Noise and Parameter Heterogeneity in Aggregate Models of Thermostatically Controlled Loads

Md Salman Nazir * **Ian A. Hiskens** *

* *Department of Electrical and Computer Engineering*
University of Michigan, Ann Arbor
(E-mail: mdsnazir,hiskens@umich.edu)

Abstract: Aggregate models are used in the analysis and control of large populations of thermostatically controlled loads (TCLs), such as air-conditioners and water heaters. The fidelity of such models is studied by analyzing the influences of noise and parameter heterogeneity on TCL aggregate dynamics. While TCLs can provide valuable services to the power systems, control may cause their temperatures to synchronize, which may then lead to undesirable power oscillations. Recent works has shown that the aggregate dynamics of TCLs can be modeled by tracking the evolution of probability densities over discrete temperature ranges or bins. To accurately capture oscillations in aggregate power, such bin-based models require a large number of bins. The process of obtaining the Markov state transition matrix that governs the dynamics can be computationally intensive when using Monte Carlo based system identification techniques. Existing analytical techniques are further limited as noise and heterogeneity in several thermal parameters are difficult to incorporate. These challenges are addressed by developing a fast analytical technique that incorporates noise and heterogeneity into bin-based aggregate models. Results show the identified and the analytical models match very closely. Studies consider the influence of model error, noise and parameter heterogeneity on the damping of oscillations. Results demonstrate that for a specific bin width, the model can be invariant to quantifiable levels of noise and parameter heterogeneity. Finally, a discussion is provided of cases where existing bin models may face challenges in capturing the influence of heterogeneity.

Keywords: Thermostatically controlled loads, Aggregate models, Noise, Heterogeneity, Oscillations, Synchronization.

1. INTRODUCTION

Thermostatically controlled loads (TCLs) offer tremendous potential for enhancing power system responsiveness in the presence of fluctuating outputs from renewable generation (see Callaway and Hiskens (2011)). Large groups of TCLs can be controlled for providing various power systems services. However, controlling large groups of TCLs may cause their temperatures to synchronize, which might then lead to undesired power oscillations (see Ihara and Schweppe (1981), Kundu and Sinitsyn (2012), Perfumo et al. (2013)). Since aggregate models are commonly used to characterize TCL dynamics and to control their aggregate power (see Bashash and Fathy (2013), Mathieu et al. (2013)), this paper focuses on analyzing the ability of such models to capture power oscillations, especially in the presence of noise and parameter heterogeneity. The need for analytical approaches to modeling is addressed.

1.1 Literature review

Physically based models of TCLs were developed initially by Chong and Debs (1979), Ihara and Schweppe (1981), Mortensen and Haggerty (1988). Malhame and Chong (1985) showed that the aggregate dynamics of TCLs can be modeled using coupled Fokker-Plank equations. More recently, Bashash and Fathy (2013) modeled the aggregate dynamics of TCLs by bilinear partial differential equations and used finite-difference based discretization. Kundu et al. (2011) modeled the steady state aggregate temperature densities for ON/OFF states and applied it to develop controllers for set point variation. Koch et al. (2011) presented a Markov chain based statistical modeling approach to describe the evolution of the probability masses over "temperature bins". Each bin is defined by a specified temperature range and whether the TCLs in it are in ON/OFF states. Such a model will from here on be referred to as a *bin model*. Mathieu et al. (2013) used a similar model to design state estimators and develop control strategies for large population of TCLs. Recently, Ghaffari et al. (2014) also modeled temperature evolution over discretized temperature ranges using advection-diffusion PDEs and studied the model's numerical stability. A large body of recent work has been utilizing such bin models in designing new control techniques (see for example Crocker and Mathieu, 2016, Esmaeil Zadeh Soudjani and Abate (2015), Nazir et al. (2016)). Our modeling and analysis framework is consistent with these existing bin models.

TCLs are heterogeneous and their dynamic behavior is affected by several sources of uncertainty. Therefore, aggregate models need to capture those influences. Simulations have shown that heterogeneity and noise can add additional desirable damping to the system and the steady state is therefore reached faster (see Malhame and Chong (1985), Callaway (2009), Perfumo (2013), Ghaffari et al. (2014)). Perfumo (2013) provides an

⋆ The authors gratefully acknowledge the contribution of the Natural Sciences and Engineering Research Council of Canada (NSERC) and the U.S. National Science Foundation through grant CNS-1238962.

analytical model to approximate the influence of heterogeneity, specifically for the thermal capacitance parameter. Ghaffari et al. (2014) analyzes the impact of heterogeneity by considering different sub-groups of homogeneous TCLs. Several other studies have also used the bin model to consider probabilistic distributions of various parameters governing the TCL population (see Koch et al. (2011), Mathieu et al. (2013)). However, there remain ambiguities concerning the performance of bin models when considering heterogeneity and noise.

To model the system behavior using a bin model, it is necessary to find the Markov transition matrix that governs the population dynamics. While system identification (SI), based on sampling from Markov chains, may be used to identify the transition matrix (see Koch et al. (2011), Mathieu et al. (2013)), it can be computationally intensive for reasonably large systems. Moreover, the process must be repeated whenever parameters change. While an analytical approach was presented by Koch et al. (2011), it ignored noise and makes restrictive assumptions on parameter heterogeneity. Hence, the objective of this paper is to provide an analytical (though approximate) approach to incorporating noise and heterogeneity into bin models. This provides a basis for analyzing their impact on the population performance.

1.2 Contributions

The primary contributions of the paper are:

- We highlight some key challenges in the use of system identification (SI) based techniques to obtain the state transition matrix and discuss limitations in existing analytical techniques. We then develop an analytical technique to overcome such limitations, mainly by incorporating noise and parameter heterogeneity in the bin model. Simulations show that the aggregate power output obtained using the identified and analytical models match closely. For large systems, solutions were obtained within seconds using the proposed analytical technique, compared to several hours using the SI approach, thus significantly reducing the computational burden.
- Using the analytical modeling insights, it is shown that state transition matrices can be invariant to changes in noise, depending on the width of the temperature bins.
- The influence of model error, arising from temperature discretization and probabilistic transitions, is highlighted. A discussion suggests that certain temperature initial conditions (such as if TCL temperatures are synchronized) may require higher order bin models for accurately predicting oscillatory response.
- Finally, the relative effects of model error, noise and heterogeneity are compared in terms of their influence on aggregate power oscillations and damping. Important distinctions in the modeling of noise and heterogeneity are highlighted.

1.3 Organization

The remainder of the paper is organized as follows. Section 2 presents the model preliminaries, discusses the various ways of obtaining the aggregate model and highlights their advantages or disadvantages. Section 3 presents an analytical approach to modeling the A-matrix taking into account noise and heterogeneity. Section 4 provides several analytical as well as experimental results on model performance, invariance properties of the bin model, and compares the influences of model error, noise and heterogeneity. Section 5 concludes by summarizing the main contributions and future directions.

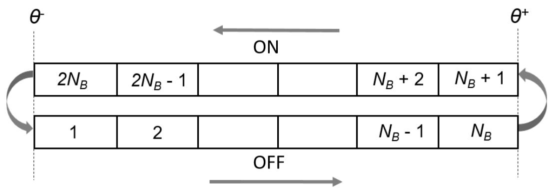

Fig. 1. Temperature bin model.

2. MODELING

2.1 TCL model preliminaries

Consider a large population of N_{tcl} TCLs. The set-point, dead-band, internal and ambient temperatures corresponding to each load i are denoted $\theta_{s,i}$, δ_i, θ_i and $\theta_{a,i}$, respectively (°C). Each load can be modeled as a thermal capacitance C_i (kWh/°C) in series with a thermal resistance R_i (°C/kW). Finally, the binary variable m_i denotes the on or off status of the load, and P_i (kW) is the energy transfer rate when a cooling (or heating) TCL is switched on. The dynamic behavior of each TCL can be modeled using a first-order difference equation Mortensen and Haggerty (1988), Callaway (2009),

$$\theta_{i,t+h} = a_i\theta_{i,t} + (1-a_i)(\theta_{a,i} - m_{i,t}\theta_{g,i}) + w_{i,t} \tag{1}$$

where h is the time step, $a_i = \exp(-h/C_iR_i)$ is the parameter governing the thermal characteristics of each TCL, $\theta_{g,i} = P_iR_i$ is the temperature gain when a cooling TCL is on, and $\mathbf{w}$ is a noise process (usually modeled as white noise with zero mean and known standard deviation). To consider heterogeneity, the parameters P, C and R may follow particular probability distributions, rather than all TCLs having the same values. Additionally, the variable $m_{i,t}$ for TCL i captures the TCL's switching behavior according to,

$$m_{i,t+h} = \begin{cases} 0, & \text{if } \theta_i < \theta_{min} \\ 1, & \text{if } \theta_i > \theta_{max} \\ m_{i,t} & \text{otherwise} \end{cases} \tag{2}$$

where $\theta_{min} = \theta_{s,i} - \delta_i/2$ and $\theta_{min} = \theta_{s,i} + \delta_i/2$.

With coefficients of performance η_i, the aggregate electrical power consumed is given by,

$$P_t^{tot} = \sum_{i=1}^{N_{tcl}} m_{i,t}P_i/\eta_i. \tag{3}$$

Typical parameter values are given in Table 1 (similar to Callaway (2009), Bashash and Fathy (2013)).

For dealing with a population of TCLs, Kundu et al. (2011), Bashash and Fathy (2013) and Mathieu et al. (2013) showed

Table 1. TCL parameters.

Parameter	Meaning	Value	Units
θ_s	Temperature set-point	20	$^\circ C$
δ	Temperature dead-band width	0.5	$^\circ C$
θ_a	Ambient air temperature	32	$^\circ C$
P	Energy transfer rate	14	kW
R	Thermal resistance	2	$^\circ C$/kW
C	Thermal capacitance	10	kWh/$^\circ C$
η	Coefficient of performance	2.5	

how to reduce the model from N_{tcl} copies of (1) to a simplified aggregate model by using discrete temperature bins. A bin is defined by its temperature bounds as well as whether the TCLs in it are on or off. As shown in Fig. 1, the normalized dead-band can be discretized by uniformly dividing it into N_B bins, so in total $2N_B$ bins are used. The system's state x describes the population density of each bin. The evolution of the population can be expressed as,

$$x(k+1) = Ax(k). \tag{4}$$

2.2 Obtaining the A-matrix

Matrix A is the transpose of the Markov transition matrix that is composed of the probabilities of transitioning from bin to bin. Let p_{ij} be the probability of transitioning from bin i to bin j. The A-matrix has the form,

$$A = \begin{bmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,2N_B} \\ p_{2,1} & p_{2,2} & \cdots & p_{2,2N_B} \\ \vdots & \vdots & \ddots & \vdots \\ p_{2N_B,1} & p_{2N_B,2} & \cdots & p_{2N_B,2N_B} \end{bmatrix}. \tag{5}$$

Two approaches to obtaining the A-Matrix are summarized below.

System identification (SI) approach: By simulating a system of TCLs using (1) and using full state information from all of the simulated TCLs, the number of TCL transitions from each starting bin to each ending bin can be counted. By collecting a large number of samples, a Monte-Carlo estimate of the Markov transition matrix can be obtained. To ensure conservation of probability mass, the resulting matrix must be normalized such that each column sums to one. Finally, the transpose of the identified Markov matrix is the desired A-matrix (see Koch et al. (2011), Mathieu et al. (2013)).

The main issue related to the SI approach is that it can be computationally intensive, especially when one aims to obtain a higher order A-matrix. With no noise or heterogeneity, we ran experiments with 1000 TCLs over 24 hours with $h = 10$ second time-steps and obtained the A-matrices for N_B=10, 20, 40, 80, 160, and 320 bins. The process of obtaining the A-matrix estimate with small N_B was relatively fast, for $N_B = 10$ it took 50 seconds. However, the required time increased significantly for large N_B, with $N_B = 160$ requiring 186 minutes.

Since the system identification process is parameter specific, obtaining the A-matrix for a different set of parameters requires repetition of the entire process. As will be shown in our proposed approach, this often can be avoided if analytical insights are used to obtain the A-matrix. It should also be noted that to obtain a good estimate using the SI-based approach the sampling process needs to be exhaustive. Depending on the initial conditions, some bins may have a low number of samples and therefore the transition rate is not truly indicative of behavior. The identified A-matrix will perform poorly when it is applied to a population with a different initial condition.

Analytical approach: The A-matrix can also be built analytically by considering evolving probabilities. Koch et al. (2011) proposed a technique which allowed heterogeneity in the C parameter but ignored noise. In this case, the only uncertain parameter in (1) is a_i. Under these assumptions, the thermal equation (1) can be used to establish transition probabilities.

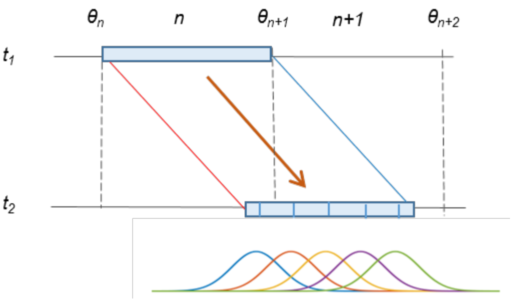

Fig. 2. Top figure: TCLs propagating over a time-step; Bottom figure: Approximate probability distribution of TCLs at the end of a time-step.

To summarize, the probability of TCLs evolving from a starting temperature θ_{start} to a different temperature θ_{end} in one time step is,

$$P(\theta_{end}|\theta_{start}) = P(a_i), \tag{6}$$

where a_i can be obtained from (1) as,

$$a_i = \frac{\theta_a - \theta_{end} - m_t\theta_g}{\theta_a - \theta_{start} - m_t\theta_g}. \tag{7}$$

With known bin boundaries, the fraction of TCLs lying in a specific receiving bin (indexed with j), given they all started at θ_{start} can be expressed as,

$$P(\theta_j \le \theta_{end} < \theta_{j+1}|\theta_{start}) = \int_{a_1}^{a_2} p(a)\mathrm{d}a \tag{8}$$

where $p(\cdot)$ is the probability density function associated with parameter a, θ_j is the lower temperature boundary of bin j, θ_{j+1} the upper end, a_1 and a_2 follow from (7) with θ_j and θ_{j+1} replacing θ_{end}, respectively. Repeating this process for all sending and receiving bins gives the desired A-matrix (see Koch et al. (2011)).

This technique cannot take into account noise $\mathbf{w}$ and heterogeneity in other important parameters such as P and R (see Mathieu et al. (2013)). The next section presents an analytical technique that incorporates noise and can also be extended to parameter heterogeneity. The intuition obtained from this approach will be applied in Section 4 to study bin model fidelity, the effects of bin widths, and to provide a comparison of the influence of noise and heterogeneity.

3. ANALYTICAL MODELING WITH NOISE AND HETEROGENEITY

3.1 Incorporating noise

Initially, consider a homogeneous TCL population. Figure 2 shows two specific OFF bins, indexed n and $n+1$, and their contents at two subsequent time period t_1 and t_2 (where $t_2 = t_1 + h$). Assume all TCLs lie *uniformly* inside bin n at time t_1, where the uniform assumption inside any bin is consistent with Koch et al. (2011), Mathieu et al. (2013). Then, at time t_2, governed by (7), TCLs will transition to a range of bins. If no noise is present, the analytical approach described in the previous section can be used to obtain the A-matrix. But with a non-zero noise process, $\mathbf{w} \sim N(0, \sigma_w^2)$, the previous method cannot be applied.

With noise, intuitively, the range of temperature reached at time t_2 should be wider. Depending on the noise variance, TCLs may reach the bins in different proportions to the noise-free and may also reach neighboring bins not reached previously. Thus, there is a need to compute the fractions that evolve to various bins given that they all start from the same bin. At time t_1, consider a specific sending bin n. The probability of lying inside the receiving bin j, with boundaries θ_j and $\theta_{(j+1)}$, is,

$$\begin{aligned} P(\theta_j \leq \theta_{end} \leq& \theta_{j+1} | \theta_n \leq \theta_{start} < \theta_{n+1}) \\ &= 1 - \Big(P(\theta_{end} \leq \theta_j | \theta_n \leq \theta_{start} < \theta_{n+1}) \\ &\quad + P(\theta_{end} > \theta_{j+1} | \theta_n \leq \theta_{start} < \theta_{n+1}) \Big). \end{aligned} \tag{9}$$

Instead of a continuous uniform distribution over the n-th bin, consider the bin population concentrated at M discrete values $\theta_{nl}, l = 1,...,M$ that are evenly spaced across the width of the bin, with $\theta_{n1} = \theta_n$ and $\theta_{nM} = \theta_{n+1}$. Hence, TCLs inside bin n are equally distributed to each of these M discrete pulses. This process can be thought of as subdividing the n-th bin into M sub-bins. Inside each sub-bin, consider a TCL population with the same parameters as considered for the bin.

Each of these pulses propagates forward in time accordin to (7). With $\mathbf{w} = 0$, the pulse at θ_{nl,t_1} will move to $\bar{\theta}_{nl,t_2}$, which may or may not lie in the n-th bin. With noise, $\mathbf{w} \sim N(0, \sigma_w^2)$, there will be a Gaussian distribution around each center, $\bar{\theta}_{nl,t_2}, l = 1,...,M$. This gives a Gaussian mixture distribution, i.e. a mixture of $N(\bar{\theta}_{nl,t_2}, \sigma_w^2), l = 1,...,M$.

The contribution to each of the recceiving bins must be computed. This can be achieved by computing tail probabilities, according to $Q(\cdot) = 1 - \Phi(\cdot)$, where $\Phi(\cdot)$ is the Gaussian cumulative distribution function (CDF).

Starting from bin n, the probability of TCLs falling above the upper boundary of receiving bin j at time t_2 can be computed by summing over the tail probabilities of Gaussian distributions centered at $\bar{\theta}_{nl,t_2}, l = 1,...,M$ with variance σ_w^2,

$$P(\theta_{end} > \theta_{j+1} | \theta_n \leq \theta_{start} < \theta_{n+1}) = \frac{1}{M} \sum_{l=1}^{M} Q\Big(\frac{\theta_{j+1} - \bar{\theta}_{nl,t_2}}{\sigma_w} \Big). \tag{10}$$

Similarly, the probability of the tail falling below the lower boundary is given by,

$$\begin{aligned} P(\theta_{end} \leq& \theta_j | \theta_n \leq \theta_{start} < \theta_{n+1}) \\ &= 1 - P(\theta > \theta_j | \theta_n \leq \theta_{start} < \theta_{n+1}) \\ &= 1 - \frac{1}{M} \sum_{l=1}^{M} Q\Big(\frac{\theta_j - \bar{\theta}_{nl,t_2}}{\sigma_w} \Big). \end{aligned} \tag{11}$$

This same process is repeated for every receiving bin, giving the entries of the transition matrix for sending bin n. The overall algorithm can be summarized as:

Algorithm I
For all sending bins $n = 1,...,2N_B$:

(i) For the specific sending bin n, divide the bin population evenly among M discrete temperature values at $\theta_{nl,t_1}, l = 1,...,M$.
(ii) Using (7) and $\mathbf{w} = 0$, find the corresponding time-propagated values $\bar{\theta}_{nl,t_2}, l = 1,...,M$.
(iii) With $\mathbf{w} \sim N(0, \sigma_w^2)$, establish M normal distributions $N(\bar{\theta}_{nl,t_2}, \sigma_w^2), l = 1,...,M$.
(iv) Given the pre-specified temperature boundaries for receiving bin j, use (9)-(11) to compute the fractions of TCLs that propagate to each bin $j = 1,....2N_B$.

Normalize all the propagated values to give the desired A-matrix. □

When noise is high, it may cause transitions from n to $j < n$, i.e. in the opposite direction to normal. This must be handled carefully at the switching thresholds, bins 1 and $N_B + 1$. TCLs falling below j=1 when starting from any off bin remain off. Since we do not consider bins outside of the dead-band, this fraction of TCLs is small and can be captured in bin $j = 1$. Similarly, TCLs starting in bin $N_B + 1$ and that experience a temperature rise remain in that bin.

If noise is relatively low and the number of bins N_B is small (i.e. the bins are wide), we have observed that noise has limited effect on the A-matrix coefficients. This suggests that given a specific bin width, the A-matrix coefficients are invariant to noise up to a certain level. This property is analyzed in Section 4 by varying the noise levels and the numbers of bins (i.e. the bin widths).

3.2 Incorporating parameter heterogeneity

Algorithm I can be modified to consider parameter heterogeneity rather than noise. Consider heterogeneity only in P, and assume P is distributed with $N(\mu_P, \sigma_P^2)$. Let σ_P can be expressed as a fraction σ_r of the mean value of P,

$$\sigma_P = \sigma_r \mu_P. \tag{12}$$

Starting from temperature θ_t, a TCL's cooling and heating rates, $\alpha_{ON,t}$ and $\alpha_{OFF,t}$, are given by the expressions (see Callaway (2009), Bashash and Fathy (2011)):

$$\alpha_{ON,t} = \frac{\theta_t - \theta_A + PR}{CR} \tag{13}$$

$$\alpha_{OFF,t} = \frac{\theta_t - \theta_A}{CR}. \tag{14}$$

Given $P \sim N(\mu_P, \sigma_p^2)$, the expected value of $\alpha_{ON,t}$ can be computed using (13),

$$\begin{aligned} \mathrm{E}\big[\alpha_{ON,t}\big] &= \mathrm{E}\Big[\frac{\theta_t - \theta_a + PR}{CR} \Big] \\ &= \mathrm{E}\Big[\frac{\theta_t - \theta_a}{CR} + \frac{P}{C} \Big] \\ &= \frac{\theta_t - \theta_a}{CR} + \frac{\mu_P}{C} \end{aligned} \tag{15}$$

and the variance of $\alpha_{ON,t}$ is,

$$\mathrm{Var}(\alpha_{ON,t}) = \mathrm{Var}\Big(\frac{\theta_t - \theta_A + PR}{CR} \Big) = \frac{\sigma_P^2}{C^2}. \tag{16}$$

Hence, α_{ON} is distributed as $N\big(\frac{(\theta_t - \theta_a)}{CR} + \frac{\mu_P}{C}, \frac{\sigma_P^2}{C^2}\big)$. From (14), $\alpha_{OFF,t}$ is independent of P.

Starting from a sending bin n at time t_1, with $\theta_{nl,t_1}, l = 1,...,M$, use (1) with μ_P to obtain $\bar{\theta}_{nl,t_2}$. That again results in a mixture of Gaussian distributions centered at $\bar{\theta}_{nl,t_2}, l = 1,...,M$ and each with variance σ_P^2/C^2.

Following the same procedure for computing the tail probabilities, using (9)-(11) but replacing σ_w^2 by σ_P^2/C^2, the fractions inside each receiving bin can be computed. Repeating the procedure for all sending bins and normalizing gives the transition matrix. Due to α_{OFF} being independent of P, the OFF-bin transition probabilities are identical to the homogeneous case. Hence, transitions through ON bins have variance σ_P^2/C^2 and through OFF bins have zero variance.

Algorithm II
Replace steps (iii)-(iv) in **Algorithm I** with:

(iii) With $P \sim N(\mu_P, \sigma_p^2)$, establish M normal distributions:
(a) $\mathrm{N}(\bar{\theta}_{nl,t_2}, \sigma_P^2/C^2), l = 1,...,M$ if $\bar{\theta}_{nl,t_2} \in j = N_B+1,...,2N_B$.
(b) $\mathrm{N}(\bar{\theta}_{nl,t_2}, 0), l = 1,...,M$ if $\bar{\theta}_{nl,t_2} \in j = 1,...,N_B$.
(iv) With the variance terms in step (iii) for receiving bin j, use (9)-(11) to compute the fractions of TCLs that propagate to each bin $j = 1,....2N_B$.

While with high noise, it is possible for TCLs to move to a lower indexed bin while remaining ON (or OFF), without noise, transitions to lower indexed bins while remaining ON (or OFF) should not occur. Otherwise, transitions would contradict the thermal dynamics in (1). Hence, the level of heterogeneity that is acceptable in P has a limit which will be discussed in Section 4.4.

Algorithm II allows for heterogeneity in P. Heterogeneity in other parameters, such as θ_a, C, R can be incorporated following a similar approach. Depending on where each parameter appears in (13) and (14), slight modifications may be necessary. Additionally, if parameters appear in the denominator, dealing with log-normal or uniform distributions may be easier since their reciprocal distributions have closed form expressions.

4. ANALYSIS AND RESULTS

4.1 Model performance

To perform the experiment on TCLs, let us consider parameter values according to Table 1. Let, σ_w = 0.005. We obtained the analytical Matrix for $N_B = 160$ using Algorithm-I. Next, we simulated 10,000 TCLs using 1 over 24 hours and obtained the identified A-matrix. Aggregate power outputs using the analytical matrix and the identified matrix against the output using the full simulation are shown in Fig. 3. We see that the aggregate power output matched exactly with the identified and the analytical matrix and also agrees closely with the reference output from the full simulation. The analytical matrix was obtained in *2.5 seconds*, whereas the analytical one in 145 minutes. Thus, clearly demonstrating the significant computational benefit. Next, consider a heterogeneous population of

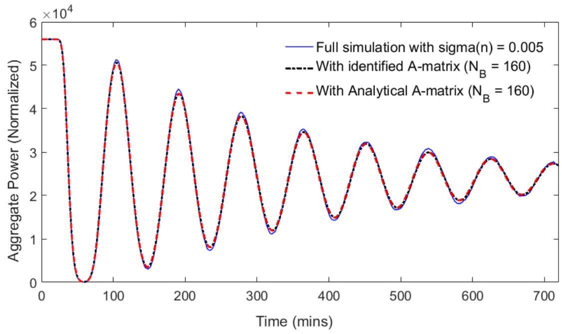

Fig. 3. Simulation of TCLs with identified, analytical and full model with noise.

10,000 TCLs. Assume P is distributed with $N(\mu_P, \sigma_P^2)$, where $\sigma_P = \mu_P \sigma_r$ with μ_P = 14 kW, $\sigma_r = 0.02$. We computed the analytical matrix using Algorithm-II. Similarly as above, the identified matrix was also obtained. Aggregate power output using the analytical matrix and the identified matrix as well as

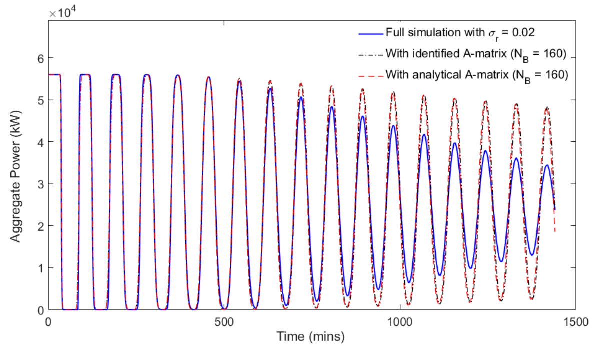

Fig. 4. Simulation of TCLs with identified, analytical and full model with heterogeneity in P.

with full simulation of 10000 heterogeneous TCLs are shown in Fig. 4. Again, we see that the aggregate power outputs from the identified and the analytical matrix matched exactly. However, when compared against the output from full simulation, we see that results match closely up to 620 minutes, and then output from full simulation starts to damp at a faster rate than obtained by the bin model. This deviation is mainly due to bin model's limitations to deal with heterogeneity, as will be analyzed in Section 4.4. Besides noise and heterogeneity, bin model is also prone to model error due to temperature discretization, which causes damping of oscillations. Hence, before analyzing the influence of noise and heterogeneity, in the next section, we briefly highlight some important features of the bin model and discuss model error.

4.2 Influence of model error

Without any noise and heterogeneity, when a homogeneous TCL population is simulated using (1), the power output shows undamped oscillations (the rectangular pulse with N_B=640 in Fig. 5). Simulations with different N_B were carried out. The responses are shown in Fig. 5. We see that for N_B= 320 and 640, the response is accurately a rectangular pulse, whereas with N_B less than 160, there is considerable damping.

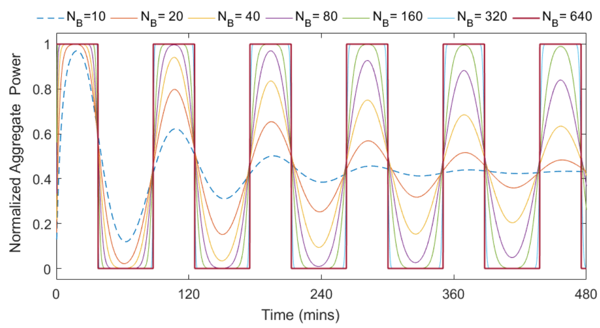

Fig. 5. Simulation of homogeneous TCLs with varying N_B.

The primary sources contributing to model error are typically in discrezation of time and temperature space. To avoid error due to time discretization, (1) is solved using an exact model. However, the bin model is prone to temperature discretization. Since bin model is based on probabilistic transitions, using a low number of bins cannot accurately capture the aggregate dynamics. If a homogeneous group of TCLs start at same temperature $\theta_{start} = \theta_{max}$ (i.e. bin N_B) at t_1, with all in off states, one time step later, by (4) states evolve. Using the A-matrix coefficients (5), p_{N_B,N_B} fraction will stays at bin N_B, whereas $(1 - p_{N_B,N_B})$ fraction moves to bin other bins. Thus, starting from a single bin, the process of filling the bins is similar to a geometric RV. Since we consider N_B bins over the dead-band range, if all TCLs start at $n = N_B$ (OFF bin), within additional N_B time steps, all ON bins will have received some fraction of TCLs. Subsequently OFF bins start to get filled which causes damping in power. Thus, if N_B is low, the damping will appear sooner. Therefore, in Fig. 5, we that using lower number of bins result in considerable damping.

Additionally, bin model performance also varies with temperature initial conditions. For example, if TCL temperatures uniformly distributed over the dead-band range, the power output is expected to have smoother transitions. Readers may refer to Bashash and Fathy (2011) where authors compared the performance of bin model similarly by varying N_B. Since initial conditions were dispersed, the resulting output power was triangular shaped. With 200 bins outputs matched closely, whereas for a rectangular shaped pulse more bins are necessary. Thus, if temperatures are synchronized, it is advisable to use models with larger number of bins to reduce the output error, again demonstrating the need for fast computational tools to compute such higher order models.

4.3 Comparing the influence of noise and heterogeneity on damping

To analyze the influence of noise on damping, noise levels were varied from $\sigma_w = 0$ to 0.01 with 0.002 step increments and 10,000 TCLs were simulated using (1). Results are shown in Fig. 6. To compare the decay of oscillations under the influence of noise against decay due to model error (shown in Fig. 5), we can use log decrement as a metric to see how fast oscillations are decaying. Using the first peak and a subsequent peak after n periods, the log decrement is defined as,

$$\delta_{decay} = \frac{1}{n} \ln \frac{x_1}{x_n} \tag{17}$$

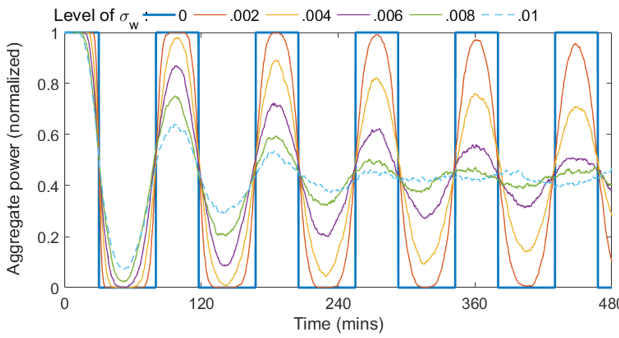

Fig. 6. Aggregate power response with varying σ_w.

In our case, the first and the third peaks were chosen. We compute δ_{decay}'s for both cases (Fig. 5 and Fig. 6). In Fig. 7, the influence of model error and noise on decay rates are shown. The x-axis values, representing noise standard deviation and the number of bins, are normalized so that we can visually compare their influence on the same plot. The y-axis represents the log decrements (i.e. damping). We see that when below $\sigma \leq 0.0025$ and $N_B \leq 80$ model error is more dominant. Any model of less than 80 bins decays faster than the influence of having noise

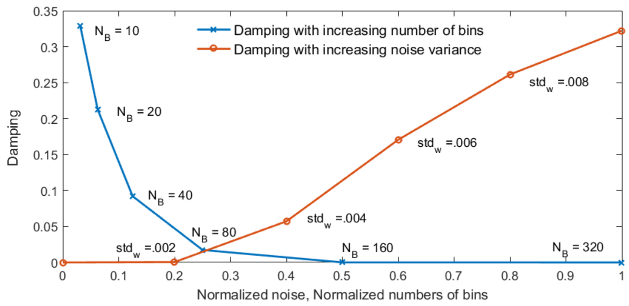

Fig. 7. Damping due to varying N_B, σ_w.

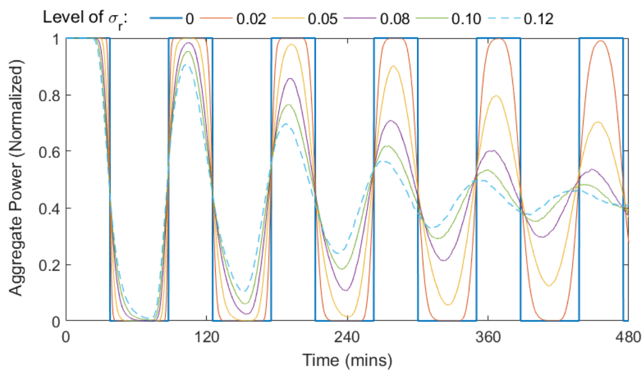

Fig. 8. Aggregate power response with varying σ_r.

of $\sigma <= 0.0025$. As noise level rises above 0.0025, noise has increasingly higher impact on damping. This also implies that if we arbitrarily choose $N_B = 40$ for simulating a population with known $\sigma_w = 0.003$, the resulting damping in power would be due to model error, hence, a larger order model should be used to capture the influence of noise.

Similar to above, Fig. 8 shows the effect of heterogeneity in P by varying σ_r and Fig. 9 compares relative damping due to σ_r and N_B. Notice that the oscillations due to heterogeneity are asymmetric compared to the ones observed under noise and in homogeneous case. Additionally, the period of oscillation varies as we vary σ_r. Fig.9 suggests that with $\sigma \leq 0.0025$ and $N_B \leq 80$, the observed damping of oscillation should be attributed to model error.

4.4 Noise invariance in bin model

In this section, we demonstrate that due to the fact that bins have non-zero width, bins are inherently tolerant to some noise

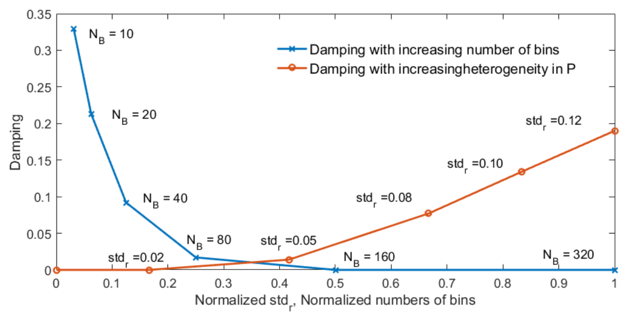

Fig. 9. Damping due to varying N_B, σ_r.

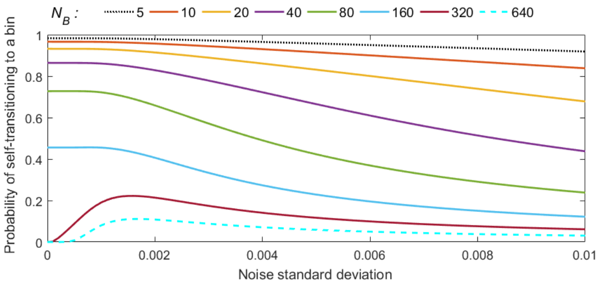

Fig. 10. Probability of TCLs self-transitioning to a bin.

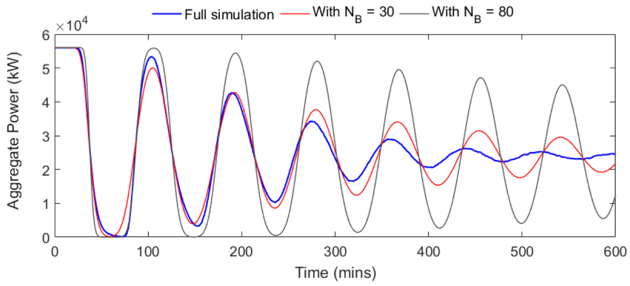

Fig. 11. Aggregate power with full simulation of TCLs when $\sigma_r = 0.1$ and with different order A-matrices.

level. In Fig. 10, the probability of self-transition to a bin (i.e. one diagonal element of an A-matrix) is shown for varying N_B, and σ_w. In our case, self-transition for bin j=3 was chosen.

Fig. 10 suggests that for a particular N_B, where curve is flat, the diagonal elements of A-matrix are invariant to noise change. For example, when N_B=5, 10, 20 and for $0 < \sigma_w < 0.001$, the change in A-matrix diagonal elements is infinitesimally small. Moreover, with $N_B = 5, 10, 20$, diagonal elements are dominant. Hence, we can assume that the effect is also negligible on off-diagonal bins.

One the other hand, for models with very large number of bins, the diagonal elements are sensitive to low noise. This is because, without noise self-transition does not happen as bins are very narrow. As noise level rises, the left tail of the normal distribution (see Fig. 2) enters the bin, causing self-transition. Also since diagonal elements are not dominant for large N_B, remarks on invariance cannot be readily made.

Following the same procedure shown for noise we can also find tolerance to heterogeneity given the width of the bin. Note that for TCLs starting at the same initial temperature, their temperature distribution after one time step has a variance of σ_w^2 under the noise process, whereas from (16), variance is σ_P^2/C^2 (only for ON-bins) under heterogeneity in P.

Assuming $\sigma_r = 0.1$ and $\mu_P = 14$, standard deviation in temperature is $\sigma_P \mu_P / C = 0.00038$ (comparable to $\sigma_w = 0.00038$). From analysis of Fig. 10, we expect such low value of standard deviation will have negligible impact on A-matrix coefficients (with N_B less than 320). Such A-matrices are thus invariant to $\sigma_r = 0.1$.

Additionally, note that to preserve the natural thermal dynamics in (1), heterogeneity in parameters should not cause shifts from higher indexed ON bins to lower indexed ON bins. This suggests (13) should be larger than 0. Then, using three sigma of standard deviation of μ_P and values from Table 1, we find $\sigma_r \geq 0.1905$. Again, using (16), the standard deviation in temperature is then 0.00074. From Fig. 10 we see again that even at this level of uncertainty, A-matrices would still be almost invariant to heterogeneity in P.

Note that the remarks about bin model performance apply to both identified and analytical matrices whose responses match closely in all cases. Our analysis of heterogeneity, at least in P, suggests that in literature while heterogeneity is often included in bin models, one should be aware of the limitations. Fig. 11 shows that with $\sigma_r = 0.1$, the output using N_B=30, instead of N_B=80 is more similar to results obtained by full simulation of TCLs using (1). This is because damping due to model error with $N_B = 30$ closely matches damping due to $\sigma_r = 0.1$.

The bin model treats noise and heterogeneity almost in a similar way. The only notable differences are that for heterogeneity in P, non-zero variance only occurs when transitioning to ON bins, and variance is scaled compared to noise variance. However, we argue that the higher error in dealing with heterogeneity can be attributed fundamental differences in uncertainty due to noise process vs. heterogeneity. Noise $\mathbf{w}$ is a truly random process, hence the random transitions can be captured well using a Markov transition matrix (equivalently by A-matrix). The uncertainty in heterogeneity, on the other hand, appears only at the initial step when using the full model (1). But with bin model it appears at each step since at end of each transition, bin model assumes TCLs mix uniformly inside the bin. Additionally, we saw from our above analysis that $\sigma_r \geq 0.19$ for valid thermal dynamics in (1). This results in limited spreading of temperature (standard deviation of less than 0.001), thus having limited impacts of the coefficients of A-matrix, irrespective of obtained analytically or by SI approach.

5. CONCLUSIONS

While there has been considerable work in modeling and control of TCLs using aggregate models, to our best knowledge, there has been limited work that focus on identifying the various sources of error when using bin based models, especially by exploiting analytical based techniques. Hence, here we incorporate noise and parameter heterogeneity analytically in the A-matrix governs the aggregate TCL dynamics, which is a unique contribution compared to existing literature. We show that matrices obtained under computationally intensive SI approach can be obtained within seconds using our analytical approach. Using insights gained through the analytical modeling, we then study the influence of noise and parameter heterogeneity on bin models. We showed how bins are invariant to varying levels of noise. We compared damping of power oscillations due to model error, noise and heterogeneity. By comparing the influence of different sources, we showed how an appropriate number of bins should be chosen. Another implication of our analysis is that if parameters change one can often avoid re-identifying the bin model, thus saving valuable computational resources, knowing that the set of new parameters would have negligible impact on the existing model. We also pointed out bin models limitations in healing with cases of heterogeneity, irrespective of if A-matrix is obtained analytically or by system identification.

Future work could include analyzing the impact of non-Gaussian distributions of noise or parameter heterogeneity. We are also interested in using the analytical tools developed in this paper to predict, and thus, hopefully avoid load synchronization and oscillations.

REFERENCES

Bashash, S. and Fathy, H.K. (2011). Modeling and control insights into demand-side energy management through setpoint control of thermostatic loads. *Proceedings of the 2011 American Control Conference*, 4546–4553.

Bashash, S. and Fathy, H.K. (2013). Modeling and control of aggregate air conditioning loads for robust renewable power management. *IEEE Transactions on Control Systems Technology*, 21(4), 1318–1327.

Callaway, D.S. (2009). Tapping the energy storage potential in electric loads to deliver load following and regulation, with application to wind energy. *Energy Conversion and Management*, 50(5), 1389–1400. doi: 10.1016/j.enconman.2008.12.012.

Callaway, D.S. and Hiskens, I.A. (2011). Achieving Controllability of Electric Loads. *Proceedings of the IEEE*, 99(1).

Chong, C. and Debs, A. (1979). Statistical synthesis of power system functional load models. *IEEE Conference on Decision and Control*, 18, 264–269.

Crocker, S.J. and Mathieu, J.L. (2016). Adaptive State Estimation and Control of Thermostatic Loads for Real-Time Energy Balancing. 3557–3563.

Esmaeil Zadeh Soudjani, S. and Abate, A. (2015). Aggregation and Control of Populations of Thermostatically Controlled Loads by Formal Abstractions. *IEEE Transactions on Control Systems Technology*, 23(3), 975–990.

Ghaffari, A., Moura, S., and Krstic, M. (2014). Analytics Modeling and Integral Control of Heterogeneous Thermostatically Controlled Load Populations. *Proceedings of ASME 2014 Dynamic Systems and Control Conference (DSCC 2014)*.

Ihara, S. and Schweppe, F.C. (1981). Physically based modeling of cold load pickup. *IEEE Transactions on Power Apparatus and Systems*, PAS-100(9), 4142–4150.

Koch, S., Mathieu, J.L., and Callaway, D.S. (2011). Modeling and Control of Aggregated Heterogeneous Thermostatically Controlled Loads for Ancillary Services. *Proceedings of the 17th Power Systems Computation Conference*.

Kundu, S. and Sinitsyn, N. (2012). Safe Protocol for Controlling Power Consumption by a Heterogeneous Population of Loads. *2012 American Control Conference*, (July), 2947–2952.

Kundu, S., Sinitsyn, N., Backhaus, S., and Hiskens, I. (2011). Modeling and Control of Thermostatically Controlled Loads. In *Power Systems Computation Conference*. IEEE, Stockholm.

Malhame, R. and Chong, C.Y. (1985). Electric Load Model Synthesis by Diffusion Approximation of a High-Order Hybrid-State Stochastic System. *IEEE TRANSACTIONS ON AUTOMATIC CONTROL*, (9).

Mathieu, J.L., Koch, S., and Callaway, D.S. (2013). State estimation and control of electric loads to manage real-time energy imbalance. *IEEE Transactions on Power Systems*, 28(1), 430–440.

Mortensen, R.E. and Haggerty, K.P. (1988). Stochastic computer model for heating and cooling loads. *IEEE Transactions on Power Systems*, 3(3), 1213–1219.

Nazir, M.S., Galiana, F.D., and Prieur, A. (2016). Unit Commitment Incorporating Histogram Control of Electric Loads With Energy Storage. *IEEE Transactions on Power Systems*, 31(4), 2857–2866.

Perfumo, C., Braslavsky, J., Ward, J.K., and Kofman, E. (2013). An analytical characterisation of cold-load pickup oscillations in thermostatically controlled loads. *2013 3rd Australian Control Conference, AUCC 2013*, 195–200.

Perfumo, C.N. (2013). Dynamic modelling and control of heterogeneous populations of thermostatically controlled loads.