidehpour, and Q. Zheng, “Co-optimization of Transmission and Other Supply Resources,” 2013.

[82] A. D. Belegundu and T. R. Chandrupatla, *Optimization Concepts and Applications in Engineering*. Upper Saddle River, N.J: Prentice Hall, 1999.

[83] X. Xi, R. Sioshansi, and V. Marano, “A stochastic dynamic programming model for co-optimization of distributed energy storage,” *Energy Syst.*, vol. 5, no. 3, pp. 475–505, Sep. 2014.

[84] W. B. Powell, A. George, H. Simão, W. Scott, A. Lamont, and J. Stewart, “SMART: A Stochastic Multiscale Model for the Analysis of Energy Resources, Technology, and Policy,” *INFORMS J. Comput.*, vol. 24, no. 4, pp. 665–682, Nov. 2011.

[85] E. a Feinberg and D. Genethliou, “Applied Mathematics for Restructured Electric Power Systems,” in *IEEE Transactions on Automatic Control*, vol. 50, no. 9, 2005, pp. 269–285.

[86] H. Chen, “A Multiscale Forecasting Methodology for Power Plant Fleet Management Hongmei Chen,” 2005.

[87] M. Espinoza and C. Joye, “Short-term load forecasting, profile identification, and customer segmentation: a methodology based on periodic time series,” *Power Syst. IEEE*, pp. 1–8, 2005.

[88] R. E. Walpole, R. H. Myers, S. L. Myers, and K. Ye, *Probability and Statistics for Engineers and Scientists*, 8th ed. Pearson Education, 2007.

[89] M. Hydeman and K. Gillespie, “Tools and techniques to calibrate electric chiller component models,” *ASHRAE Trans.*, pp. 1–9, 2002.

[90] D. A. York, E. F. Tucker, and C. C. Cappiello, “DOE-2 Reference Manual Version 2.1.” Los Almos Scientific Laboratory, 1980.

[91] M. G. Balchanos, J. Kim, S. J. Duncan, and D. N. Mavris, “Modeling and Simulation-based Analysis for Large Scale Campus Chilled Water Networks,” in *13th International Energy Conversion Engineering Conference*, 2015.

[92] Department of the Army, “TM 5-698-5 - Survey of Reliability and Availability Information for Power Distribution , Power Generation , and Heating , Ventilating and Air Conditioning (HVAC) Components for Commercial , Industrial , and Utility Installations,” no. July, 2006.

[93] “METASYS® BUILDING AUTOMATION SYSTEM.” [Online]. Available: http://www.johnsoncontrols.com/buildings/building-management/building-automation-systems-bas. [Accessed: 08-Apr-2018].

[94] “NOAA National Centers for Environmental Information.” [Online]. Available: https://www.ncdc.noaa.gov/. [Accessed: 08-Apr-2018].

[95] F. W. H. Yik, W. L. Lee, J. Burnett, and P. Jones, “Chiller plant sizing by cooling load simulation as a means to avoid oversized plant,” *HKIE Trans. Hong Kong Inst. Eng.*, vol. 6, no. 2, pp. 19–25, 1999.

[96] M. Trčka and J. L. M. Hensen, “Overview of HVAC system simulation,” *Autom. Constr.*, vol. 19, no. 2, pp. 93–99, 2010.

[97] Q. Cheng, S. Wang, C. Yan, and F. Xiao, “Probabilistic approach for uncertainty-based optimal design of chiller plants in buildings,” *Appl. Energy*, vol. 185, pp. 1613–1624, Jan. 2017.

[98] S. Wilcox and W. Marion, “Users manual for TMY3 data sets,” 2008.

[99] N. POPOVICH and A. PEARCE, “It’s Not Your Imagination. Summers Are Getting

Hotter.," *New York Times*, 2016.

[100] M. Drury and O. Dynamics, "Climate Change and its Effect on Weather Data How We Use Weather Data," in *ACEEE Summer Study on Energy Efficiency in Buildings*, 2016.

[101] J. Hansen, M. Satoa, R. Ruedy, G. A. Schmidtc, and C. Ken Lob, "Global Temperature in 2015," *Columbia University Earth Institute*, 2016. [Online]. Available: http://csas.ei.columbia.edu/2016/01/19/global-temperature-in-2015/.

[102] R. Schaeffer, A. S. Szklo, A. F. Pereira de Lucena, B. S. Moreira Cesar Borba, L. P. Pupo Nogueira, F. P. Fleming, A. Troccoli, M. Harrison, and M. S. Boulahya, "Energy sector vulnerability to climate change: A review," *Energy*, vol. 38, no. 1, pp. 1–12, Feb. 2012.

[103] D. J. Sailor, "Risks of summertime extreme thermal conditions in buildings as a result of climate change and exacerbation of urban heat islands," *Build. Environ.*, vol. 78, pp. 81–88, 2014.

[104] H. Wang and Q. Chen, "Impact of climate change heating and cooling energy use in buildings in the United States," *Energy Build.*, vol. 82, pp. 428–436, 2014.

[105] M. F. Jentsch, P. A. B. James, L. Bourikas, and A. B. S. Bahaj, "Transforming existing weather data for worldwide locations to enable energy and building performance simulation under future climates," *Renew. Energy*, vol. 55, pp. 514–524, 2013.

[106] M. F. Jentsch, A. S. Bahaj, and P. A. B. James, "Climate change future proofing of buildings-Generation and assessment of building simulation weather files," *Energy Build.*, vol. 40, no. 12, pp. 2148–2168, 2008.

[107] L. Troup, "MORPHING CLIMATE DATA TO SIMULATE BUILDING ENERGY CONSUMPTION Luke Troup and David Fannon," in *Building Performance Modeling Conference*, 2016, pp. 439–446.

[108] "WeatherShift™." [Online]. Available: http://www.weather-shift.com/.