Abstract—We consider a system where inelastic demand for electric power is met from three sources: the grid, in-house renewables such as wind turbines or solar panels, and an in-house energy storage device. In our setting, power demand, renewable power supply, and cost for grid power are all time-varying and stochastic. Further, there are limits and efficiency issues for charging and discharging the energy storage device. Under such a scenario, at all times across an infinite horizon, we need to determine how to split power demand among the various sources. For that, we formulate a dynamic programming framework that minimizes the long-run average operational cost subject to satisfying demand and meeting capacity constraints. To determine the optimal actions, we propose a dynamic program with deterministic and stochastic components where parameters estimated from a training data set (that uses real demand, supply and price) is used. The program is computationally intensive for large-scale problems, we explore algorithms based on approximate dynamic programming, and apply them to a test data set. We compare its performance against other heuristics that may or may not use the training data. In addition, we ascertain the value of storage as well as the value of installing a renewable source.

I. INTRODUCTION

Renewable generation capacity is expanding rapidly to potentially reduce carbon dioxide emissions and dependence on fossil fuels. As non-dispatchable generation, renewable energy introduces variability into the energy portfolio, and further amplifies the difficulty of matching demand with supply in real time. Energy storage devices are environmentally friendly candidates that can provide flexibility to the system and mitigate the impact of volatile renewable generations.

The focus of this paper is on the operation of electric storages operated by the electricity consumers who own distributed renewable generation and face time-varying (and stochastic) electricity prices. Our motivation stems from the potential of electricity consumers to own and use storage devices (e.g., major consumers like data centers [1], [2] and individual consumers who own PHEVs [3], [4], and from a recent study that shows consumer ownership of storage can be socially beneficial [5]. We also note that there is a growing trend for residential consumers and data centers to own distributed renewable generation [6], [7], [8].

In this paper, we consider a consumer of electricity with inelastic power demand, i.e. demand must be satisfied instantaneously and cannot be either postponed or cut back. Demand for power is time-varying and stochastic. Part of the demand can be met by a renewable energy source (such as photovoltaic (PV) solar panels or wind turbines) that is situated locally and owned by the consumer. Note that renewable power supply is time-varying and stochastic. Remaining demand (if any), beyond what the renewable source can supply, is satisfied either from the grid or by an in-house Energy Storage Device (ESD) or both. It is also possible to charge the ESD, thus the ESD is both a source and a sink of power. Like power demand and renewable supply, price for power from the grid is also time-varying and stochastic.

In the last few years this problem has received a lot of attention in the literature. The literature can be broadly divided into two categories: (i) articles that assume that all parameters are deterministic time-varying quantities and make one-shot decisions such as how much to invest in renewable sources and ESDs; what fraction of demand is met by renewable supply; whether it is possible to go net zero, etc., and (ii) articles that assume that demand, supply and/or costs are either IID or follow stationary Markov chains; these result in Markov decision processes (MDPs) or Lyapunov-based techniques to model the storage process.

For case (i), the authors of [9] consider the storage operation problem faced by a data center. Besides the modeling nuances we consider, [9] also models the use of diesel generators and external renewable sources (besides on-site renewables). They evaluate strategic decisions such as whether to use renewables, from where, whether to invest in storage, whether to use diesel generators, etc. The (marginal) value of energy storage devices is assessed in [10], [11], for two different settings with and without renewable generation. These two works use the assessment on value of storage to determine the sizing of storage devices as well as the influence of large-scale electricity storage on electricity prices.

For case (ii), there exists a substantial literature on the operation of energy storage owned by system operators [12], [13]. Another well studied application of energy storage is the use of storages to arbitrage [14], [3]. A few recent works conduct a dynamic programming approach to derive the arbitrage value of electric storage, in the presence of dynamic pricing [15], [16]. The aforementioned works assume that owners (of electric storages) put no value on electricity consumption and do not have their own generating sources.
There have been recent studies on the operation of consumer-owned ESDs. Leveraging on techniques from queuing theory, a few recent papers propose a variety of on-line algorithms that are shown to be asymptotically optimal, as the storage capacity increases to infinity [17], [18]. Closer to the present paper, the authors of [18] consider a similar storage operation problem faced by a consumer who has additional options to sell power back to the grid (but through the ESD) and to shift demand across time. Unlike MDP, which is computationally complex and requires substantial statistical information of the system dynamics, their method requires no statistical knowledge and uses an extremely light linear program. The algorithm is expected to perform well when the storage capacity is significantly larger than the maximum charging/discharging rates, i.e., when it takes many hours to fully charge and discharge the storage.

While our article considers a consumer, [19] takes the perspective of a renewable source (such as a wind farm) with large storage device. They consider renewable supply to be controlled and sold/stored, and also negative prices. Their objective is to develop a fast algorithm to determine how much to generate and what fraction of that must be sold versus stored so that revenue is maximized. At each discrete time instant, they develop three thresholds (by solving a large optimization problem) on the ESD contents to determine when to: generate, buy and store; generate and store; do nothing; or sell.

The authors of [20] study the joint optimization of wind generation and energy storage through a stochastic programming approach. The computation of such a stochastic program becomes intractable as the number of stages increases, because the number of scenarios grows exponentially with the number of stages.1 Closely related to the present paper, a few recent papers establish structural properties on optimal storage operation policies in a variety of MDP settings [21], [22], [23]. We note, however, that due to the well known curse of dimensionality, exactly solving the MDP is usually intractable for practical settings with stochastic renewable generation and electricity prices.

The contribution of this paper is twofold: (i) we construct a dynamic programming framework that incorporates the randomness in consumer demand, renewable generation, and electricity prices; (ii) using a set of real training data on electricity prices, solar generation, and consumer demand, we numerically compare approaches ranging from Markovian models, to hybrid methods based on statistics, optimization and heuristics, to those that require no historical data. The insight we obtain from these numerical experiments will guide the design of more efficient heuristics. Interestingly, most of the aforementioned articles in the literature consider wind energy while our focus for the numerical study is on solar PV. We summarize some of our key findings in the following.

1) Approximate dynamic programming (ADP) based algorithms usually yield the best performance. Algorithms based on Lyapunov techniques (e.g., the one proposed in [18]) requires no statistical knowledge and minimum computation power, and performs reasonably well when the storage capacity is significantly larger than the maximum charging/discharging rates. For fast-charging storage devices that can be fully charged within 2 hours, on the other hand, ADP-based algorithms significantly outperforms the one proposed in [18].

2) The value of storage (VoS, which is measured as the net benefit obtained by the consumer if she operates the storage according to an ADP algorithm) is much higher under 5-minute real-time pricing than that under hourly pricing, due the higher variability in cost in the former case. VoS increases sharply with the storage capacity only when the maximum charging/discharging rates grow in proportion to the storage capacity. In other words, the value of storage does not increase appreciably with increase in storage size, if the maximum charging/discharging rates remain fixed.

We finally note that the storage operation problem is intimately related to inventory control problems with random production cost and uncertain demand [24], [25]. We note, however, our model is significantly different from the setting in the inventory control literature. In our model, there is no “inventory” holding cost that is proportional to the storage level; instead, the major losses resulting from storage operation are due to energy injection and withdrawal (e.g., battery charging and discharging), and there are constraints on the charging and discharging rate of a certain storage device.

The rest of the paper is organized as follows. We describe our problem in Section II. In Section III we develop a probabilistic model and suggest ADP-based approaches to solve it. We consider other heuristic algorithms in Section IV. We compare and discuss the performance of these algorithms in Section V by obtaining parameters using a real training data set and testing them.

II. PROBLEM DESCRIPTION

Before describing our model, we state a few assumptions. Recall that demand is inelastic. There are no operational costs for the consumer other than price of electricity from grid. Price is exogenous and not affected by the consumer’s decisions. There are inefficiencies in charging and discharging the ESD. We assume no leakages in the ESD, since the storage efficiency of many different types of modern batteries (e.g., Lead acid, Li-ion and Vanadium redox batteries) is close to 100% [26]. The ESD has a finite energy storage capacity. Power cannot be sold to the grid (we will relax this assumption in an extension of this work). There are no diesel generators or external renewable sources.

Next we describe some of the notations used in this work (pictorially described in Fig. 1). We consider a discrete-time model where time periods are indexed by $t = 0, 1, \ldots$. The stochastic uncontrollable variables are: $D_t$, the demand for energy in period $t$ (in kWh); $S_t$, the energy supply from renewable source in period $t$ (in kWh); $C_t$, the cost in period $t$ for a unit of energy from grid (in $ per kWh). Let $U_t$ be the amount of energy in the ESD at the beginning of period $t$ (in kWh).
There are constraints and inefficiencies in the ESD charging and discharging processes. A maximum of $K$ kWh of energy can be stored in the ESD at any time. The ESD can be discharged and charged at a maximum rate of $c_{\text{dis}}$ and $c_{\text{char}}$ (in kW) respectively. Also, the ESD discharging and charging efficiencies are $\eta_{\text{dis}} \leq 1$ and $\eta_{\text{char}} \leq 1$ respectively (which we explain next). If $\rho$ kW of power is used to charge the ESD in period $t$, then $U_t = -1$ ($\forall t$ and $\rho \leq c_{\text{char}}$) is $\rho \eta_{\text{char}}$ hours per unit time. Likewise if $\rho$ kW of power is needed from the ESD, then $U_t = 1$ ($\forall t$ and $\rho \leq c_{\text{dis}}$) is $-\rho / \eta_{\text{dis}}$ hours per unit time. Next we describe the decision variables under the control of the consumer. Let $X_t$, $Y_t$ and $Z_t$ be the energy drawn (in kWh) from the grid, renewable source and ESD respectively at period $t$. While $X_t \geq 0$ and $Y_t \geq 0$ for all $t$, $Z_t$ can be positive or negative (with negative value denoting power used to charge ESD).

Having described the notation, we are in a position to state the problem we consider. During every period $t$, given demand $D_t$, renewable supply $S_t$, cost $C_t$ and ESD charge level $U_t$, we need to determine the supply from grid $X_t$, the draw from renewable source $Y_t$, and contribution from the ESD $Z_t$ so that the long-run average cost per time period is minimized subject to satisfying demand, staying within ESD capacities and other constraints such as dynamics and non-negativity. This sequential decision making problem can be formulated mathematically as follows:

\[
\text{Minimize } \lim_{\tau \to \infty} \frac{1}{\tau} \sum_{t=0}^{\tau} \mathbb{E} \{ C_t X_t \}
\]

Subject to the following $\forall t \in \{0, 1, 2, \ldots\}$

\[
\begin{align*}
X_t + Y_t + Z_t &\geq D_t \\
0 &\leq Y_t \leq S_t \\
Z_t &\leq c_{\text{dis}} \\
-\min\{Z_t, 0\} &\leq c_{\text{char}} \\
\psi(U_{t+1} - U_t) &\geq -\max\{Z_t, 0\} / \eta_{\text{dis}} + \eta_{\text{char}} \min\{Z_t, 0\} \\
0 &\leq U_t \leq K \\
X_t &\geq 0
\end{align*}
\]

where $\psi$ is a constant for time-unit conversion, i.e. number of time units per hour (viz. since $U_t$ is in kWh and $Z_t$ in kW, if length of period $t$ is 1 second then $\psi = 3600$). Note that $C_t$, $D_t$, and $S_t$ are random variables.

Remark 1: The above formulation has a trivial solution if $C_t$ stays constant over $t$. In particular, if $D_t \geq S_t$ for all $t$ then $Z_t = 0$, $Y_t = S_t$ and $X_t = D_t - S_t$ is the optimal solution with no need for storage. However, if $C_t$ stays constant over $t$ but $D_t$ is not always greater than $S_t$, then the greedy solution of charging or discharging at maximum feasible capacities when $D_t$ is less or more respectively than $S_t$ is optimal. Thus for the rest of this article we consider the non-trivial case of $C_t$ varying with time and the possibility that for some time periods $S_t$ can be greater than $D_t$.

Remark 2: We can reduce the above problem to a 1-dimensional control in $X_t$ or $Z_t$ by realizing that we can let $Y_t = S_t$ and $Z_t = D_t - X_t - Y_t$. However, for ease of presentation we will use all variables, not just $X_t$ or $Z_t$.

Remark 3: Due to the linear nature of the objective function and constraints, the value function at every period $t$ is piecewise-linear in $U_t$ [23]. As a result, given the system state at period $t$, namely $D_t$, $S_t$, $C_t$, and $U_t$, the optimal policy is such that $Z_t$ is one of the following: (i) zero, (ii) the minimum of $c_{\text{dis}}$, $\psi U_t$ and $(D_t - S_t, 0)^+$, or (iii) negative of certain threshold that depends on future demand and cost.

III. MDP: Probabilistic Model with Cycles

In this section, we first introduce the way we fit real data into an MDP model, and then discuss approaches to solve it. Analyzing the data described in [9], [6] and the NREL labs, it is evident that demand, solar PV supply and cost are time-varying and stochastic. However, it is also not unreasonable to assume that there are daily or weekly seasonality effects. In other words, there is a deterministic variability as well as stochastic variability. To model such a phenomenon we consider what we call probabilistic model with cycles.

Definition 1: An uncontrolled process $\{V_t, t \geq 0\}$ is cyclic with cycle length $N$ if $V_t$ is stochastically identical to $V_{t+N}$ for all $t \in \{1, 2, \ldots\}$, where $N$ is the number of periods in a cycle of length $T$ hours, i.e. $N = \psi T$.

Based on the above definition, we assume that $\{D_t, t \geq 0\}$, $\{S_t, t \geq 0\}$ and $\{C_t, t \geq 0\}$ are cyclic with cycle length $T$ (one cycle typically is the equivalent of one day). Further, we write down for all $t \in \{0, 1, \ldots\}$, with $n = (t \mod N)$,

\[
\begin{align*}
D_t &= d_n Z_t^d + \delta_n \\
S_t &= s_n Z_t^s \\
C_t &= c_n Z_t^c + \gamma_n,
\end{align*}
\]

where \{d_0, \ldots, d_N\}, \{\delta_1, \delta_2, \ldots, \delta_N\}, \{s_1, s_2, \ldots, s_N\}, \{c_1, c_2, \ldots, c_N\}, and \{\gamma_1, \gamma_2, \ldots, \gamma_N\}, are sets of deterministic constants while \{Z_t^d, t \geq 0\}, \{Z_t^s, t \geq 0\}, and \{Z_t^c, t \geq 0\}, are stationary and independent discrete time Markov chains on discrete state spaces $S^d$, $S^s$, and $S^c$ and transition probability matrices $P^d$, $P^s$, and $P^c$ respectively.

One can think of \{s_1, s_2, \ldots, s_N\} as the power supplied by PV panels on a perfectly sunny day while $s_n$ as a discrete set of values between 0 and 1. The demand and cost terms do not have such a nice interpretation and one would have to model them carefully based on data. In Section V, we will use training data to estimate $d_n, \delta_n, s_n, c_n$ and $\gamma_n$ for all $n \in \{1, \ldots, N\}$ as well as $P^d, P^s$, and $P^c$. However, for the rest of this section we take a probabilistic approach assuming all the aforementioned parameters are known and formulate the system as an MDP.
We denote the system state at time $t$ as a 5-tuple
\[ x_t = \{ (t/N) + 1, Z^d_t, Z^e_t, Z^c_t, U_t \}, \]
where $\{ t/N \}$ denotes $(t \mod N)$ with state space $S$ given by the cartesian product
\[ \{1, 2, \ldots, N\} \times S^d \times S^e \times S^c \times S^u \]
where $S^u$ is the discrete set of values between 0 and $K$ that $U_t$ can take. It is worth noting that the time dependency and correlations of consumer demand and renewable generation are incorporated by including in the system state a periodic Markov chain that describes time evolution.

Let the action at time $t$ denote the amount of power to be supplied from the grid, i.e. $X_t$, with action space $A$ corresponding to the set of all possible non-negative values that are less than or equal to
\[ \psi \min \left\{ K - \frac{U_{(t/N)}}{\eta_{\text{char}}}, c_{\text{char}} \right\} + \max \{ D_{(t/N)} - S_{(t/N)}, 0 \}, \]
for all $t \in \{1, 2, \ldots, N\}$. As noted in Remark 2, the energy drawn from renewable source and ESD at period $t$ is determined by $X_t$, i.e., $Y_t = S_t$ and $Z_t = D_t - X_t - Y_t$.

For every action $X_t \in A$, $x \in S$, and $y \in S$, we can compute the transition probability $P_{xy}(X_t)$ using appropriate Kronecker products of $P^d$, $P^e$, $P^c$ and other matrices of zeros and ones (which are not explained due to space constraints). The stage cost at time $t$ is the product of the corresponding power cost in state $i$ times $X_t \in A, C_i X_t$.

By incorporating the time element into the state of the dynamic program, we have indeed formulated a “stationary” MDP (the quotes are because the state transition is stationary from one cycle of $N$ values to the next cycle, but not within a cycle). For a given stationary policy which maps every point in the state space $S$ to a point in the action space $A$, the long-run average cost per unit time corresponds to the objective function of our optimization problem defined in Section II (and also results in a feasible solution).

Next we outline methods to solve the MDP.

A. Exact MDP Solution

Note that the above MDP has a finite state-space and a finite action-space. There are many methods to obtain the optimal action $a \in A$ in state $i$ for all $i \in S$. We consider a linear program (LP) based method. We use (a different) LP:

LP: Minimize
\[
\sum_i \sum_a c_{ia} x_{ia}
\]
subject to
\[
\sum_i x_{ia} = 1
\]
\[
\sum_a x_{ia} = \sum_i \sum_a p_{ij}(a) x_{ia} = 0, \quad \forall j
\]
\[
x_{ia} \geq 0, \quad \forall i \in S, \quad \forall a \in A.
\]

Let $a^*$ be the optimal solution to the above LP. Then the optimal randomized policy $u^*_{ia}$ can be computed as $u^*_{ia} = x^*_{ia} / \sum_b x^*_{ib}$. As described in [27], the LP produces for each $i$ optimal values $x^*_{ia}$ that are all zero except one $a$ which would be 1 (the optimal action). In this manner (in theory) it is possible to determine the optimal action in each state. However in practice, one could encounter difficulties.

The above LP can be solved (especially by packages such as MATLAB) when the state and action spaces are small, $\eta_{\text{char}} = \eta_{\text{dis}} = 1$, and $D_n$ and $S_n$ belong to a small discrete set for all $n \in \{1, 2, \ldots, N\}$. MATLAB runs out of memory when the dimension becomes large and approximating by rounding off to integer values sometimes results in solutions that are too far from being optimal. In the next subsection, we adopt a common procedure usually referred to as Approximate Dynamic Programming (ADP) (see [28]) to deal with the curse of dimensionality.

B. Approximate MDP Solution

We start from the simplest ADP algorithm that looks only a single stage ahead. In this One-Step Look-ahead algorithm (OLA), given the current state, we determine the best action so that the expected cost for this state and the next state is minimized. Formally, given the current system state $x_t$, the algorithm chooses an action $X_t \in A_t$ that minimizes the following cost:

\[ C_t X_t + \sum_y P_{xy}(X_t) \cdot C_{t+1} \cdot (D_{t+1} - S_{t+1} - \psi \eta_{\text{dis}} U_{t+1})^+, \]

where $y$ is the system state at time $t + 1$, which includes the parameters $C_{t+1}, D_{t+1}, S_{t+1}$ and $U_{t+1}$. In this myopic version of ADP, the stage $t + 1$ is treated as the terminal stage and therefore the stage cost at time $t + 1$ is given by $C_{t+1} (D_{t+1} - S_{t+1} - \psi \eta_{\text{dis}} U_{t+1})^+$.

A natural improvement of the OLA is to replace the myopic stage cost at state $y$ by some approximated cost-to-go at this state. Formally, given the current system state $x_t$, a One-step Roll-out algorithm (ORA) chooses an action $X_t \in A_t$ that minimizes the following cost:

\[ C_t X_t + \sum_y P_{xy}(X_t) \cdot \tilde{J}_{t+1}(y), \]

where $\tilde{J}_{t+1}(y)$ is an approximation of the cost-to-go at system state $y$ and stage $t + 1$. We note that OLA simply treats stage $t + 1$ as the terminal stage and let $\tilde{J}_{t+1}(y)$ be the stage cost at time $t + 1$ in state $y$. The ORA algorithm, on the other hand, takes into account future system dynamics and approximates the cost-to-go at system state $y$ by solving a deterministic optimization problem where all random variables take the expected value, conditioned that the system state at time $t + 1$ is $y$.

\[ \tilde{J}_{t+1}(y) = \text{Minimize} \sum_{\tau=t+1}^{t+N} \mathbb{E}[C_{\tau} \mid y] \cdot \bar{X}_{\tau} \]

Subject to the following $\forall \tau \in \{ t+1, t+2, \ldots, t+N \}$
\[ \bar{X}_{\tau} + \bar{Y}_{\tau} + Z_{\tau} \geq \mathbb{E}[D_{\tau} \mid y], \]
\[ 0 \leq \bar{Y}_{\tau} \leq \mathbb{E}[S_{\tau} \mid y], \]
\[ \bar{Z}_{\tau} \leq \bar{C}_{\text{dis}}, \]
\[ -\min(\bar{Z}_{\tau}, 0) \leq \bar{C}_{\text{char}} \]
\[ \psi(\bar{U}_{\tau+1} - \bar{U}_{\tau}) = -\max(\bar{Z}_{\tau}, 0)/\eta_{\text{dis}} + \eta_{\text{char}} \min(\bar{Z}_{\tau}, 0) \]
\[ 0 \leq \bar{U}_{\tau} \leq K, \quad \bar{X}_{\tau} \geq 0, \]
where $\mathbb{E}[\cdot | y]$ denotes conditional expectation, and the minimization is taken over the variables $\bar{X}_t$, $\bar{Y}_t$, and $\bar{Z}_t$.

### IV. Other Heuristics

Before describing another two heuristics that we develop to approximately solve the MDP, we revisit Huang, Walrand and Ramchandran [18] (details explained in Section I). We call their algorithm **HWR**. It is a remarkable online algorithm that does not use any historical information and makes decisions based only on current state information (such as $D_t$, $S_t$, $C_t$ and $U_t$). As the authors of HWR argue, it is an ideal algorithm under such situations when there is tremendous variability, and there may be too many parameters to fit while modeling as an MDP. However, it is our belief that when historical information is available, it appears like a waste to not use all that especially when there are cyclic patterns.

With that understanding, we are in a position to describe two heuristics that use the cyclic structure in Section III. Under such a setting, we first obtain a fluid model of the system by taking expectation of $D$, $S_t$, and $C_t$ for all $\tau \in \{1, 2, \ldots, N\}$. Then we have that the relative straightforward to estimate $\mathbb{E}[D_t]$, $\mathbb{E}[S_t]$ and $\mathbb{E}[C_t]$ as the sample mean from historical observations. Then we can obtain the resulting policy as the fluid model (FL) with the variables $\bar{X}_t$, $\bar{Y}_t$, $\bar{Z}_t$, and $\bar{U}_t$,

**FL:** Minimize $\sum_{\tau=1}^{N} \mathbb{E}[C_t]\bar{X}_t$

Subject to the following $\forall \tau \in \{1, \ldots, N\}$

$$\begin{align*}
\bar{X}_t + \bar{Y}_t + \bar{Z}_t &\geq \mathbb{E}[D_t] \\
0 &\leq \bar{Y}_t \leq \mathbb{E}[S_t] \\
\bar{Z}_t &\leq \epsilon_{\text{dis}} \\
-\min\{\bar{Z}_t, 0\} &\leq \epsilon_{\text{char}} \\
\psi(\bar{U}_{t+1} - \bar{U}_t) &= -\max\{\bar{Z}_t, 0\}/\psi_{\text{dis}} + \eta_{\text{char}} \min\{\bar{Z}_t, 0\} \\
0 &\leq \bar{U}_t \leq K \\
\bar{X}_t &\geq 0,
\end{align*}$$

where $\bar{U}_{t+1} = \bar{U}_t$.

The two heuristics that we are about to explain essentially use the above fluid model’s solution. The idea for the heuristics is based on Remark 3. The state at time $t$ is the 5-tuple $\left(\{t/N\}, D_t, S_t, C_t, U_t\right)$. However, it is not easy to determine in each state which case to choose among (i), (ii) or (iii) of Remark 3. Thus as an approximation we consider guidelines provided by the fluid model and propose two heuristics: TBA (threshold-based approximation) and NOA (naive opportunistic algorithm).

**Heuristic TBA:** Given the state at time $t$ namely, $\left(\{t/N\}, D_t, S_t, C_t, U_t\right)$, we determine $Z_t$ so that at time $t+1$, $\bar{U}_{t+1}$ is as close to $\bar{U}_{\{t/N\}+1}$ by appropriately charging or discharging. The goal is to reach threshold level $\bar{U}_{\{t/N\}+1}$ in the next time. Thus TBA is as follows (with $n = \{t/N\}$):

if $U_t < \bar{U}_{n+1}$, $Z_t = -\min\left\{\psi(\bar{U}_{n+1} - U_t)/\psi_{\text{char}}, \epsilon_{\text{char}}\right\}$

else if $U_t = \bar{U}_{n+1}$, $Z_t = 0$,

else $Z_t = \min\left\{\psi(U_t - \bar{U}_{n+1})/\psi_{\text{dis}}, \epsilon_{\text{dis}}\right\}$.

In all the above cases, $X_t = \max\{D_t - S_t - Z_t, 0\}$.

**Heuristic NOA:** Given the state at time $t$ namely, $\left(\{t/N\}, D_t, S_t, C_t, U_t\right)$, we adopt a naive (but intuitive) strategy – if $C_t$ is cheap, charge the ESD as much as possible; and if $D_t$ is much higher than $S_t$, discharge as much as possible; otherwise do what the fluid model suggests. For that we use $\mathbb{E}[C_t]$ as the grand average cost (computed over entire cycle $N$), and $\text{Var}[C_t]$ the corresponding grand variance; also $\phi_{\text{c}}$ and $\phi$ are parameters to be tuned. It leads to the following NOA (with $n = \{t/N\}$):

if $C_t < \mathbb{E}[C_t] - \phi_{\text{c}}\sqrt{\text{Var}[C_t]}$, then $Z_t = -\min\{\psi(K - U_t)/\eta_{\text{char}}, \epsilon_{\text{char}}\}$ otherwise,

if $D_t - S_t > \mathbb{E}[D_t] - \mathbb{E}[S_t] + \phi\sqrt{\text{Var}[D_t] + \text{Var}[S_t]}$, $Z_t = \min\{D_t - S_t - \bar{X}_n, \psi(U_t)/\epsilon_{\text{dis}}, \psi_{\text{char}}\}$

else (i.e. $D_t - S_t < \mathbb{E}[D_t] - \mathbb{E}[S_t] + \phi\sqrt{\text{Var}[D_t] + \text{Var}[S_t]}$)

$Z_t = \min\{\max(S_t - D_t, \bar{Z}_n)/\epsilon_{\text{char}}, \psi(K - U_t)/\epsilon_{\text{char}}\}$

In all the above cases, $X_t = \max\{D_t - S_t - Z_t, 0\}$.

### V. Numerical Experimentation and Results

We obtained 26 days of demand, supply and cost data in a single month. Our main purpose was to get a representative sample that adequately captures the deterministic and stochastic variability over time. In that spirit we collected demand data from households (http://www.doc.ic.ac.uk/~dk3810/data/), solar PV supply data from NREL (http://www.nrel.gov/mbic/), and cost data at 2 granularities: 1-hour pricing (https://www.nationalgridus.com) and 5-minute pricing (http://iso-ne.org/).

We used 16 days of collected data to train the model, i.e., estimate/fit parameters in the MDP model described in Section III. For 1-hour granularity we have $\psi = 1$ and use Monte-Carlo simulations for 100 days. For 5-minute granularity we have $\psi = 12$ and use raw data in the 10 remaining days (from the original 26 days) for testing.

To estimate $\delta_t$ and $\delta_t$ for any $t \in [1, N]$, we use $D(1, t), D(2, t), \ldots, D(16, t)$, the realized demands in 16 days, to compute $\delta_t = \min_{i}[D(i, t)]$ and $\delta_t = \max_{i}[D(i, t)] - \delta_t$. Likewise for $\epsilon_t$ and $\gamma_t$. In case of supply $s_t$, the minimum value is zero. Then for the DTMCs $\{Z_t^n, t \geq 0\}$, $\{Z_t^m, t \geq 0\}$, and $\{Z_t^{\text{dis}}, t \geq 0\}$, we first arbitrarily select the number of states $M$. The state space is a set of discrete values $0, 1/(M - 1), 2/(M - 1), \ldots, 1$. Then we estimate the elements of $P^d$, $P^s$, and $P^{\text{dis}}$ as the respective frequency of transition based on the 16 days’ data. For all the numerical experiments we use $\epsilon_{\text{char}} = \epsilon_{\text{dis}}$ and $\eta = \eta_{\text{char}} = \eta_{\text{dis}}$.

### A. Results of MDP-based Analysis

Recall from Section III that due to curse of dimensionality we are unable to build very large sized probability matrices. Using the version of MATLAB installed, the following is the largest size problem that can be solved before memory limits are reached: storage efficiency $\eta = 1$; 1-hour intervals; and number of states in demand, storage and cost Markov chain.
is (3,3,2). For the analysis we considered the case where the hours of “average” demand in storage, i.e. $K/E[D_i]$ is 1.64 while the hours to fully charge/discharge, i.e. $K/c_{char}$ is 2. We set the ratio of average PV supply to average demand, $E[S_i]/E[D_i]$, as 0.59 to avoid non-trivial solution (cf. the discussion in Remark 1).

Once the training data is used to model the system, we perform a simulation (sampling from fitted data). Naturally, the MDP solution would be the best and we benchmark the five heuristics ORA, OLA, HWR, TBA and NOA. The results are described in Fig. 2 where we compare the five heuristic algorithms with the y-axis denoting $(b-a)/a$ where $a$ is the minimum average cost (that is obtained by exactly solving the MDP) and $b$ is the corresponding heuristic’s average cost. While the left side display corresponds to the above conditions and models fitted with training data, the right side figure is based on 10 randomly generated examples to test. From the figure it is clear that the two ADP-based algorithms (ORA and OLA) perform the best; ORA, in particular, yields neglected performance loss compared to the optimal operation.

Based on above conditions

Avg. over 10 random cases

Fig. 2. Performance of heuristics: fraction higher than MDP solution

B. Benchmarking Heuristics Against ORA

Here we compare the five heuristics ORA, OLA, HWR, TBA and NOA but without MDP. Motivated by the excellent performance of ORA, we compare the other four algorithms against ORA in the next set of experiments. In addition we let storage charging/discharging efficiency $\eta = 0.85$, and consider both 1-hour/5-min intervals (corresponding to $N = 24$ and $N = 288$, respectively). While we will consider variations, we will mainly consider the baseline of: Hours of “average” demand in storage, i.e. $K/E[D_i]$ as 2.17; Hours to fully charge/discharge, i.e. $K/c_{char}$ as 8; Ratio of average PV supply to average demand, $E[S_i]/E[D_i]$, as 0.468. We first estimate parameters using training data. We tried various alternatives for size of the state space. We chose number of states in demand, storage and cost Markov chain to be (4,4,4). Incidentally, when we increased the number of states to (10,10,10), the results remain the same.

For the testing we used 10 days of demand, supply and cost data (from the same month as training) and let $U_0 = K/2$, i.e., the initial storage level is 50% of storage capacity. For NOA, we selected tolerance parameters $\phi_c = \phi = 0.25$

$^2$The first two ADP-based algorithms are introduced in Section III-B, HWR is proposed in Huang et al [18], and the last two algorithms are introduced in Section IV.

by testing several options. Interestingly the 0.25 value is robust and the solutions do not change with much higher or lower values. Since the ORA algorithm always outperforms other heuristics, in Fig. 3-6, we compare the four heuristic algorithms against ORA with the y-axis denoting $(b-a)/a$ where $a$ is the average cost resulting from ORA and $b$ is the corresponding heuristic’s average cost. We observe from these four figures that for almost all 1-hour cases, the ADP-based OLA algorithm performs slightly worse than ORA, and yields the minimum average cost among the four heuristics. In many 5-minute cases, however, OLA algorithm performs worse than some other heuristics (e.g., TBA). This is intuitive since OLA always treats the next stage (the next five minutes in 5-minute case) as the terminal stage and completely ignores the system dynamics after the next stage. Regarding HWR, it in general performs reasonably well, and we observe that TBA achieves a lower cost than HWR in all the 5-minute cases.

Fig. 3. Heuristics’ performance over varying storage capacity (via $K/E[D_i]$) keeping $c_{char}$ constant

Fig. 4. Heuristics’ performance over varying storage capacity (via $K/E[D_i]$) keeping $K/c_{char} = 8$

Fig. 5. Heuristics’ performance over varying hours to completely charge/discharge storage (via $K/c_{char}$) keeping $K$ fixed

In Fig. 3, we fix the maximum charging rate $c_{char}$ and vary the storage capacity $K$. Note that the ratio $K/c_{char}$ is 4, 8, 16, and 32 hours for the four cases, respectively. We observe from Fig. 3 that HWR performs reasonably well in all
cases, and yields a 2% – 10% more cost than ORA. Also, we observe that TBA achieves the minimum average cost (among the four heuristics) in the 5-minute case. In Fig. 4, we fix the ratio $K/c_{\text{char}} = 8$ and vary the storage capacity $K$. This is a more practical setting since the maximum charging rate usually grows (nearly) proportionally to the storage capacity. The performance gap between ORA and the four heuristics increases with storage capacity.

In Fig. 5 we fix the storage capacity $K$ and vary the capacity to charging rate ratio $K/c_{\text{char}}$. The parameter setting in our simulation is motivated by the development of fast-charging batteries [29]. We observe from Fig. 5 that the performance of HWR heavily depends on the ratio $K/c_{\text{char}}$: it achieves almost the same performance as ORA when this ratio is larger than 8 (i.e., it takes more than 8 hours to fully charge the storage); however, for fast-response storage devices with $K/c_{\text{char}} \leq 2$, both ORA and OLA significantly outperform HWR.

Finally, in Fig. 6, we vary the ratio of average demand to average solar supply ($E[S_t]/E[D_t]$) while fixing the other parameters. The parameter setting is motivated by the fast growing installment of solar panels on the consumer side. We note that as the solar penetration increases, ORA still outperforms the other four heuristics (including OLA), especially in the 5-minute case.

C. The Value of Storage and PV

There are costs to install a solar PV system and/or an ESD. A natural question to ask is whether the PV and/or ESD installation was worth it. For that we consider two parameters: value of storage and value of PV and storage. We use the same test data as the previous sub-section and policy based on ORA. In Fig. 7-10, the y-axis denotes $(a - b)/a$ where $b$ is the average cost using both PV and storage, while $a$ in the left bars correspond to the (optimal) use of only PV, and $a$ in the right bars correspond to the use of neither PV nor storage.

The left bars present the fractional cost savings due to the operation of storage, and can be therefore viewed as an illustrator on the value of storage. Similarly, the right bars illustrate the value of storage and PV. We observe from Fig. 7-10 that the value of storage is much higher in 5-minute cases, due to the higher variability in costs under 5-minute real-time pricing than that under hourly pricing.

As shown in Fig. 7, the values of storage and PV do not increase appreciably with increase in storage size (without increasing rates of charging/discharging). We observe from Fig. 8 that the value of storage increases sharply with the storage capacity $K$, when the maximum charging rate $c_{\text{char}}$ grows in proportional to $K$. We observe from Fig. 9 that the values of storage and PV increase with the maximum charging/discharging rate of the storage. Fig. 10 shows that the values of storage and PV increase with average solar PV generation.

VI. CONCLUSION

Although deceptively easy to state, the problem of determining energy mix from the grid, renewable source and storage device is a fairly complex one to solve. We explored five heuristics ORA, OLA, HWR, TBA and NOA, based on approaches ranging from Markovian models, to hybrid methods based on statistics, optimization and heuristics, to those that require no historical data.
The following were our findings. ORA outperforms the other four heuristics in all cases, and at the same time, requires the most computing power among the five heuristics. Except when charging/discharging rates are high, in general all algorithms perform reasonably well for 1-hour data. Indeed, another ADP-based algorithm OLA performs only slightly (about 1%) worse than ORA in all 1-hour cases, and requires much less computation than ORA. For many 5-minute cases, however, ORA significantly outperforms OLA; this is intuitive, since OLA always treats the next stage (the next five minutes in this case) as the terminal stage and completely ignores the system dynamics after the next five minutes. TBA performs well in some 5-minute cases. Tuning tolerance and coefficient parameters had virtually no effect on NOA. HWR is an easily implementable algorithm that needs no training. Its performance to a large extent depends on the number of hours to fully charge storage (i.e. $K/c_{\text{char}}$). It achieves almost the same average cost as ORA when $K/c_{\text{char}}$ is large (e.g. $> 8$). On the other hand, the two ADP-based algorithms, ORA and OLA, significantly outperform HWR for the case with $K/c_{\text{char}} \leq 2$.

Under the one-step roll-out algorithm (ORA) and hourly pricing, value of storage is not too high with $K/c_{\text{char}} = 8$. Value of storage is much higher in 5-minute cases, because the cost is much more fluctuating under 5-minute real-time pricing than that under hourly pricing. Value of storage would improve greatly if the storage size increases along with speed of charging and discharging. As solar penetration goes higher, storage has more value.

**REFERENCES**


