

An Aggregate Model of the Flexible Energy Demand of Thermostatically Controlled Loads with Explicit Outdoor Temperature Dependency

Kári Hreinsson
ECEE, Arizona State University
kari.hreinsson@asu.edu

Anna Scaglione
ECEE, Arizona State University
anna.scaglione@asu.edu

Mahnoosh Alizadeh
ECE, UC Santa Barbara
alizadeh@ucsb.edu

Abstract

In this paper we describe an aggregate model of Thermostatically Controlled Loads (TCLs) for Demand Response (DR) scheduling that, through a new approximation, makes explicit the dependency between the feasible control region and the time series of outdoor temperatures. In turn, the model can easily account for non-constant, stochastic temperatures during the control period, expressing the feasible load control through a set of linear equations and constraints with stochastic parameters. To highlight this feature we present a stochastic optimization formulation for the management of the DR-TCL and compare it with its deterministic counterpart, and with various equivalent models aimed at reducing the complexity of the constraints in the market optimization.

1. Introduction

TCLs are among the most promising candidate appliances for DR programs [1], as they can offer considerable flexibility and have virtually no ramping limits unlike conventional spinning generation. The first simplified dynamical model for a TCL population was introduced in [2]. Its aim was to capture the rebound peak observed after TCLs were interrupted during an emergency. While some of the recent work is concerned with the response of TCLs to real-time pricing (e.g. [3, 4, 5, 6, 7]), in this paper we focus on direct load control models that are decoupled from the economic signals that entice consumer participation. Instead of considering load curtailment, most of the recent work on DR of TCLs is based on changing thermostat set-points in order to adjust the load profile [8, 9, 10, 11, 12] and realize a certain collective response for load following, regulation and frequency control. Most population dynamics models in the literature are reminiscent of the model in [2]. In such models, TCLs are assumed to have a certain dead-band; the control consists of switching fractions of the TCLs population

between the ON and OFF states prematurely, relative to the time they hit the corresponding temperature threshold, in order to create the intended deviation from the otherwise-uncontrolled load profile. Work refining this basic idea is in e.g. [9], [10], [13], [14], [15]. Another relevant line of work has used battery models to capture the TCL population. Variants of this idea have been equivalent thermal battery models that are controlled to follow the least costly control trajectory [16], and generalized battery model [17] with lower and upper bounds for electric power consumption. For the latter, the control is achieved by aggregating and looping through individual TCLs in a priority-stack, switching units *early* until the desired load profile is obtained. Similar reference-signal control techniques are described by [18, 19, 20]. The authors of [21] consider generic DR *blocks* of certain power, duration and slack, prioritizing and optimizing the service of blocks to minimize costs. In [22] the authors describe a robust economic dispatch market clearing process, subject to a generic DR uncertainty bounded by a certain confidence through the scenario approach [23].

The literature cited above assumes ambient temperatures to be constant, however, knowing that temperature can both change quickly and randomly, our work attempts to capture the relationship between temperature and the range of load attainable by controlling a large number of TCLs. Most of the literature on TCL control focuses on scalable control solutions, but because of the constant ambient temperature limitation, does not offer a direct method to include TCL-DR in (stochastic) power system planning models such as Economic Dispatch (ED). Only recently work has emerged that allows for variations in temperature [24, 12, 25, 26, 27]. The authors of [24] introduce a model that tracks electric load of TCLs subject to changing ambient temperature, and later in [12] apply that model to implement a Model Predictive Controller (MPC) that reacts to a reference load signal by changing thermostat reference temperatures. A recent contribution in a similar

vein to ours is [26, 27], where the authors use a time-varying equivalent battery model for an aggregate of space heaters from [25], and solve a robust Optimal Power Flow (OPF) given the uncertainty of ambient temperature and of transmission-level wind infeed. Uncertainty is managed by ensuring feasibility against a certain number of scenarios [26] or by making the assumption that the uncertainty can be described by a jointly-Gaussian distribution [27]. In contrast to [26, 27] we do not formulate a robust problem, but a stochastic optimization, where the joint uncertainty of net-load and temperature is captured through a scenario tree, and unlike the thermal battery model of [25] we build a state-space population model which allows us to more accurately capture the complex inter-temporal relationships between temperature and constraints on power and energy. The novelty of our work can be summarized as follows: (a) we derive equivalent operational decisions/constraints of DR aggregates directly as a function of varying ambient temperature, (b) we show how to utilize this mapping in a decision problem that incorporates temperature forecasts or temperature scenario trees, capturing the uncertainty on the TCL model. Numerical results showcase the accuracy of our representation and the benefits of using our optimization framework for DR.

2. Thermal Modeling

Following most of the literature, we assume that the temperature dynamics of a heat-pump based TCL can be modeled as a first-order differential equation:

$$C\dot{\theta}(t) = (\theta_o(t) - \theta(t))R^{-1} + m(t)\eta P - \hat{\varepsilon}(t)R^{-1} \quad (1)$$

with R being thermal resistance, C thermal capacitance, θ the inside temperature, θ_o the outside temperature, η the efficiency of the heat-pump, $m(t) \in \{-1, 0, +1\}$ the operational mode of the heat pump, P its continuous electrical power rating¹ and $\hat{\varepsilon}$ denoting any noise. The authors of [11] have surveyed common values for such models. Extensions to this model include thermal mass temperatures (see e.g. [15]), adding an additional state dimension. Even though our formulation could incorporate thermal mass, in order to streamline the presentation of our work, we will focus on the simpler first order model (1).

¹Water-heaters can be described using the same principles, with an additional energy loss component describing the hot water being replaced by cold water. However, in this paper we will focus on heat-pump based TCLs, primarily because they are more dependent on external temperatures than water-boilers, which means that handling their response requires more care.

2.0.1. Discrete-Time Model Given a continuous time signal $s(t)$ and a sampling period T , in the following we adopt the conventions to denote the samples as $s[k] \triangleq s(kT)$, and their finite difference and mid-point respectively, as:

$$\dot{s}[k] \triangleq s[k] - s[k-1], \quad \bar{s}[k] \triangleq \frac{s[k] + s[k-1]}{2}. \quad (2)$$

From the RC circuit (1) we make two main observations. First, we can write the energy stored in the thermal capacity w.r.t. the reference temperature $\theta_r(t)$ as:

$$\mathfrak{e}(t) = (\theta(t) - \theta_r(t))C\eta^{-1}. \quad (3)$$

The second observation we make from (1) is that the rate at which energy is gained (or lost) from the outside environment by a TCL circuit at any given time is $(\theta_o(t) - \theta(t) - \hat{\varepsilon}(t))(\eta R)^{-1}$. Unlike most of the literature which uses θ to describe the individual state, and track the operational mode (on/off) of the heat-pump $m(t)$ defined in (1), we will use the joint energy and reference temperature $(\mathfrak{e}[k], \theta_r[k])$ as the individual TCL state. Contrary to previous models, this approach, as we will see next, yields an explicit relationship between the electric load associated to changing state and the future random ambient temperature $\theta_o[k]$. The energy $\Lambda[k]$ required by a TCL for a transition $\theta[k-1] \rightarrow \theta[k]$ ($\dot{\theta}[k]$) during the k -th interval $(k-1)T \leq t < kT$ can, using (3), be mapped into a change in state $\dot{\mathfrak{e}}[k]$ and reference temperature $\dot{\theta}_r[k]$:

$$\begin{aligned} \Lambda[k] &= \frac{C(\theta[k] - \theta[k-1])}{\eta} + \frac{1}{R\eta} \int_{(k-1)T}^{kT} [\theta(t) - \theta_o(t) + \hat{\varepsilon}(t)] dt \\ &= \dot{\mathfrak{e}}[k] + \frac{C\dot{\theta}_r[k]}{\eta} + \frac{1}{R\eta} \int_{(k-1)T}^{kT} [\theta(t) - \theta_o(t) + \hat{\varepsilon}(t)] dt. \end{aligned} \quad (4)$$

The first half of (4) captures the energy spent transitioning between different energy states, while the second half describes the energy gain over the thermal resistance due to the temperature difference. Note that since we are integrating over an interval of length T , Λ is defined in terms of energy, like the stored energy \mathfrak{e} . Also, while $\Lambda[k]$ can be either positive or negative, indicating the direction of pumping (heating or cooling), the electric energy consumption is equal to $|\Lambda[k]|$ and it is always positive. From (4), we have:

Proposition 1 Assume that, during the k -th interval:

1. The outside temperature is a non-stationary discrete time random process, i.e. $\theta_o(t) = \theta_o[k]$;
2. The reference temperature can only change between intervals, i.e. $\theta_r(t) = \theta_r[k]$ for $(k-1)T \leq t < kT$,
3. The leakage across the thermal resistance can be approximated using the average temperature, $\theta(t) \approx$

$$\bar{\theta}[k] = (\theta[k] + \theta[k-1])/2,$$

$$4. \text{ the random error: } \varepsilon[k] = (R\eta)^{-1} \int_{(k-1)T}^{kT} \hat{\varepsilon}(t) dt$$

then, the estimated energy cost is $|\Lambda[k]|$:

$$\begin{aligned} & \text{Approximate TCL energy expenditure } |\Lambda[k]|: \\ & \Lambda[k] \approx \dot{\mathbf{e}}[k] + \frac{T}{RC} \bar{\mathbf{e}}[k] + \frac{C}{\eta} \dot{\theta}_r[k] + \frac{T}{R\eta} (\bar{\theta}_r[k] - \theta_o[k]) + \varepsilon[k]. \end{aligned} \quad (5)$$

Proof: Using (3) in (4), we have:

$$\begin{aligned} \Lambda[k] &= \dot{\mathbf{e}}[k] + \frac{C}{\eta} \dot{\theta}_r[k] + \frac{1}{R\eta} \int_{(k-1)T}^{kT} [\theta(t) - \theta_o(t) + \hat{\varepsilon}(t)] dt \\ &\approx \dot{\mathbf{e}}[k] + \frac{C}{\eta} \dot{\theta}_r[k] + \frac{T}{R\eta} (\bar{\theta}[k] - \theta_o[k]) + \varepsilon[k] \\ &\approx \dot{\mathbf{e}}[k] + \frac{C}{\eta} \dot{\theta}_r[k] + \frac{T}{R\eta} \left(\frac{\eta}{C} \bar{\mathbf{e}}[k] + \bar{\theta}_r[k] - \theta_o[k] \right) + \varepsilon[k] \end{aligned}$$

which, grouping the terms, leads to (5). \square

3. Demand response model

In the model we assume that an Aggregator provides the users participating in the direct load control program with a choice of:

1. A limited set of reference temperatures \mathcal{S}_{θ_r} and of possible transitions $\mathcal{E}_{\theta_r} \subseteq \mathcal{S}_{\theta_r} \times \mathcal{S}_{\theta_r}$. For simplicity, we assume the set to contain contiguous values and allow only transitions between consecutive values with a step-size Δ_θ , i.e.

$$\mathcal{S}_{\theta_r} = \{\theta_r | \theta_r = i\Delta_\theta, i = I^{\min}, I^{\min} + 1, \dots, I^{\max}\} \quad (6)$$

$$\mathcal{E}_{\theta_r} = \{\dot{\theta}_r | \dot{\theta}_r = (i - j)\Delta_\theta, |i - j| \leq 1, (i\Delta_\theta, j\Delta_\theta) \in \mathcal{S}_{\theta_r}\} \quad (7)$$

where I^{\min}/I^{\max} denote the lowest and highest available state. Note that we assume the sequence $\theta_r[k]$ to be *user controlled*, and that a smart thermostat can translate the individual wish for a sharp reference temperature change to a gradual one, honoring \mathcal{E}_{θ_r} .

2. A temperature dead-band B , such that $|\theta(t) - \theta_r(t)| \leq B$; correspondingly the energy states are bounded by:

$$|\mathbf{e}(t)| \leq E \triangleq BC\eta^{-1}. \quad (8)$$

In controlling energy transitions, we discretize the energy state-space with a step-size of $\Delta_e = E/(2U+1)$, where U is the number of steps the energy can make upward or downward, as we assume the state-space to be symmetric (extending equally far below and above the reference temperature that corresponds to $\mathbf{e} = 0$). In addition, we assume that transitions can only cover a certain number of contiguous energy states. The sets of

states and state transitions are, respectively:

$$\mathcal{S}_{\mathbf{e}} = \{\mathbf{e} | \mathbf{e} = \Delta_e u, u = 0, \pm 1, \dots, \pm U\} \quad (9)$$

$$\mathcal{E}_{\mathbf{e}} = \{(\dot{\mathbf{e}}, \bar{\mathbf{e}}) | |\Lambda[k]| \leq TP\} \quad (10)$$

where $\mathcal{E}_{\mathbf{e}}$ is limited by the electrical power rating P .

Note that some choices of $\dot{\theta}_r \neq 0$ may not be compatible with $\mathcal{E}_{\mathbf{e}}$ since, clearly, the possible transitions for energy and reference temperature are both constrained by $|\Lambda[k]| \leq TP$, and the desired change $\dot{\theta}_r$ may be impossible to achieve in one interval duration T , i.e. the set $\mathcal{E}_{\mathbf{e}}$ as-is may be empty. To circumvent this limitation and allowing individuals to freely choose their reference temperature, we consider the Aggregator to track the *target values* for reference temperature and stored energy, and correct the load response model to account for the delay required to meet the reference temperature change. As we consider reference temperature changes to be relatively rare, we assume that eventually the actual temperature will catch up with the target value. This implies that a more general expression to (5) is required, where the restriction $\mathcal{E}_{\mathbf{e}}$ only applies when $\dot{\theta}_r[k] = 0$ and is relaxed for $\dot{\theta}_r[k] \neq 0$, while still capturing the effect of such transitions through an energy cost term that lasts multiple periods. We will return to this point after a few useful definitions.

Energy Cost Model We define the following vectors (indicated with bold face) of normalized target energy and reference temperatures:

$$\mathbf{e}[k] = \frac{1}{\Delta_e} (e[k-1], e[k])^\top, \boldsymbol{\theta}_r[k] = \frac{1}{\Delta_\theta} (\theta_r[k-1], \theta_r[k])^\top \quad (11)$$

where for now we assume the target and true stored energy to be approximately equal, $\mathbf{e}[k] \approx e[k]$. Clearly:

$$\begin{pmatrix} \dot{e}[k] \\ \bar{e}[k] \end{pmatrix} = \Delta_e \begin{pmatrix} -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \mathbf{e}[k], \quad (12)$$

and similarly for $(\dot{\theta}_r[k], \bar{\theta}_r[k])^\top$. We also define:

$$\mathbf{Y}[k] = (\mathbf{e}[k], \boldsymbol{\theta}_r[k]), \quad (13)$$

and refer to $\mathbf{Y}[k]$ as the TCL state transition. To rewrite compactly (5), we introduce the matrix:

$$\boldsymbol{\Gamma} = \begin{pmatrix} -1 & 1/2 \\ 1 & 1/2 \end{pmatrix} \begin{pmatrix} 1 & \frac{C}{\eta} \\ \frac{T}{RC} & \frac{T}{R\eta} \end{pmatrix} \begin{pmatrix} \Delta_e & 0 \\ 0 & \Delta_\theta \end{pmatrix} \quad (14)$$

and also use the following notation:

$$\boldsymbol{\Gamma} = (\gamma_e, \gamma_{\theta_r}), \quad a = T(R\eta)^{-1}. \quad (15)$$

Now consider the case of a TCL having $\dot{\theta}_r[k] = 0$. Simple algebra shows that (5) can be rearranged as:

$$\Lambda[k] = \gamma_e^\top e[k] + \gamma_{\theta_r}^\top \theta_r[k] - a\theta_o[k] + \varepsilon[k] \quad (16)$$

$$= \text{Tr}(\mathbf{\Gamma}^\top \mathbf{Y}[k]) - a\theta_o[k] + \varepsilon[k]. \quad (17)$$

The expression (16) breaks the energy spent in four terms: (a) the first term depends on $e[k]$ and is subject to the control action of the Aggregator; (b) the second term, function of $\theta_r[k]$, is the operating mode chosen by the user; (c) the third term depends on the random $\theta_o[k]$, and is due to Mother Nature; (4) the fourth is the random error $\varepsilon[k]$.

For $\theta_r[k] \neq 0$, we assume that TCLs changing their reference temperature can spread the energy expenditure for the transition over Q_k periods, as necessary to fulfill the total energy requirement, with each step not exceeding the power rating of P times T . Hence, the energy cost at time k for a TCL can, in general, be expressed as follows²:

Energy spent for transition $\mathbf{Y}[k]$ at temperature $\theta_o[k]$:

$$|\Lambda[k]| = \left| \sum_{q=0}^{Q_k} H[k-q, q] + \varepsilon[k] \right| \quad (18)$$

where the energy spent at each interval is:

$$H[k, q] = \begin{cases} PT & 0 \leq q < Q_k \\ |\text{Tr}(\mathbf{\Gamma}^\top \mathbf{Y}[k]) - a\theta_o[k]| - qPT & q = Q_k \\ 0 & \text{otherwise} \end{cases},$$

and the number of intervals required is:

$$Q_k = \left\lfloor \frac{|\text{Tr}(\mathbf{\Gamma}^\top \mathbf{Y}[k]) - a\theta_o[k]|}{PT} \right\rfloor.$$

Clearly, when $Q_k = 0$ this reduces to the expression in (17). It is important to notice the explicit dependency of the energy spent (load response) from the random ambient temperature.

3.1. Aggregate Quantized Population Model

Let us neglect the effect of $\varepsilon[k]$. The population model is built with two standard steps. The first consists of clustering of the parameters (R, C, η) , so that the matrix $\mathbf{\Gamma}$ and coefficient a in (16) are chosen from a restricted set that approximates well the most common TCL characteristics, i.e. $\mathbf{\Gamma} \mapsto \mathbf{\Omega}(\mathbf{\Gamma}) \in \mathcal{S}_\Gamma$, $a \mapsto \mathbf{\Omega}(a) \in \mathcal{S}_a$:

$$\mathbf{\Omega}(\mathbf{\Gamma}) = \underset{\hat{\mathbf{\Gamma}}_s \in \mathcal{S}_\Gamma}{\text{argmin}} \|\hat{\mathbf{\Gamma}}_s - \mathbf{\Gamma}\|, \quad (19)$$

²Naturally, all the steps that correspond to a certain reference temperature transition will have the same sign.

and similarly for $\mathbf{\Omega}(a)$. We enumerate all pairs $(\hat{\mathbf{\Gamma}}_s, \hat{a}_s) \in \mathcal{S}_\Gamma \times \mathcal{S}_a$ with an index $s = 1, \dots, S$. Note that this step yields errors in the load representation, which we will assume to be bounded. The second consists in quantizing the action space:

$$\mathbf{q}(e) \mapsto (u, v)^\top, \mathbf{q}(\theta_r) \mapsto (i, j)^\top, \mathbf{q}(\mathbf{Y}) \mapsto \begin{pmatrix} u & i \\ v & j \end{pmatrix} \quad (20)$$

where \mathbf{q} is the quantization operator. In this way, the cost of the p th TCL in the population \mathcal{P} is approximated as follows:

$$\hat{\Lambda}^{(p)}[k] = \text{Tr}(\mathbf{\Omega}(\mathbf{\Gamma}^p)^\top \mathbf{q}(\mathbf{Y}^p[k])) - \mathbf{\Omega}(a^p)\theta_o[k], \quad (21)$$

and similarly for $\hat{H}^{(p)}[k, q]$.

3.1.1. Population model To describe the expected load from the approximate aggregate model, it is convenient to break the population into the groups $\mathcal{P}_s, s = 1, \dots, S$ that belong to the different clusters for the TCL parameters:

$$\ell[k] = \sum_{s=1}^S \ell^{(s)}[k] = \sum_{s=1}^S \sum_{p \in \mathcal{P}_s} \mathbb{E} \left[|\hat{\Lambda}^{(p)}[k]| \right] \quad (22)$$

For simplicity, let us ignore momentarily transitions in reference temperature. In this case:

$$\ell^{(s)}[k] = \sum_{(u,v,i,j)} \left| \text{Tr} \left((\hat{\mathbf{\Gamma}}_s)^\top \begin{pmatrix} u & i \\ v & j \end{pmatrix} \right) - \hat{a}_s \theta_o[k] \right| \mathbb{E} \left[X_{uv}^{(s)ij}[k] \right]$$

where $X_{uv}^{(s)ij}[k]$ denotes the random variable equal to the population in cluster s that at time k has $\mathbf{q}(\mathbf{Y}^p[k]) \equiv \begin{pmatrix} u & i \\ v & j \end{pmatrix}$; using the $\delta(\mathbf{x})$ as an indicator function that is one only if the array $\mathbf{x} = \mathbf{0}$ and zero else, we can express this as follows:

$$X_{uv}^{(s)ij}[k] = \sum_{p \in \mathcal{P}_s} \delta \left(\mathbf{q}(\mathbf{Y}^p[k]) - \begin{pmatrix} u & i \\ v & j \end{pmatrix} \right). \quad (23)$$

In a nutshell, the expression for $\ell^{(s)}[k]$ clarifies that the Aggregator can control the expected load profile by controlling the expected transitions of the population, but that the response is a function of the ambient temperature realization $\theta_o[k]$.

To consider the general case in which customers can change their reference temperatures, we simply have to codify the response in terms of energy cost that corresponds to a particular array $\begin{pmatrix} u & i \\ v & j \end{pmatrix}$. That corresponds to:

Quantized profile for energy spent for transition $\begin{pmatrix} u & i \\ v & j \end{pmatrix}$:

$$H_{uv}^{(s)ij}[k, q] = \begin{cases} P^{(s)}T & 0 \leq q < Q_k^{(s)} \\ \lambda_{uv}^{(s)ij}(\theta_o[k]) - qP^{(s)}T & q = Q_k^{(s)} \\ 0 & \text{otherwise} \end{cases} \quad (24a)$$

where the total energy spent for the transition is:

$$\lambda_{uv}^{(s)ij}(\theta_o) = \left| \text{Tr} \left((\hat{\Gamma}_s)^\top \begin{pmatrix} u & i \\ v & j \end{pmatrix} \right) - \hat{a}_s \theta_o \right| \quad (24b)$$

and the duration of the load response is:

$$Q_k^{(s)} = \left\lfloor \frac{|\lambda_{uv}^{(s)ij}(\theta_o[k])|}{P^{(s)}T} \right\rfloor. \quad (24c)$$

leading to the general expression:

$$\ell^{(s)}[k] = \sum_{(u,v,i,j)} \sum_{q=0}^{+\infty} H_{uv}^{(s)ij}[k-q, q] \mathbb{E} \left[X_{uv}^{(s)ij}[q] \right], \quad (25)$$

indicating that the model has memory.

4. Control model

The key to scalability of the control lies in assuming that the Aggregator can broadcast commands that are cluster specific, but not customer specific, by controlling only the expected population trajectory as opposed to its exact values. Hence, the decision variable for the Aggregator are:

$$D_{uv}^{(s)ij}[k] \triangleq \mathbb{E} \left[X_{uv}^{(s)ij}[k] \right], \quad (26)$$

and the aggregator objective is to shape the expected load, which is a linear function of (26) (c.f. (25)):

Forecast of flexible TCL load response:

$$\ell[k] = \sum_{s=1}^S \ell^{(s)}[k] \quad (27a)$$

$$\ell^{(s)}[k] = \sum_{(u,v,i,j)} \sum_{q=0}^{+\infty} H_{uv}^{(s)ij}[k-q, q] D_{uv}^{(s)ij}[q] \quad (27b)$$

In plain English, each $D_{uv}^{(s)ij}[k]$ is the expected number³ of TCLs in cluster s transitioning from state u to state v and reference temperature i to j over period k .

Randomized control policy The expectation in (26) can be made an explicit function of the transition

³While not explicitly stated, since the policy is randomized, the case where the population participants vary can be handled with no change in the optimization formulation, by a scaling factor equal to the probability that nodes exit the control.

probabilities $\pi_k^{(s)ij}(v|u)$ that are specified as commands by the Aggregator for a randomized policy. In fact, the instructions are the probabilities for changing normalized energy from value u to value v . This presumes that the execution of the commands consists of choosing at random to switch normalized energy u to v and reference temperatures i to j with probability $\pi_k^{(s)ij}(v|u)$. Let $\Pi_k^{(s)i}(u)$ denote the *state* probability, i.e. the probability that a TCL is in normalized energy state u and normalized reference temperature i . Under the randomized control policy $\delta \left(\mathbf{q}(\mathbf{Y}^P[k]) - \begin{pmatrix} u & i \\ v & j \end{pmatrix} \right)$ is a Bernoulli random variable with probability $\pi_k^{(s)ij}(v|u) \Pi_{k-1}^{(s)i}(u)$ of being equal to 1. This implies that:

$$D_{uv}^{(s)ij}[k] = |\mathcal{P}_s| \pi_k^{(s)ij}(v|u) \Pi_{k-1}^{(s)i}(u). \quad (28)$$

From Chapman-Kolmogorov theorem for Markov chains:

$$\Pi_k^{(s)j}(v) = \sum_{(u,i)} \pi_k^{(s)ij}(v|u) \Pi_{k-1}^{(s)i}(u), \quad (29)$$

which, in turn, implies that:

$$|\mathcal{P}_s| \Pi_k^{(s)j}(v) = \sum_{(u,i)} D_{uv}^{(s)ij}[k] \quad (30)$$

Combined with (26), equation (30) means that the Aggregator can evaluate the randomized policy values based on the optimum values of $D_{uv}^{ij}[k]$ as follows:

$$\pi_k^{(s)ij}(v|u) = \frac{D_{uv}^{(s)ij}[k]}{\sum_{(u',i')} D_{u'u}^{(s)i'i}[k-1]}. \quad (31)$$

The values of $\pi_k^{(s)ij}(v|u)$ are the instructions that are broadcast to the TCLs to plan their switching.

4.1. Feasible action space and its representation

By deciding the values for $D_{uv}^{(s)ij}[k]$ over the horizon, the Aggregator can shape the expected aggregate load of (27a) within a feasible region determined by the constraints that exist on $D_{uv}^{(s)ij}[k]$. The following are known to the Aggregator before solving for D over a horizon KT :

1. A temperature forecast or scenario, $\theta_o[k]$.
2. The initial population of each cluster, and $|\mathcal{P}_s| \Pi_0^{(s)i}(u)$, i.e. the initial spread of individuals across reference temperatures (i) and states (u), for all s, i, u .
3. The thermostat schedule of all individuals; more

specifically, the number of TCLs going from one reference temperature to another at time k :

$$\rho^{(s)ij}[k] = \sum_{(u,v)} X^{(s)ij}[k] = \sum_{p \in \mathcal{P}_s(u,v)} \delta \left(q(Y^p[k]) - \begin{pmatrix} u & i \\ v & j \end{pmatrix} \right) \quad (32)$$

The feasible action space is described, in order, by the following constraints on $D_{uv}^{(s)ij}[k]$: (33a) indicates that population transitions are non-negative; (33b) comes from the fact $\sum_{(v,j)} \Pi_k^{(s)j}(v) = 1$ and (29), and can be interpreted as the conservation of population mass; (33c) comes considering (31) and because $\sum_{(v,j)} \pi_k^{(s)ij}(v|u) = 1$; (33d) comes from accounting for the thermostat plans of the TCL population; (33e) forces the model to spread evenly the population transitioning from i to j throughout the departing reference temperatures state-space; finally, (33f) expresses the Aggregator restrictions on the reference temperature state transitions, as discussed in Section 3. Note that the set (33g) improves on (10) by forcing those individual changing $\theta_r[k-1] \rightarrow \theta_r[k]$ to arrive in the top/bottom of the dead-band, depending on the reference temperature change.

Feasible action space for TCL population:

$$D_{uv}^{(s)ij}[k] \geq 0 \quad \forall s, i, j, u, v \quad (33a)$$

$$\sum_{(v,j)} \sum_{(u,i)} D_{uv}^{(s)ij}[k] = |\mathcal{P}_s|; \quad \forall s \quad (33b)$$

$$\sum_{(v,j)} D_{uv}^{(s)ij}[k] = \sum_{(u',i')} D_{u'u}^{(s)i'i}[k-1] \quad \forall s, i, u \quad (33c)$$

$$\rho^{(s)ij}[k] = \sum_{(u,v)} D_{uv}^{(s)ij}[k] \quad \forall k, i, j. \quad (33d)$$

$\forall k \in \mathcal{K}$, and $\forall s, u, i, j$ such that $i \neq j$ and $\rho^{(s)ij}[k] > 0$:

$$\sum_{(v)} D_{uv}^{(s)ij}[k] = \frac{\rho^{(s)ij}[k]}{\sum_{(j')} \rho^{(s)ij'}[k]} \sum_{(j',v)} D_{uv}^{(s)ij'}[k] \quad (33e)$$

$$D_{uv}^{(s)ij}[k] = 0 \quad \forall \begin{pmatrix} u & i \\ v & j \end{pmatrix} \notin \mathcal{E}^{(s)}(\theta_o[k]), \quad (33f)$$

where:

$$\mathcal{E}^{(s)}(\theta_o) = \left\{ \begin{pmatrix} u & i \\ v & j \end{pmatrix} \left(\left| \text{Tr} \left(\hat{\Gamma}_s \begin{pmatrix} u & i \\ v & j \end{pmatrix} \right) - \hat{a}_s \right| \leq PT \text{ and } i = j \right) \cup \left(v = (i-j)U \text{ and } |i-j| = 1 \right) \right\} \quad (33g)$$

Alternatively, in vector form, we refer to the vector of the TCL load as $\ell \in \mathbb{R}^K$, and use D for the tensor of the decision variables, i.e. all values $D_{uv}^{(s)ij}[k]$, $\forall s, u, v, i, j, k$ in their respective sets. We can describe the constrained problem that includes (27) and

(33) generally as $\ell \in \mathcal{L}$ where:

$$\mathcal{L}(\theta_o) = \{\ell \mid \ell(\theta_o) = M(\theta_o)D, A(\theta_o)D \leq b, D \geq 0\} \quad (34)$$

where $\ell \in \mathbb{R}^K$, $(A(\theta_o), b)$ represent (33), while the decision space of D is high dimensional, with $\dim(D) = S \cdot \sum_k |\mathcal{E}^{(s)}(\theta_o[k])|$, and $M(\theta_o)$ maps a decision to load for a particular temperature realization.

4.2. Reduction of Constraints

The complexity of the full model described remains quite high, making it undesirable for direct inclusion in conventional operational models, such as a Unit Commitment (UC) or OPF. To make the life of Independent System Operators (ISOs) easier, the expectation is often that such models can fit into existing market structures; for example, that they behave and are constrained like a generator. The problem of fitting the generator model to an aggregate of TCLs, is similar to that of fitting such models to large scale storage, to bid in the energy markets. Conservative approaches use a price forecast to determine the exact amount offered to sell or buy ahead of time, or limit the bid min/max to avoid energy constraints violations. Riskier approaches increase the min/max range to better leverage the market, resorting to real-time markets or paying penalties when deliverability becomes an issue.

DR models, in general, require defining a load baseline profile $\ell^b = (\ell^b[0], \dots, \ell^b[K-1])$. Using the model in Section 4 one can estimate the baseline consumption of a TCL aggregate by constraining to zero the stored energy states and estimate the aggregate load that corresponds to all individuals following the reference temperature closely⁴. With that in mind, below we propose different TCL-DR representations. Both of them will depend on the random realization of the temperature profile.

4.2.1. Negative Generator Model The output of this model is a set of constraints $\ell^{\min} \in \mathbb{R}^K$ and $\ell^{\max} \in \mathbb{R}^K$, where we allow the ISO to choose any load profile within the hyper-rectangle $\ell^{\min} \leq \ell \leq \ell^{\max}$. To obtain the min/max values one can, for example, solve:

$$\max_{\ell^{\min}, \ell^{\max}} \|\ell^{\max} - \ell^{\min}\| \quad \text{s.t.} \quad \ell^{\min} \leq \ell^b \leq \ell^{\max}. \quad (35)$$

The choice of norm (more generally, *distance*) should reflect the general objective of obtaining a vector that

⁴Note that the Aggregator can still control the shape of the baseline, pricing the customer consumption based on their choice of reference temperature. However, we assume that the construction of bids takes the ℓ^b as a given.

is as far away from the baseline load, so as to have the largest possible generator capacity range, with the exact measure possibly differing on case-by-case basis, since particular periods may be more profitable. For a TCL aggregate the generator ramping constraint is not helpful, as there is little inertia preventing sharp jumps in consumption between periods.

4.2.2. Arbitrary Constraint Set Model

Unfortunately, the practice of mapping flexible load to a negative generator is bound to capture a very small portion of the full feasible region. It is important to notice, however, that ISOs are not concerned with the individual decisions \mathbf{D} , but only with ℓ . Also, many decisions \mathbf{D} give the same load ℓ , i.e. the mapping is not injective. Finding a direct and compact description of the feasible space \mathcal{L} is, however, not a simple problem, which deserves its own treatment. Here, for the sake of comparison, we propose the following heuristic. Given a particular temperature profile θ_o , we use a set of cost vectors $\mathcal{C} = \{c_1, \dots, c_W\}$ and solve:

$$\arg \min_{\ell} c_w^T \ell \quad \text{s.t.} \quad \ell \in \mathcal{L}(\theta_o). \quad (36)$$

If we take the convex hull of the vertices $[\ell_1, \dots, \ell_W]$, we obtain a region $\mathcal{L}' \subset \mathcal{L}$. As $W \rightarrow \infty$, $\mathcal{L}' \rightarrow \mathcal{L}$; the problem becomes choosing the \mathcal{C} that reveal the vertices of \mathcal{L} that are *most efficient*. First, note that the problem (36) can be solved in parallel. Secondly, the problem (36) is simply a conventional minimization problem; given a price vector c_w we obtain the load profile ℓ_w that is the least costly way to meet the constraint set (34). If \mathcal{C} are price forecasts that contain the future price, we know that the optimal load profile will be contained in the set \mathcal{L}' , making them reasonable choices for search directions. We convert the descriptions of these feasible regions from vertices (V) of the polytope to a half-space (H) representation, i.e.:

$$\mathbf{A}' \ell' \leq \mathbf{b}'. \quad (37)$$

In principle this conversion is an NP-hard problem [28], but with a reasonably small set of vertices in \mathcal{L}' and $K \leq 24$ the conversion is quick on a modern computer. The constraints (37) can be placed directly into the standard OPF/ED problems, and as long as the optimal price (or a scalar thereof) of the ISO formulation is contained in the set \mathcal{C} , we can obtain the optimal solution ℓ as if we directly placed (34) into the formulation. Note that constraints on energy, such as those that may come out of TCL equivalent battery models are a special case of this constraint set \mathbf{A}' , but imposing that limitation on the structure

of \mathbf{A}' only decreases the expressible flexibility which is particularly problematic if the temperature θ_o is changing throughout the time horizon.

5. A Stochastic Security-Constrained Economic Dispatch (SSCED) with responsive TCLs

A straightforward way to incorporate uncertainty in power system models is to apply standard stochastic programming techniques, where one optimizes a collection of plans for different scenarios that reflect the underlying uncertainty on the future. More specifically, this section describes a rolling horizon *two stage* Security-Constrained Economic Dispatch (SCED) optimization, where we consider the joint uncertainty of net-load and outdoor temperature, while being secure against the loss of any single generator ($G - 1$). In the first stage, we make decisions based on the realized uncertainty; these decisions account for a second stage consisting of several possible future trajectories. We consider a time horizon of K intervals, each of length T seconds such that we look KT seconds into the future. We describe the formulation in *nodal form*, that is, instead of indexing variables and parameters by time k and scenario s we simply index them by the node number n , where each node has the parent node n^- . We denote the set of generators by \mathcal{G} , the set of generator outages by \mathcal{G}^{out} , and the set of TCL aggregates as \mathcal{M} , and define the vector of decision variables as $\chi = [x, \hat{x}, \bar{x}, \ell, \hat{\ell}, \bar{\ell}]$, which we seek to choose to optimize the expected cost by solving:

$$\min_{\chi} \sum_{g \in \mathcal{G}} \sum_{n \in \mathcal{V}} \pi\{n\} (C_X^g(x^g\{n\}) + \bar{C}_X^g(\bar{x}^g\{n\})) \quad (38a)$$

$$\sum_{g \in \mathcal{G}} x^g\{n\} = \sum_{m \in \mathcal{M}} \ell^m\{n\} + L\{n\} \quad \forall n \in \mathcal{V} \quad (38b)$$

$$\underline{P}^g \leq x^g\{n\} \leq \bar{P}^g - \bar{x}^g\{n\} \quad \forall n \in \mathcal{V}, g \in \mathcal{G} \quad (38c)$$

$$|x^g\{n\} - x^g\{n^-\}| \leq \bar{P}_{\text{ramp}}^g \quad \forall n \in \mathcal{V}, g \in \mathcal{G} \quad (38d)$$

$$\ell^m \in \mathcal{L}^m(\theta_o) \quad \forall m \in \mathcal{M} \quad (38e)$$

$$\ell^m - \bar{\ell}^m \in \mathcal{L}^m(\theta_o) \quad \forall m \in \mathcal{M} \quad (38f)$$

$$\sum_{g \in \mathcal{G}} \hat{x}_{g'}^g\{n\} + \sum_{m \in \mathcal{M}} \hat{\ell}_{g'}^m\{n\} = x^g\{n\} \quad \forall n \in \mathcal{V}, g' \in \mathcal{G}^{\text{out}} \quad (38g)$$

$$0 \leq \hat{\ell}_{g'}^m\{n\} \leq \bar{\ell}^m\{n\} \quad \forall n \in \mathcal{V}, m \in \mathcal{M}, g' \in \mathcal{G}^{\text{out}} \quad (38h)$$

$$0 \leq \hat{x}_{g'}^g\{n\} \leq \bar{x}^g\{n\} \leq \bar{P}_{\text{resv}}^g \quad \forall n \in \mathcal{V}, g \in \mathcal{G}, g' \in \mathcal{G}^{\text{out}} \quad (38i)$$

$$x_g^g\{n\} = 0 \quad \forall n \in \mathcal{V}, g \in \mathcal{G}^{\text{out}} \quad (38j)$$

The objective (38a) is to minimize the cost of generation $x^g\{n\}$ and reserves $\bar{x}^g\{n\}$ subject to the cost functions C_X^g and \bar{C}_X^g . Constraint (38b) describes the base-case (non-outage) balance between generation $x^g\{n\}$, TCL load $\ell^m\{n\}$ and the remaining net-load $L\{n\}$. Further,

(38c) and (38d) describe conventional min/max power and ramp constraints, where $\bar{x}^g\{n\}$ is the reserve allocation of unit g at node n , with n^- indicating the parent of node n . Constraints (38e) and (38f) ensure that the base-case TCL load and its upper reserve allocation of TCL aggregate m are contained within the feasible set \mathcal{L}^m . Constraint (38g) ensures we have enough reserves to replace any single generator outage, while (38h) and (38i) ensure that the maximum reserve allocation \bar{x}^g and $\bar{\ell}^m$ is greater or equal to the individual outage responses \hat{x}^g and $\hat{\ell}^m$. Finally (38j) ensures an outaged unit can not contribute to its own replacement reserves. The feasible region \mathcal{L}^m can be the full model described by (34), the negative/virtual generator model obtained through (35), the constraint set defined in (37), or another convex inner approximation of the actual feasible space.

Remark 1 Note that with a convex \mathcal{L}^m , constraints (38e) and (38f) are not sufficient to guarantee that any combination of the individual reserve responses $\hat{\ell}_g^m\{n\}$ are feasible. However, due to the nature of the TCL model lacking inherent ramping constraints while having prominent energy constraints, we can interpret (38f) and (38e) as bounds on not only power but also energy, and any power profile contained within these limits thus also respects this energy bound. This assumption means that the probability of falling outside the feasible region \mathcal{L}^m during some sequence of reserve events is small, but non-zero. This approximation is also apparent for generators; the standard SCED model does not properly bound generator ramps between any possible sequence of reserve events. The message here is that (a) these reserve bounds are approximations and (b) many $N - 1$ events captured by conventional SCED/SCUC models are rare.

6. Numerical Simulations

Numerical simulations were performed using Python and a collection of scientific programming libraries, as well as utilizing Gurobi 8 for solving LP models.

The load and renewable infeed data are from California ISO (CAISO); the corresponding temperature data for seven California locations from the National Oceanic & Atmospheric Administration (NOAA) [29]. From [11] we assume the population to be clustered, with a centroid reflecting any combination of the parameters $R \in \{1.5, 2, 2.5\}$ °C/kW, $C \in \{1.5, 2, 2.5\}$ kWh/°C and $S_{\theta_i} = \{69, 72, 75, 78, 81\}$ °F. From [30] we assume 7,000,000 households use heat-pump air-conditioning. In order to obtain realistic parameters for the simulations, we reverse the model in (33) and use a simple linear regression to estimate how the population is spread between R/C clusters as well as

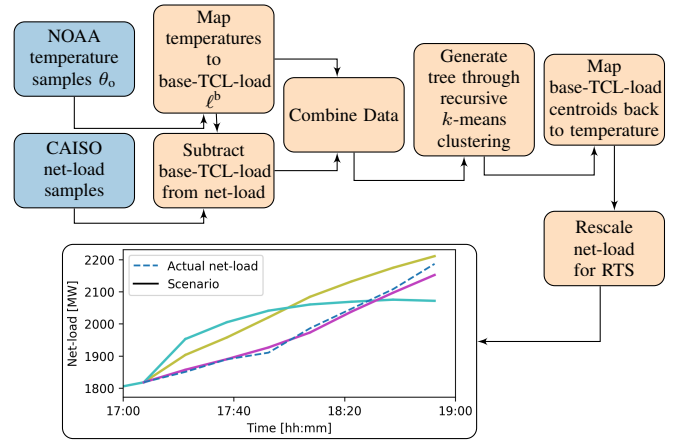


Figure 1. The process from input data to multi-variate scenario tree.

Table 1. Costs for the RTS case generators.

Unit Type	C_X	\bar{C}_X
Oil/Steam (U12)	[55.15, 57.98]	[27.57, 28.99]
Oil/CT (U20)	[127.40, 132.60]	[63.70, 66.30]
Hydro (U50)	[0.00, 0.00]	[0.00, 0.00]
Coal/Steam (U76)	[15.76, 16.40]	[7.88, 8.20]
Oil/Steam (U100)	[43.01, 44.32]	[21.50, 22.16]
Coal/Steam (U155)	[12.14, 12.64]	[6.07, 6.32]
Oil/Steam (U197)	[47.85, 49.31]	[23.93, 24.65]
Coal/Steam (U350)	[11.79, 11.79]	[5.90, 5.90]
Nuclear (U400)	[4.38, 4.47]	[2.19, 2.23]

how reference temperature changes over the course of the day. In practice, an Aggregator would know the individual $R/C/\theta_r$ parameters of its population, and would cluster based on that knowledge. The way the scenario trees were generated is illustrated in Figure 1.

Simulation Setup We solve a rolling-horizon SCED with a two hour look-ahead window and a resolution of 15 minutes ($T = 900s$, $K = 8$). For 7 sample days we build scenario trees of net-load and temperature for each starting interval, using the closest (in terms of aggregate load) 50 sample days (excluding the target day) from our original data-set of 293 summer week-days (see Figure 1). We test two variants of (38), in both cases considering any single generator outage $\mathcal{G}^{\text{out}} \leftarrow \mathcal{G}$:

- (i) A stochastic SCED with 3 future scenarios.
 - (ii) A deterministic SCED with a single forecast.
- For the SCED \leftrightarrow TCL interface we simulate:
- (a) An inflexible TCL aggregate, that consumes ℓ^b .
 - (b) A virtual/negative generator model (35).
 - (c) A set of arbitrary constraints (37).
 - (d) The full set of constraints (34) included in the SCED.
- All TCLs start at midnight in an energy neutral state ($u = 0$) and finish each rolling horizon window at $u = 0$. We use the IEEE RTS system as a starting-point for the simulations, but model it as a single bus system,

Table 2. Summary of average costs for different solutions, and the percentage savings compared with the no DR solution.

Interface	(i) Stoc. SCED	(ii) Determ. SCED
(a) No DR	3, 407, 291 (0.0%)	3, 407, 028 (0.0%)
(b) Virtual Gen	3, 300, 076 (3.1%)	3, 293, 324 (3.3%)
(c) Constr. Set	3, 101, 865 (9.0%)	3, 103, 868 (8.9%)
(d) Full Model	3, 004, 401 (11.8%)	3, 024, 382 (11.2%)

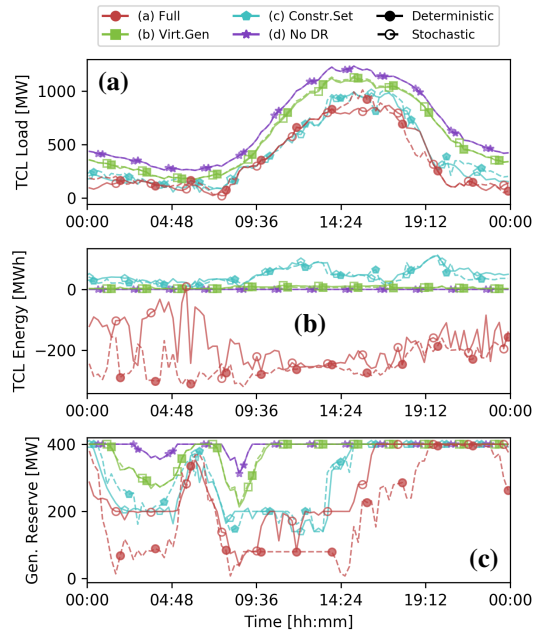


Figure 2. Using scaled CAISO data from July 20th 2017, Figure (a) shows the TCL load for the different models, (b) shows the energy stored/borrowed from the TCL aggregates, while (c) shows the reserves provided by generators.

with linear generator cost functions whose parameters are shown in Table 1.

Results Table 2 shows the average cost over the 7 sample days for different model combinations. Unsurprisingly, there is a clear trend where increased flexibility reduces costs, with the full model saving over 11% compared with no DR, and the constraint set tripling the savings from the virtual generator, from approximately 3% to 9%. Comparing the deterministic with the stochastic we see a 0.6% improvement in the stochastic results, but this improvement is largely lost when using the constraint set approximation, and slightly negative for the virtual generator. The small improvement can be explained by (a) the small uncertainty over the two-hour look-ahead horizon reflected by a Pearson autocorrelation coefficient in the range of [0.88, 0.98] for the different locations, (b) the high flexibility of the TCLs allowing the operator to

react quickly and cheaply to forecast deviations and (c) the small number of scenarios considered and the relatively primitive approach to scenario generation.

Figure 2 shows how the different models react for the sample day of July 20th 2017. First, in (a), we can see that all the solutions incorporating DR consume less power throughout the day compared with the non-controllable (no DR) counterpart. In (b), we see the aggregate energy stored in the aggregate population, where the sign indicates whether the average temperature sits above (positive) or below (negative) the reference temperature. We can see the inflexible and virtual gen. models hardly deviate from the reference temperature, while the constraint set increases it slightly to reduce thermal losses. There are more fluctuations in the full model, as it utilizes variations in temperature to cool more efficiently, but it also decreases the average temperature (“storing energy”) to allow for more reserve potential. The reserve capacity coming from generators is shown in Figure 2(c), where we see significant portion of the reserves being replaced by the TCLs as the DR model becomes more flexible, with the TCLs at times covering between 75% and 90% of the reserve requirements. The generator reserve reduction varies considerably throughout the day, but interestingly, the stochastic full model seems to be able to deliver substantially more reserves compared with its deterministic counterpart.

7. Conclusions

This paper provides a model of the aggregate feasible operation region of TCLs that allows to capture its uncertainty directly, as a function of the random outdoor temperature. The paper further shows how this model can be used, either directly or indirectly through approximations, inside power system planning models such as SCED or UC, subject to weather forecasts or a set of potential future scenarios. The numerical section explores the different modeling interfaces with a rolling-horizon SCED model, and shows how such inclusion can significantly reduce costs by decreasing overall TCL load as well as allowing TCLs to partially cover generation reserve requirements, without significant impact on the residential DR participants. Future work will target improving the interface between DR and power system models, with a special consideration for stochastic programming applications.

References

- [1] D. S. Callaway and I. A. Hiskens, “Achieving controllability of electric loads,” *Proceedings of the*

- IEEE*, vol. 99, no. 1, pp. 184–199, 2011.
- [2] C.-Y. Chong and R. P. Malhamé, “Statistical synthesis of physically based load models with applications to cold load pickup,” *IEEE transactions on power apparatus and systems*, no. 7, pp. 1621–1628, 1984.
 - [3] N. Lu and D. P. Chassin, “A state-queueing model of thermostatically controlled appliances,” *IEEE Transactions on Power Systems*, vol. 19, no. 3, pp. 1666–1673, 2004.
 - [4] S. Behboodi, D. P. Chassin, N. Djilali, and C. Crawford, “Transactive control of fast-acting demand response based on thermostatic loads in real-time retail electricity markets,” *Applied Energy*, vol. 210, pp. 1310–1320, 2018.
 - [5] J. H. Yoon, R. Baldick, and A. Novoselac, “Dynamic demand response controller based on real-time retail price for residential buildings,” *IEEE Transactions on Smart Grid*, vol. 5, no. 1, pp. 121–129, 2014.
 - [6] H. Zhong, L. Xie, and Q. Xia, “Coupon incentive-based demand response: Theory and case study,” *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 1266–1276, 2013.
 - [7] C. Zhao, J. Wang, J.-P. Watson, and Y. Guan, “Multi-stage robust unit commitment considering wind and demand response uncertainties,” *IEEE Transactions on Power Systems*, vol. 28, no. 3, pp. 2708–2717, 2013.
 - [8] S. Kundu, N. Sinitzyn, S. Backhaus, and I. Hiskens, “Modeling and control of thermostatically controlled loads,” in *Proceedings of the 2011 Power System Computation Conference*, 2011.
 - [9] D. S. Callaway, “Tapping the energy storage potential in electric loads to deliver load following and regulation, with application to wind energy,” *Energy Conversion and Management*, vol. 50, no. 5, pp. 1389–1400, 2009.
 - [10] S. Bashash and H. K. Fathy, “Modeling and control insights into demand-side energy management through setpoint control of thermostatic loads,” in *Proceedings of the 2011 American Control Conference*. IEEE, 2011, pp. 4546–4553.
 - [11] J. Mathieu, M. Dyson, and D. Callaway, “Using residential electric loads for fast demand response: The potential resource and revenues, the costs, and policy recommendations,” in *Proceedings of the ACEEE Summer Study on Energy Efficiency in Buildings*, 2012.
 - [12] N. Mahdavi, J. H. Braslavsky, M. M. Seron, and S. R. West, “Model predictive control of distributed air-conditioning loads to compensate fluctuations in solar power,” *IEEE Transactions on Smart Grid*, vol. 8, no. 6, pp. 3055–3065, 2017.
 - [13] S. Koch, J. L. Mathieu, and D. S. Callaway, “Modeling and control of aggregated heterogeneous thermostatically controlled loads for ancillary services,” in *Proceedings of the 2011 Power System Computation Conference*, 2011.
 - [14] M. Alizadeh, A. Scaglione, A. Applebaum, G. Kesidis, and K. Levitt, “Reduced-order load models for large populations of flexible appliances,” *IEEE Transactions on Power Systems*, vol. 30, no. 4, pp. 1758–1774, 2015.
 - [15] W. Zhang, J. Lian, C.-Y. Chang, and K. Kalsi, “Aggregated modeling and control of air conditioning loads for demand response,” *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 4655–4664, 2013.
 - [16] J. L. Mathieu, M. Kamgarpour, J. Lygeros, and D. S. Callaway, “Energy arbitrage with thermostatically controlled loads,” in *Proceedings of the 2013 European Control Conference*, 2013.
 - [17] H. Hao, B. M. Sanandaji, K. Poolla, and T. L. Vincent, “Aggregate flexibility of thermostatically controlled loads,” *IEEE Transactions on Power Systems*, vol. 30, no. 1, pp. 189–198, 2015.
 - [18] L. C. Totu, R. Wisniewski, and J. Leth, “Demand response of a tcl population using switching-rate actuation,” *IEEE Transactions on Control Systems Technology*, 2016.
 - [19] S. H. Tindemans, V. Trovato, and G. Strbac, “Decentralized control of thermostatic loads for flexible demand response,” *IEEE Transactions on Control Systems Technology*, vol. 23, no. 5, pp. 1685–1700, 2015.
 - [20] S. Meyn, P. Barooah, A. Busic, and J. Ehren, “Ancillary service to the grid from deferrable loads: The case for intelligent pool pumps in florida,” in *2013 IEEE 52nd Annual Conference on Decision and Control (CDC)*. IEEE, 2013, pp. 6946–6953.
 - [21] F. Elghitani and W. Zhuang, “Aggregating a large number of residential appliances for demand response applications,” *IEEE Transactions on Smart Grid*, 2017.
 - [22] H. Ming, L. Xie, M. Campi, S. Garatti, and P. Kumar, “Scenario-based economic dispatch with uncertain demand response,” *IEEE Transactions on Smart Grid*, 2017.
 - [23] M. C. Campi, S. Garatti, and M. Prandini, “The scenario approach for systems and control design,” *Annual Reviews in Control*, vol. 33, no. 2, pp. 149–157, 2009.
 - [24] N. Mahdavi, J. H. Braslavsky, and C. Perfumo, “Mapping the effect of ambient temperature on the power demand of populations of air conditioners,” *IEEE Transactions on Smart Grid*, 2016.
 - [25] J. L. Mathieu, M. Kamgarpour, J. Lygeros, G. Andersson, and D. S. Callaway, “Arbitraging intraday wholesale energy market prices with aggregations of thermostatic loads,” *IEEE Transactions on Power Systems*, vol. 30, no. 2, pp. 763–772, 2015.
 - [26] M. Vrakopoulou, B. Li, and J. L. Mathieu, “Chance constrained reserve scheduling using uncertain controllable loads part i: Formulation and scenario-based analysis,” *IEEE Transactions on Smart Grid*, 2017.
 - [27] B. Li, M. Vrakopoulou, and J. L. Mathieu, “Chance constrained reserve scheduling using uncertain controllable loads part ii: Analytical reformulation,” *IEEE Transactions on Smart Grid*, 2017.
 - [28] S. Barot and J. A. Taylor, “A concise, approximate representation of a collection of loads described by polytopes,” *International Journal of Electrical Power & Energy Systems*, vol. 84, pp. 55–63, 2017.
 - [29] H. J. Diamond, T. R. Karl, M. A. Palecki, C. B. Baker, J. E. Bell, R. D. Leeper, D. R. Easterling, J. H. Lawrimore, T. P. Meyers, M. R. Helfert, G. Goodge, and P. W. Thorne, “US climate reference network after one decade of operations: Status and assessment,” *Bulletin of the American Meteorological Society*, vol. 94, no. 4, pp. 485–498, 2013.
 - [30] Energy Information Administration, U.S. Department of Energy, “2009 Residential Energy Consumption Survey (RECS),” 2009.