Strategy and Modeling for Building DR Optimization

Richard Lau, Sami Ayyorgun, Siun Chuon Mau
Sharanya Eswaran, Archan Misra*
Telcordia Technologies Inc.,
Singapore Management University*

Steven Bushby, David Holmberg
National Institute of Standards and Technology

Abstract— **While it is well recognized that renewable microgrid generation and intelligent storage can significantly reduce a building's need for grid power and its peak loading, there is currently no sound and generalized approach to combine these two technologies. Further, loads are becoming increasingly responsive, in terms of both sheddability and shiftability. In this paper, we formulate the building energy management problem based on a demand-response (DR) adaptation framework that extends the traditional "supply-demand matching" to a "supply-store-demand-time-shift-utility adaptation" paradigm. Stochastic modeling of distributed-energy resources (DER) and measurement-based stochastic models of loads are used to approximately optimize building DR actions. Compared to systems that have no DR, the proposed optimization achieves savings in the range of approximately 35-70%, depending on the amount of energy storage, the flexibility of the loads, and the accuracy of the stochastic models.**

Keywords - demand response; optimization policy; local renewable; energy storage; load modeling; commercial buildings

I. INTRODUCTION

Demand Response (DR) in commercial buildings (such as office complexes, shopping malls or educational campuses) is an important problem for future Smart Grids, especially as such buildings constituted approximately 20% of the overall U.S. energy consumption and 35% of the total U.S. *electricity* consumption in 2006 [1]. Demand response is defined here as consumer loads changing consumption patterns in response to changes in the price of energy or other electricity grid events. This can include a response to reduce or defer energy consumption as well as to increase consumption. We believe that DR optimization of commercial buildings must, in the future, handle the opportunities and challenges that arise from two emerging realities:

1. The increasing penetration of renewable generation technologies (e.g., wind, PV) that exhibit stochastic variability in power output due to dynamic fluctuations in weather parameters.
2. The gradual deployment of localized storage (e.g., thermal storage demonstrations in [2]) as a means for absorbing transient mismatches between renewable supply and power demands.

We think that these developments will transform the decision making in DR from the current "supply-demand adjustment" paradigm to one of "**supply-store-demand–time-shift**", where the DR controller must not only decide on what (and by how much) to control building loads, but also decide whether and when to store energy for future use. We view this problem as one of "stochastic optimal control", where the control decisions aim to minimize the non-renewable energy consumption (or cumulative electricity charges) over a future time horizon (in this paper, we employ hourly decision adaptation using a 24 hour lookahead horizon) under uncertainty in future local production and demand.

In this paper, we first introduce a mathematical framework for DR optimization that explicitly models different types of building loads (such as sheddable, controllable and time-shiftable) and the stochastic uncertainty in renewable generation. We then view DR as a form of resource utility maximization problem and propose a policy-based approach to solve what is, in general, a complex mixed-integer optimization problem. Having built the time horizon-based utility maximization framework, we then study the performance of this optimization mechanism under a range of renewable generation, local storage and building load characteristics.

While investigating the relative benefits of incorporating variable local generation and storage in the DR framework, we realized that performance gains would depend on the ability to accurately model and predict the future demands of building loads (especially ones that constitute a major fraction of a commercial building's demand). Accordingly, we then focus on utilizing real-life power consumption traces to derive a measurement-based power consumption model of the Telcordia office building (whose HVAC chiller typically constitutes 30% to 40% of the building's electricity bill), and subsequently incorporate this model into our DR framework to help quantify the benefits that accrue from such model-based estimation of in-building loads. In particular, we shall show that a simple time-dependent first order Markov model of the chiller's load variation provides high predictive accuracy of a chiller's demand in the next 24 hours, and that the incorporation of such a model helps to decrease the grid electricity consumption.

A. Key Contributions of This Paper

We believe that our work makes the following key contributions:

1. It proposes a formal framework that views DR as a finite-horizon stochastic optimization problem that incorporates different types of building loads.
2. It uses the framework to study the benefits of having different amounts (as a fraction of the building load) of renewable local generation and storage.
3. It uses real measurement data to build and validate a model for a representative building load (an HVAC chiller), and demonstrates that such model-based prediction of future load behavior can lead to additional gains in DR optimization.

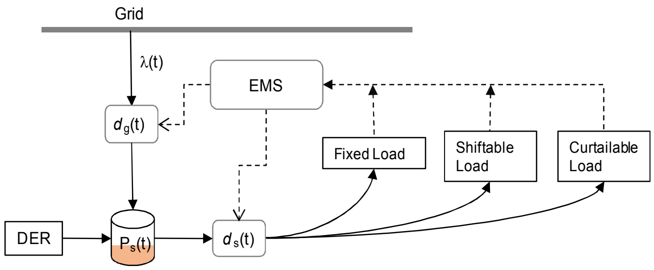

Figure 1. Demand Response framework

II. RELATED WORK

There has been a lot of work on *instantaneous* building DR optimization (at the distribution network or electric utility level), as opposed to optimization over a future time horizon. For example, the PowerMatcher framework [3] performs distributed DR adaptation using an electric utility framework, where loads and generation sources describe their functional responsiveness to price changes [4] and a centralized agent then sets an instantaneous price that matches supply and demand. For DR optimization over a time horizon, [5] applied a real-time dynamic minmax optimization technique to minimize a building's peak power consumption, while [6] applied a *Model Predictive Control* approach for energy generation optimization, where the uncertainty in the production profiles of renewables is captured over a finite look-ahead horizon and an explicit production adaptation strategy is derived that maximizes the expected reward over the whole horizon. More recently, [7] has focused on how a price-based framework may be used to manage such stochastically variable DERs with price-responsive loads. While such works treat resource fluctuations, they generally do not factor in the availabilities of local energy *storage* and load *shiftability*. In this paper, we formulate the DR optimization problem as a local decision problem, taking into account both local energy storage and load shiftability. Our approach is simple, and can be implemented in real time.

III. DEMAND-RESPONSE (DR) FRAMEWORK

A. Formulation

Fig. 1 shows the logical framework for a DR system. It consists of an Energy Management System (EMS), where Demand Response optimization is performed. The time (*t*) unit is discrete and can be in increments of any number of minutes, or hours. We assume a 24-hour (T) of optimization window, but a rolling time window can also be used. An energy storage device, either thermal or electrical, provides local storage ($P_s(t)$) of renewable energy (DER) and energy bought from the grid ($d_g(t)$). We will assume a simplified storage model in this paper, e.g. we ignore conversion losses among other things. Various loads are grouped according to their characteristics. In addition to traditional fixed loads, the shiftable loads can be served at a later time slot before the deadline (t_e). A shiftable load may also reduce the satisfaction proportional to the deferred duration. Another type of load is the curtailable load, which can be partially shed, but with a reduction in satisfaction depending on the curtailed amount. An example of a curtailable load is room temperature setpoint adjustment, which reduces energy usage, but also makes people less comfortable (reduction in satisfaction). The EMS makes decision on how much energy to buy and to use to serve the loads ($d_s(t)$). To represent the tradeoff among the controls for these loads, we will use the following utility function,

$$\text{Utility} = \text{E}\,[\text{Satisfaction - Cost}] \tag{1}$$

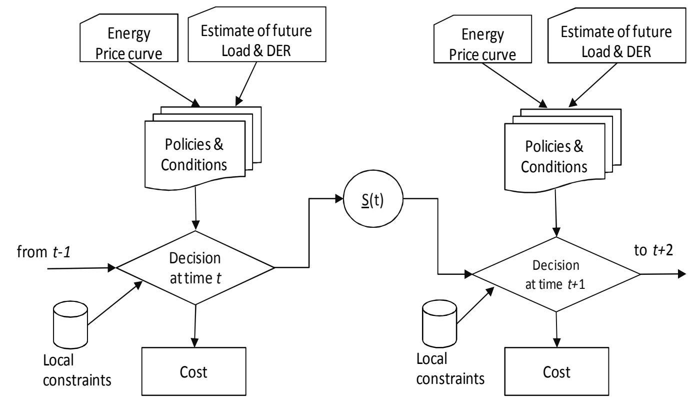

Figure 2. High-level DROP architecture

Satisfaction depends on how much and how soon the requested load is served. If part of the load is shed, satisfaction decreases. Similarly, if the load is deferred to be served, the satisfaction value decreases as a function of the deferred time. Cost is simply the product of price and the energy bought. We use expectation (E[.]) in (1) to take into account the stochastic nature of the random inputs.

In this paper, we are interested in creating an effective algorithm for DR so that the total utility of (1) is maximized. We will describe a flexible and effective heuristic optimization framework, which will allow a first order performance analysis and comparison with respect to systems that do not take advantage of DR or have no optimization.

B. Demand Response Optimization Policy

1) Approach

We propose an algorithm called Demand Response Optimization Policy (DROP), which is based on an hour-ahead heuristic local decision that takes into account global requirements and constraints. DROP is different from other numerical optimization approaches such as dynamic programming as it is expressed in the form of a set of policy and decisions to be executed at each time slot. Computation for DROP is simple and can easily be implemented in real-time. While DROP operates on real data, it takes into account the stochastic nature of the inputs (load and renewable energy source) via prediction of the future input. At a high level, the DROP algorithm incorporates the following attributes:

1. It is a form of "greedy algorithm" in the sense that local decisions are made that influence the solution,
2. Global requirements such as expected load, renewable energy, and price curve, are taken into consideration for creating the local rules,

3. Constraints such as storage limit and serving of essential loads are factored into the local decision,
4. Real-time data are used to update decision at each step,
5. Utility function, based on expected values of the cost, and satisfaction, treats various inputs as stochastic variables, and
6. At any step in time, the algorithm does not go back and change the prior decisions.

TABLE I. DROP Policy

	Condition				Decision	
	1	2	3	4		
Policy	L(t)> R(t)	λ_t < wt. ave. of {λ_{LNS}}	λ_t < min {λ_{SLD}}	λ_t < min {λ_{LS}}	Buy to:	Load serving
1	Yes	Yes	don't care	don't care	fill storage	serve curr. acc. load
2	Yes	No	Yes	don't care	fill storage	serve curr. acc. load
3	Yes	No	No	Yes	fill storage	serve curr. acc. load
4	Yes	No	No	No	serve essent-ial load	serve essential load
5	No	don't care	don't care	don't care	serve essent-ial load	serve essential load

λ_t - Price of grid energy at time t

{λ_{LNS}} - Set of grid energy price during period of non-shiftable load

{λ_{SLD}} - Set of grid energy price during period of shiftable load with deadline (SLD)

{λ_{LS}} - Set of grid energy price during period of shiftable load

L(t) – Total expected load from t+1 onward

R(t) – Total expected renewable energy from t+1 onward

2) *Algorithm Architecture*

The algorithm is based on the idea of making the best decision at each time slot t, using statistical information about the future load, the future renewable energy, and the price curve. The decisions are the amount of grid energy to buy and the quantity of various loads to serve. As shown in the high-level architecture of Fig. 2, the DROP algorithm evaluates a set of policies and executes the decision at each time slot. The policies are made up of a set of conditions that compares the current energy price and the expected cost and demand in the future. Based on the evaluation result of the policies and the current constraints (loads that required to be served immediately and storage constraints), decisions are made for the current time and the system state vector, $\underline{S}(t)$, which includes the accumulated load, current loads, current storage, current DER, and future predictions, is updated. The cost incurred for the current time is then computed, which contributes to the total utility function.

One advantage of this architecture is that it is extensible to future changes with respect to new load types, new heuristic rules for optimization, and new utility functions. For example, if a new load emerges that needs to be treated differently from other loads, new condition(s) can be added to as new entries in the policy table, without changing the overall architecture.

3) *Policies Rules and Decisions*

As shown in Table I, a policy is composed on a set of conditions and decisions. Five policies are shown, but new policies can be added as needed. The first condition suggests that if the total expected load, which include all different types of loads starting from the next hour (t+1) to the end of the day (T), is larger than the total expected resource (renewable energy) from t+1 to T, it encourages buying grid energy, otherwise it discourages buying grid energy. We use "encourage" and "discourage", as other conditions need to be met to make final decision of buy or no-buy is made. The first condition is intuitive as one does not want to buy more than needed, regardless of the price. The second condition compares the current energy price λ_t to the weighted average of the energy prices, {λ_{LNS}}, for the future times when there is non-shiftable load (LN_i) such that,

$$\lambda_t < \sum_{i=t+1}^{T} LN_i \,.\, \lambda_i \;/\; \sum_{i=t+1}^{T} LN_i \tag{2}$$

If condition 2 is satisfied , one is encouraged to buy since on the average, it would cost less to buy energy now, than when it is forced to buy later to meet the non-shiftable load demand. Condition 3 deals with shiftable load with deadline (SLD), which are shiftable loads that may be served any time between current time t and a deadline at t_e. The condition follows similar logic as that of condition 2, but has a different form because the EMS has the flexibility of serving the load any time before the deadline. Therefore condition 3 is given by,

$$\lambda_t < min\ \{\lambda_{SLD}\} \tag{3}$$

The fourth condition is for loads that are shiftable for the rest of the day (time T). It can be viewed as the same as condition 2 with t_e = T. For each policy defined in Table I, there are two types of decisions: 1) buy/no-buy; 2) ways to serve the load.

When the decision is to *buy*, the objective is to fill the storage, unless the excess of future load to future resource is smaller than the storage. The rationale is that if current price is lower than that of the future, one should accumulate as much energy as possible. To fill the storage, we first need to decide if we use existing storage or grid energy to serve the load. We keep track of the *average cost of stored energy,* λ_S , which is computed in every time slot, to guide the decision. If λ_S is lower than the current grid price, stored energy is used, otherwise, grid energy is used. Once the load is served, the buy decision fills the storage with grid energy. When the decision is *no-buy*, the algorithm computes what is necessary to serve the essential load, which includes the non-shiftable load, the last time slot of the shiftable load with deadline, and the last time slot of the shiftable load, once again using λ_S for decision of using storage vs. grid energy for the essential load.

4) DROP Algorithm Details

In a centralized EMS environment, DROP uses the exogenous input at each time slot t, makes decision of what and how much load to serve and how much energy, if any, to buy. The overall goal of the algorithm is to maximize the total utility function for the entire T duration. Fig. 3 shows the flow chart of the DROP algorithm, which is to be implemented in the EMS. At each time slot t, the algorithm first updates the system state vector S(t). Next, it checks the policies of Table I in ascending order. Policies 1-3 are buy policies and 4,5 are no-buy policies. Buy policy includes serving all the accumulated loads. The decision of using storage or grid energy to serve accumulated load depends on whether the average cost (λ_S) of the stored energy is higher (use grid energy) or lower (use stored energy) than that of the grid cost (λ_t). Once the accumulated load is served, DROP buys grid energy to fill up the storage. Non-buy policy buys enough energy from the grid to serve all the essential loads. After executing the decisions, the algorithm then updates the state vector to S(t+1), and computes the current utility.

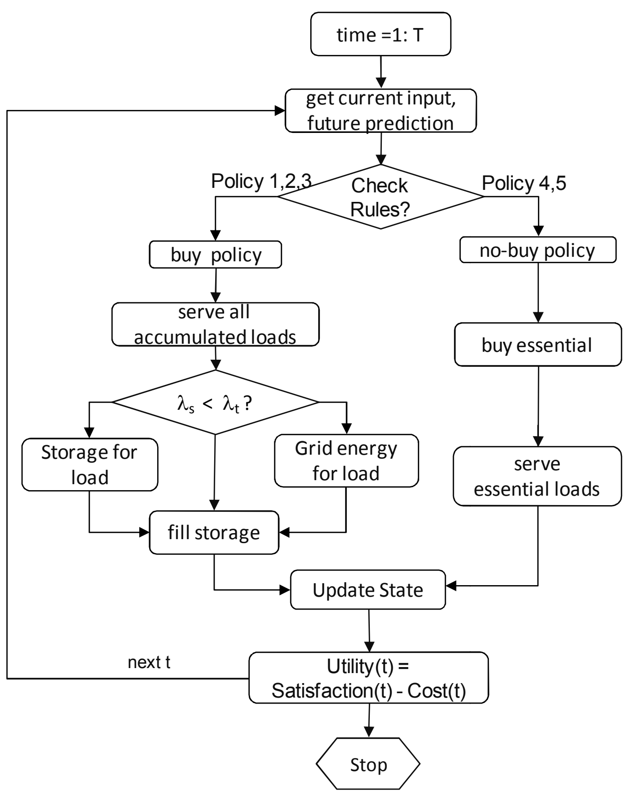

Figure 3. DROP algorithm flow chart

IV. Stochastic Load Modeling

In this section, we present a first-order Markov model for an HVAC chiller data, which will be used in the DROP algorithm for the simulation to be discussed in section V.

The modeling procedure involves the following three main steps, as shown in Fig. 4: 1) Smoothing, 2) Quantization, and 3) Model derivation. The inputs to the modeling suite consist of the real-time measurement data, such as power consumed by a load, and exogenous data that impact the measurement data, such as information about outside temperature, building occupancy, and so on. These inputs are collected in real-time and are a live feed to the modeling suite, which is an online device. However, before the measurement data can be used for modeling or prediction purposes, there is some preprocessing that must be performed on the data. Firstly, we must eliminate any noise in the collected data by applying appropriate smoothing techniques. Secondly, the smoothed data must be quantized into a well-defined set of states. The number of states depends on factors such as the memory capacity of the microprocessor on which the modeling suite runs. Finally, the state-space model must be represented in an accurate and efficient manner.

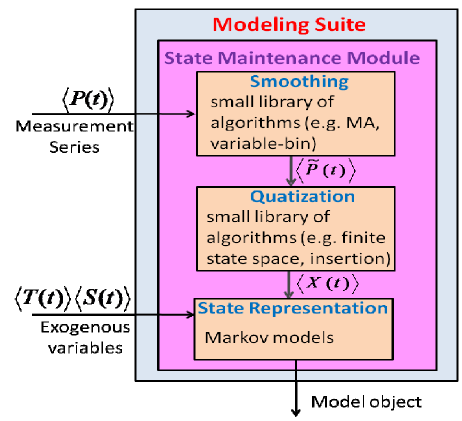

Figure 4. Design of a modeling suite

A. Smoothing

Smoothing is a common procedure for noise removal, typically used in the domain of signal processing. Conventional smoothing techniques used in that domain include Fast Fourier Transform (FFT) smoothing, moving average smoothing, etc. However, such conventional smoothing techniques are not necessarily suitable for our purpose of operational modeling of individual electric devices. The reason for this is that, unlike in signal processing, the objective of modeling in our domain is not to mimic or reproduce the original signal with little distortion, but instead to represent as much *information* as possible about the device's load/generation characteristics. In other words, our purpose for smoothing (and hence quantization) is to ensure that i) the significant or important measurement data-points (e.g. a load's frequent operating level) are preserved as much as possible, and ii) the less important data are eliminated. From this, we can understand that conventional smoothing such as FFT are not suitable because they distort the data (which we would need to preserve). Moreover, they do not have any notion of which data point (i.e., state value) is more important than the other. Consequently, we develop a family of smoothing techniques that are based on the probability distribution of the measurement data. The reasoning for this is that if a device returns to a specific state value very frequently, then it is a significant state in the operation of the device that must be preserved. One such algorithm, called the Variable-bin smoothing, leverages the measurement probability distribution. A fixed number of states (e.g. 8) is entered as input to Variable-bin Smoothing scheme, which then bins the measurement values into these bins according to their probability distribution (where roughly speaking less frequent measurements are packed together, and more frequent measurements are packed in bins by themselves as much as possible). Variable-bin Smoothing (with only 8 states) of an actual HVAC chiller data is shown in Fig. 5.

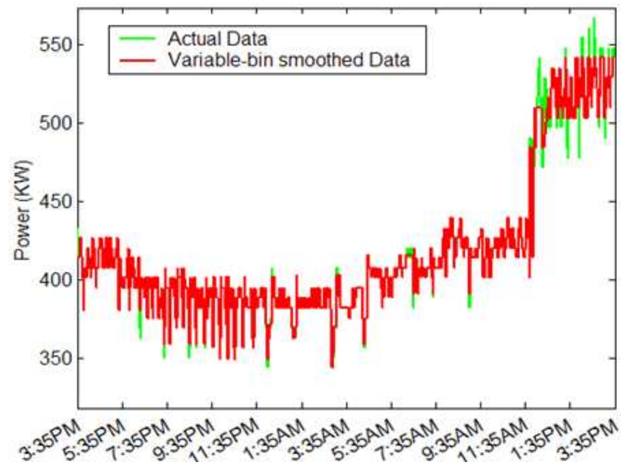

Figure 5. Actual vs. smoothed data with 8 state variable-bin smoothing. Fluctuations represent actual power variation during day, and much greater than rated meter uncertainty (+/- 0.1 %) hence not due to the measurement uncertainty.

B. *Quantization*

Quantization is the process of limiting the state values to a fixed number, depending on the amount of memory available. It must be noted that the probability-based smoothing algorithms that we developed implicitly quantize the data as well. For example, in variable bin smoothing, we use only *n* bins where *n* corresponds to the number of states, and each bin value (computed as the weighted average of the values contained in the bin) denotes a state.

C. *Model Deriviation*

The smoothed (and quantized) data is used to derive the state transition probabilities, and thereby the Markov model. For example, Fig. 6 shows the transition probability derived from the smoothed data.

Thus the state set and the corresponding transition probabilities constitute the simple first order Markov model for the data. Furthermore, we maintain a windowed-Markov model, where independent transition probabilities are derived for each t-hour window, (example, t=3). This helps improve the accuracy of the model. The length of the window depends on the memory resources available in the system, since this requires additional matrices to be stored.

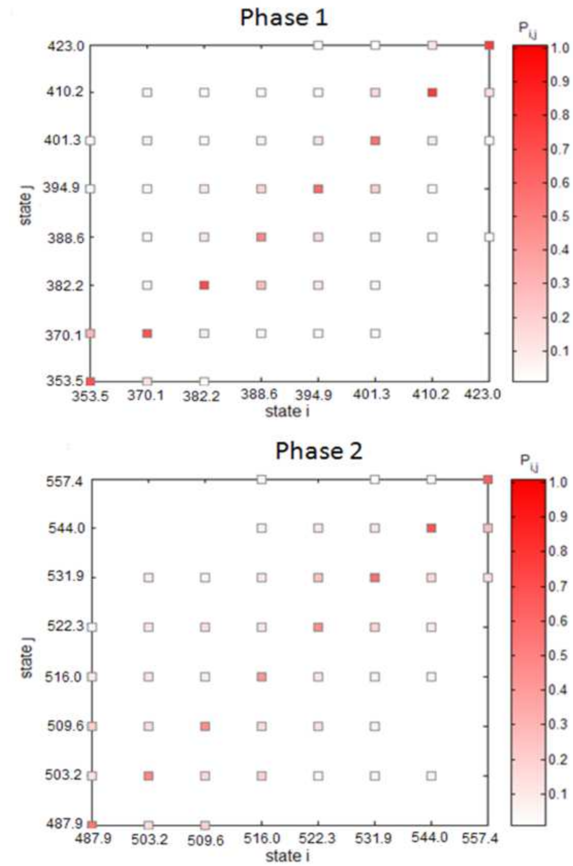

Figure 6. 8-state transition probability for smoothed chiller data

V. SIMULATION RESULTS

The potential cost benefit of using DROP is analyzed using simulations of a typical commercial building on a September day in New Jersey. DROP re-optimizes hourly to account for the latest observations. From DROP's viewpoint, energy price, both shiftable and non-shiftable loads and renewable generation are all exogenous variables requiring forecast (hourly). The hourly energy price is modeled as a deterministic time sequence approximating the day-ahead bids on Sept. 1, 2008 published by PJM [8] with an average of 77.8 \$/MWhr. Hourly total load statistics is assumed to be the measured chiller loads at the Telcordia Piscataway campus on a Sept. 1. Chiller power demand is nearly 0.5 MW and corresponds to about 40 % of total building energy demand. The chiller load is modeled using the Markov model described in Sec. IV. For comparison, we also include a "perfect forecast" using real chiller data. While Telcordia chiller load is non-shiftable, shiftable load is simulated as a fraction of total chiller load. Renewable generation is modeled as hourly-Gaussian fitting approximately the solar data shown in [9]; but scaled up to equal 10% of the total load, in daily aggregate.

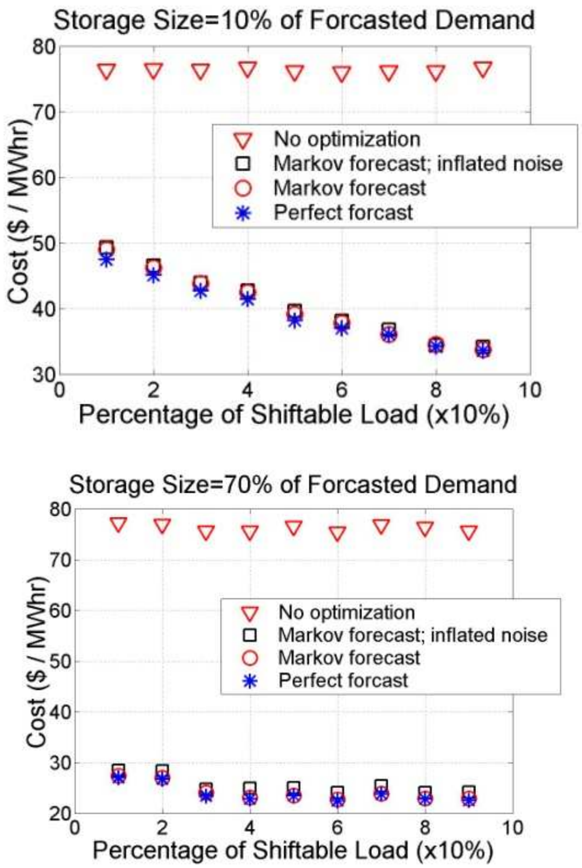

Figure 7. Cost vs. Shiftable load % at two energy storage sizes

We measured the cost per unit energy bought ($/MWhr) for scenarios with different allocations between shiftable and non-shiftable loads and different storage sizes (we did not include storage cost in this initial study). Expected total loads in each scenario are identical, while the actual values vary mildly. Fig. 7 shows costs corresponding to using DROP for optimization for 3 exogenous models (*Perfect*, in which exact load is known, *Markov*, and *Markov with inflated noise*). These are compared to the no- optimization case in which loads are served in the hour as they appear, using first renewable, and then if exhausted, grid energy. The cost comparison is shown with respect to storage sizes of 10 % and 70 % of the total load. We see from Fig. 7 that DROP yields significant cost savings ranging from about 35 % to 70 %, which increases with both percentage of load being shiftable and the storage size. Our simulation also shows that saving rates of shiftability and storage, when used *alone*, are approximately 0.16 $/MWhr/percent and 0.33 $/MWhr/ percent, respectively; which means shiftability is about 48 % as effective as energy storage in terms of enabling cost saving. These allow the calculations of breakeven points of investments in both direct energy storages, e.g., batteries, and shiftability. Given the non-trivial cost-saving effectiveness of shiftability, it deserves similar considerations as energy storage as part of an integrated strategy for energy efficiency.

From Fig. 7, we notice that the gain from using the Markov model is slightly smaller than that of the perfect forecast. When inflated noise (standard deviation equals twice the mean) is added to the Markov model, it shows less gain, although the difference is small. One explanation is that DROP leverages forecasts of highly aggregated quantities, such as the predicted energy demand aggregated over all consumers for the rest of day, rather than depending on details of hourly forecasts. This tends to reduce sensitivity to uncertainty, and hence increases reliability, as shown. Further research will focus on how DROP may better use detailed forecast to improve the savings.

VI. CONCLUSION AND FUTURE WORK

This paper explores the potential cost savings of a dynamical DR system using a simple but effective algorithm. To obtain a more realistic result, we have used real chiller data as the non-shiftable load. We also presented a first order Markov model for such data for the simulation. Our simulation results show that the DROP algorithm has potential savings up to 70% compared to a baseline case of no-optimization. The results also quantify the cost-saving rates of both energy storage and the novel strategy of shiftability. In particular, our results suggest that shiftability, with savings similar to that of storage, deserves serious considerations as part of an integrated strategy for energy efficiency.

While opportunities for incremental improvements abound, more interesting research extension would be decentralized DR optimization and models appropriate for decentralization. We would like to understand better the tradeoff between centralized and distributed approaches, and the impact of the degree of decentralization. Would accurate modeling play a larger role in the decentralized DR strategy? Should devices be myopically reactive, and a centralized manager be the smart broker between energy consumers and producers? Or should devices interact directly with the grid and renewable generators?

REFERENCES

[1] US Department of Energy, Buildings Energy Data Book, Department of Energy, March 2009. http://buildingsdatabook.eren.doe.gov/.

[2] ZBB Energy Corporation, Advanced Electricity Storage Technologies program, December 2010. Available at: http://www.zbbenergy.com/index.php/download_file/view/152/228/.

[3] K. Kok, M. Scheepers, and R. Kamphuis, "Intelligence in electricity networks for embedding renewables and distributed generation," in R.R. Negenborn, et. al., *Intelligent Infrastructures*. Springer, 2009

[4] http://gridwise.pnl.gov/docs/op_project_final_report_pnnl17167.pdf

[5] H. Sane and M. Guay, "Minmax Dynamic Optimization over a Finite-time Horizon for Building," 2008 ACC, June 2008.

[6] Xie, Le, and Marija D. Ilic, "Model Predictive Economic/Environmental Dispatch of Power Systems with Intermittent Resources," IEEE PES General Meeting 2009, July 2009.

[7] Xie, Le, Jhi-Young Joo, and Marija D. Ilic, "Integration of Intermittent Resources with Price-Responsive Loads," North American Power Symposium (NAPS 2009), October 2009.

[8] PJM wholesale electricity market. http://www.pjm.com/markets-and-operations/energy/real-time/historical-bid-data.aspx.

[9] A. KansalL, et. al., "Performance aware tasking for environmentally powered sensor networks," In *ACM SIGMETRICS*. 2004.