Modeling Thermostatically Controlled Loads to Engage Households in the Smart Grid: Lessons Learned from Residential Refrigeration Units

E. C. Kara[1], M. Bergés[2] and G. Hug[3]

[1] PhD Student, Civil and Environmental Engineering, Carnegie Mellon University, 5000 Forbes Ave., Pittsburgh, PA 15213; e-mail: eckara@cmu.edu
[2]Assistant Professor, Civil and Environmental Engineering, Carnegie Mellon University, 5000 Forbes Ave., Pittsburgh, PA 15213; e-mail: marioberges@cmu.edu
[3]Assistant Professor, Electrical and Computer Engineering, Carnegie Mellon University, 5000 Forbes Ave., Pittsburgh, PA 15213; e-mail: ghug@ece.cmu.edu

ABSTRACT

As renewable generation capacity in the power grid increases, keeping the balance between the supply and demand becomes difficult. This threatens the grid's stability and security. Existing power reserve assets and regulation methodologies fail to provide the short-term responses required to keep the load and generation balanced as the amount of renewable generation increases. Hence, researchers proposed to increase the information exchange within the power network and to introduce real-time demand control to ensure robustness while accommodating the intermittent nature of these generation resources. Constituting a significant portion of the electrical demand of buildings, thermostatically controlled loads (TCLs) are well-suited to provide real-time demand control. In this paper, we shed light on challenges associated with engaging TCLs to the power grid using a centralized control strategy. We focus on the challenges associated with simulating a realistic TCL population using the models that are proposed in the literature. Specifically, we use data collected from residential refrigeration units operating in 214 different households to propose a strategy to select parameters when simulating a TCL population.

INTRODUCTION

With the increasing presence of renewable energy resources in the power grid, additional reserves are needed to remedy the supply and demand imbalance caused by the intermittent nature of renewable power (Gellings et al. 2004). Traditionally, fossil fuel power plants are idled to provide additional capacity and maintain reliable operation of the power system when necessary (i.e. ancillary services). Therefore, higher penetration of renewable generation sources can cause carbon emissions to increase if the additional capacity is provided via traditional generation resources (Faruqi et al. 2009; Gellings et al. 2004). As an alternative to traditional generation dispatch, Demand Response (DR) mechanisms are garnering interest in the research community due to their ability to decrease or eliminate the need for additional capacity by managing the demand in the power grid (Callaway 2011; Faruqi et al.

2009; Kiliccote et al. 2011). One way to provide DR mechanisms is to leverage the thermal energy storage capacity of TCLs in buildings via direct load control (DLC). Previous researchers revealed the motivation behind using DLC mechanisms to engage an aggregate population of TCLs, and showed that such mechanisms are capable of providing services necessary to maintain reliable operation of the power system in the presence of high renewable energy penetration (Callaway et al. 2011a; Mathieu et al. 2012; Kara et al. 2012; Kara et al. 2013a).

Despite their promise, existing DLC mechanisms using TCLs include strict assumptions in their problem set-up. Specifically, assumptions are made when simulating a population of TCLs using individual load models. Often, the disturbances to individual TCL operations, such as user interactions and changes in ambient temperature, are either ignored or introduced as white noise on the temperature dynamics. The thermal parameters are assumed to be uniformly and/or normally distributed. Furthermore these parameters are assumed to stay constant during the course of load simulation (Callaway et al. 2011a; Mathieu et al. 2012; Kara et al. 2012; Kara et al. 2013b).

Recent work proposed more dynamic assumptions to address these problems. For example, different thermal resistance values for different load interaction scenarios are introduced by Kamgarpour et al. (2012). Specifically, the authors argue that thermal resistance depends on user interaction with the loads—for example, refrigerator door openings—and suggest using two different values when modeling it. The alternation in between these two values is assumed to follow a homogeneous Poisson process. As suggested by Kamgarpour et al. (2012) an improvement in the modeling approach is required to realistically model an aggregate TCL population in a way that considers time-dependent disturbances to their operation.

As addressed above, research on DLC made assumptions that failed to capture the dynamics of TCLs under operation. In this paper, we investigate the power consumption of refrigeration units available in residential households to better understand the dynamics of the populations of these loads. In particular, we leverage a dataset from the Household Energy Survey of the UK (Zimmermann et al. 2012) to propose a modeling scheme that represents the working conditions of TCLs.

THE DATA

The Household Energy Survey of the UK monitored the energy consumption and electrical power demand of 251 households over the period of May 2010 to July 2011. It consists of multiple datasets with different numbers of households, data collection durations, and sampling rates (Zimmermann et al. 2012). In this study, we use the energy demand measurements collected every 2 minutes from refrigeration units over the course of a month within the May 2010 to July 2011 period. We use data obtained from 373 different refrigeration units present in 214 different households. Table 1 shows the characteristics of this dataset along with the types and quantities of refrigeration units. The electricity demand is measured in deciwatt hours, and to estimate the rated power of each refrigeration unit, we assumed that the rated power of a unit remains constant between samples.

Table 1. Sampling period and data collection duration of different refrigeration units

Sampling Period	Collection Duration	Refrigeration Unit	Total Number
2 minutes	~27 days	Freezer (Upright)	89
	~28 days	Chest Freezer	39
	~28 days	Refrigerator	101
	~27 days	Fridge + Freezer	141
	~28 days	Beer/Wine Chiller	3

INDIVIDUAL TCL MODEL

To motivate the use of real-world data for parameter estimation, we introduce a two-state TCL modeling scheme used in the literature. In the interest of simplicity, we ignore the noise terms (Callaway 2009; Callaway et al. 2011b). In the following equation, $\theta_{i,t}$, $\theta_{a,i,t}$ and $\theta_{i,t+T_i}$ denote the current temperature of the TCL, the ambient temperature at time t, and the temperature of the TCL after a period of time T_i, respectively.

$$\theta_{i,t+T_i} = a_i\theta_{i,t} + (1 - a_i)(\theta_{a,i,t} - m_{i,t}\theta_{g,i}) \qquad (1)$$

The temperature gain, $\theta_{g,i}$, depends on the appliance type, the resistance, R_i and the rated power, $P_{rated,i}$ of the appliance: for heating devices $\theta_{g,i} = -R_iP_{rated,i}$ and for cooling devices $\theta_{g,i} = R_iP_{rated,i}$. The parameter a_i governs the thermal characteristics of each TCL and is an exponentially decaying function defined as $a_i = e^{-T_i/C_iR_i}$ where, C_i denotes the thermal capacitance and $m_{i,t}$ is a binary variable representing the state of each appliance i at time t.

In a cooling scenario, we replace the $\theta_{i,t+T_i}$ in Equation (1) with the lower thermostatic dead-band, which is calculated using the temperature set point $\theta_{i,set}$ and the thermostatic dead-band width δ_i as $\theta_{i,set} - \delta_i/2$. Assuming ambient temperature remains constant until $t + T_i$ and rearranging Equation (1) gives the time the appliance takes to reach the lower thermostatic dead-band (i.e. $T_{ON,i}$). The equation for $T_{OFF,i}$ is also written similarly.

$$T_{ON,i} = -R_iC_i\ln\left(\frac{\theta_{i,set} - \delta_i/2, -\theta_{a,i,t} + R_iP_{rated,i}}{\theta_{i,set} + \delta_i/2, -\theta_{a,i,t} + R_iP_{rated,i}}\right) \qquad (2)$$

$$T_{OFF,i} = -R_iC_i\ln\left(\frac{\theta_{i,set} + \delta_i/2, -\theta_{a,i,t}}{\theta_{i,set} - \delta_i/2, -\theta_{a,i,t}}\right) \qquad (3)$$

Rewriting R_iC_i in terms of $T_{OFF,i}$ using Equation (3) and rearranging Equation (2) we get:

$$R_i = \frac{(\theta_{a,i,t} - \theta_{i,set})(K_i - 1) - \delta_i/2(K_i + 1)}{P_{rated,i}(K_i - 1)} \qquad (4)$$

where, K_i is defined as follows:

$$K_i = \left(\frac{\theta_{i,set} + \delta_i/2 - \theta_{a,i,t}}{\theta_{i,set} - \delta_i/2 - \theta_{a,i,t}}\right)^{\frac{T_{ON}}{T_{OFF}}} \quad (5)$$

Hence, Equations (4) and (5) imply that if the T_{ON}, T_{OFF}, $\theta_{a,i,t}$, δ_i, $\theta_{i,set}$, m_i and $P_{rated,i}$ parameters are known ahead of time, then R_i can be mathematically obtained under certain conditions. Following this, C_i can be obtained using Equation (3).

If necessary, these parameters can then be used to create realistic distributions of TCLs for simulation purposes. By doing so, the estimated distributions of R_i and C_i will absorb the variability in the ambient temperature and similar time dependent disturbances are reflected in these estimated parameters.

DATA ANALYSIS AND TCL MODELING STRATEGY

In order to obtain T_{ON} and T_{OFF} values using the data collected from individual refrigeration units, we first run an event detection algorithm to infer the state transitions of the units from power measurements. The algorithm compares the change in the actual power consumption of each appliance $P_{i,t}$, which is defined as $\Delta P_{i,t} = P_{i,t+1} - P_{i,t}$, with a power threshold $\xi > 0$. If $\Delta P_{i,t} > 0$ and $\Delta P_{i,t} > \xi$, it assumes that the refrigeration unit is turned ON, and if $\Delta P_{i,t} < 0$ and $\Delta P_{i,t} < -\xi$, it assumes that the unit is turned OFF. If two consecutive ON or OFF states are detected, the algorithm selects the last state transition and calculates T_{ON} and T_{OFF} accordingly. Due to concerns regarding data collected from some refrigeration units and appliance labels assigned in the dataset, any T_{ON} and T_{OFF} value over four hours are removed from the dataset.

For modeling the dynamics of a refrigeration unit, we propose that $T_{ON,i}$ and $T_{OFF,i}$ follow a two parameter Weibull distribution for each appliance i. To make sure that the fitted Weibull distributions are obtained using a reasonable number of samples, we define the parameter N_{DC}, which is the minimum number of duty cycles required to label the collected data as valid. In this paper, we use $N_{DC} = 100$.

The Weibull distribution is commonly used in industrial engineering to represent manufacturing, lead and delivery times (Trivedi et al. 1982 and Kelle et al. 1990), and its probability density function, f, is given in terms of the scale parameter, λ and the shape parameter, k as follows.

$$f(x;\, \lambda, k) = \begin{cases} \frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1} e^{-(x/\lambda)^k} & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (6)$$

To show that our proposition is acceptable, we generated a Weibull plot from T_{ON} and T_{OFF} data collected from two refrigeration units using a power threshold, ξ of 40 Watts. Figure 1 shows these plots. The x-axis value is the natural logarithm of each data point T_{ON} or T_{OFF}. The y-axis in this case is the $ln\left(-ln\left(1 - \hat{F}(x)\right)\right)$, where $\hat{F}(x)$ is the empirical cumulative distribution. Using these axes results in a quantile-quantile (Q-Q) type of plot, in which the linearity of the scattered data points indicates that the Weibull distribution assumption is reasonable (Nelson 1982).

So far we suggested that T_{ON} and T_{OFF} parameters for each refrigeration unit follow a Weibull distribution. Assuming $T_{ON,i}$ and $T_{OFF,i}$ are independent parameters, given $k_{ON,i}$, $k_{OFF,i}$, $\lambda_{OFF,i}$ and $\lambda_{ON,i}$ for each appliance i, one can model the dynamics of each refrigeration unit via randomly sampling T_{ON} and T_{OFF} values after each state transition. However, this leads to recreation of a monitored unit in simulation.

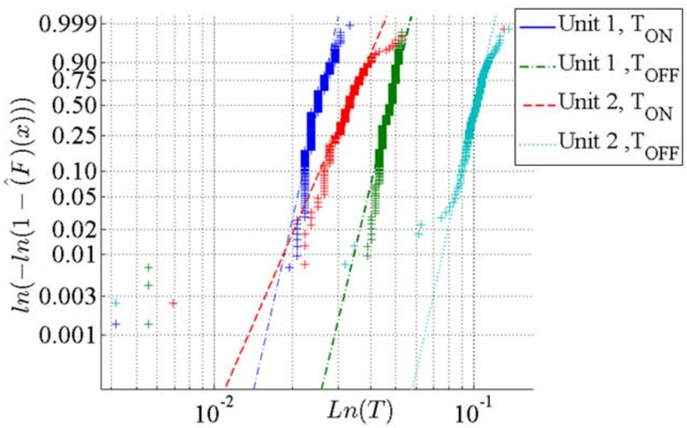

Figure 1. Weibull probability plot for T_{ON} and T_{OFF} obtained from two refrigeration units

In order to generalize the modeling strategy and to be able to simulate appliances similar to those that are monitored, one can investigate the distributions of the parameters $k_{ON,i}$, $k_{OFF,i}$, $\lambda_{OFF,i}$ and $\lambda_{ON,i}$ among all appliances. To do so, we analyzed all 373 refrigeration units, obtained T_{ON} and T_{OFF} values for each duty cycle completed during the data collection, fitted a Weibull distribution to these values, and obtained $k_{ON,i}$, $k_{OFF,i}$, $\lambda_{OFF,i}$ and $\lambda_{ON,i}$ values for each unit. Following that, we fitted various continuous probability distributions commonly used in statistical modeling to these Weibull parameters. To measure the goodness of the fit and to compare the different models obtained, we use the Bayesian Information Criterion (BIC) (Schwarz 1978). For each parameter, we rank the BIC obtained from each model. Then, we select the first three models with the lowest BIC (i.e. the top three performers). Figure 2 shows the selected models for each parameter and their performance as measured by the BIC. The dot plot on the left shows the BIC values for different models fitted to $k_{ON,i}$ and $\lambda_{ON,i}$ and the dot plot on the right shows the BIC values for different models fitted to $k_{OFF,i}$ and $\lambda_{OFF,i}$.

The results obtained for $\lambda_{OFF,i}$ and $\lambda_{ON,i}$ suggest that the top three performing models all seem to capture the distribution of these parameters similarly. For $k_{ON,i}$ and $k_{OFF,i}$, there is more variability in the performance of these models. For all of the parameters, we select the model with the best performance (i.e. with the minimum BIC). The estimated parameters for each of these models and their 95% confidence intervals are given in Table 2.

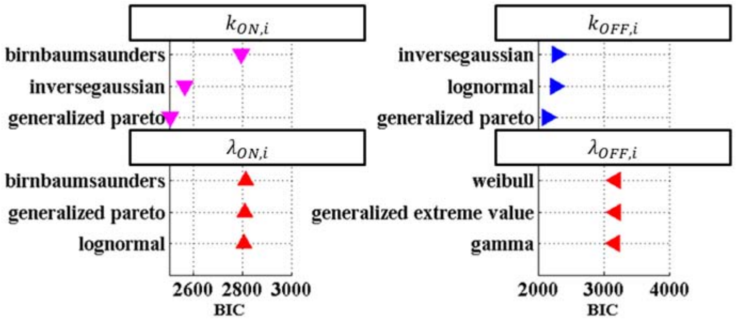

Figure 2. Selected models for $k_{ON,i}$, $k_{OFF,i}$, $\lambda_{OFF,i}$ and $\lambda_{ON,i}$ parameters vs. BIC

The proposed modeling strategy is as follows. For each refrigeration unit, we sample and store $k_{ON,i}$, $k_{OFF,i}$, $\lambda_{OFF,i}$ and $\lambda_{ON,i}$ randomly from the distributions given in Table 2 for each unit. Then, T_{ON} and/or T_{OFF} values are sampled whenever a state transition occurs using the stored parameters. This would ensure that each time the unit starts a new cycle the time dependent disturbances are incorporated in the model. Similar approaches to alternate the T_{ON} and/or T_{OFF} values can be introduced by assuming a random or controlled sampling process.

Table 2. Parameters for distributions of $k_{ON,i}$, $k_{OFF,i}$, $\lambda_{OFF,i}$ and $\lambda_{ON,i}$

Variables	Distribution	Parameter Definitions	Parameters	95% Confidence Interval
$k_{ON,i}$	Generalized Pareto	k: tail index σ: scale Θ: location	k=1.879 σ =2.232 Θ=0.793	$k_{\%95}$= [1.617, 2.141] $\sigma_{\%95}$= [1.827, 2.726]
$k_{OFF,i}$	Generalized Pareto		k=0.433 σ =5.5290 Θ=0.5404	$k_{\%95}$= [0.325, 0.542] $\sigma_{\%95}$= [4.784, 6.390]
$\lambda_{ON,i}$	Log-normal	μ: mean σ: standard deviation	μ=2.8476, σ=0.8813	$\mu_{\%95}$= [2.753, 2.942] $\sigma_{\%95}$= [0.819, 0.953]
$\lambda_{OFF,i}$	Gamma	a: shape b: scale	a=2.231, b=20.704	$a_{\%95}$= [1.938, 2.569] $b_{\%95}$= [17.678, 24.247]

COMPARISON WITH EXISTING STUDIES

Mathieu (2013) uses a range of parameters to model individual refrigerators in order to estimate their potential to provide demand response services. We adopt these parameters to match the ones defined in Equations (3-5) as shown in Table 3.

Table 3. Refrigerator parameters adopted from Mathieu (2013)

	thermal resistance, R_i (^{o}C/kW)	thermal capacitance, C_i (kWh/^{o}C)	rated power, $P_{rated,i}$ (kW)	ambient temperature, $\theta_{i,a}$(^{o}C)	thermostatic dead-band width, δ_i (^{o}C)	temperature set point, $\theta_{i,set}$ (^{o}C)
Distribution	Uniform	Uniform	Uniform	Constant	Uniform	Uniform
Range	80-100	0.4-0.8	0.2-1.0	20	1-2	1.7-3.3

In order to evaluate the variability in the R_i and C_i values given by Mathieu and the modeling strategy proposed in this paper, we created appliance populations that consist of a hundred thousand appliances. For the first population (P1), we randomly select $k_{ON,i}$, $k_{OFF,i}$, $\lambda_{OFF,i}$ and $\lambda_{ON,i}$ values and sample a T_{ON} and a T_{OFF} value from the corresponding Weibull distributions for each unit. Then we calculate the R_i and C_i values using Equations (3-5) and the $P_{rated,i}$, $\theta_{i,a}$, δ_i and $\theta_{i,set}$ distributions given in Table 3. For the second population (P2), we randomly sampled R_i and C_i parameters using the distributions given in Table 3. Then, we estimated the sample means $\bar{R}$ and $\bar{C}$, the standard deviations s_R and s_C, and the coefficients of variation $\hat{c}_{v,R} = s_R/\bar{R}$ and $\hat{c}_{v,R} = s_C/\bar{C}$ for each population. Table 4 shows the results.

Table 4. Estimated parameters per population

Population	$\bar{R}$ (°C/kW)	$\bar{C}$ (kWh/°C)	s_R (°C/kW)	s_C (kWh/°C)	$\hat{c}_{v,R}$	$\hat{c}_{v,C}$
P1	419.41	0.07	9205.7	0.07	**21**	**1**
P2	90.00	0.60	175.76	0.12	**0.06**	**0.19**

CONCLUSIONS AND FUTURE WORK

Although mean R and C values among different populations should not be compared, the coefficient of variations given in Table 4 show that the assumptions made for the parameter values and distributions in P2 results in a squeezed and less varying distribution. Appliances simulated based on assumptions followed by P2 will provide similar availability for demand response services during each cycle. Furthermore, the dispersion around the estimated sample mean, measured by the $\hat{c}_{v,ON}$ and $\hat{c}_{v,OFF}$, appears to be more significant for P1, which indicates a relatively "stretched" distribution. We believe that this might be due to the time dependent disturbances incorporated in the proposed model and the variety of different refrigeration unit types in the dataset. However, further investigation is needed to clarify the source of this dispersion.

In this study, we assumed that T_{ON} and T_{OFF} values can be modeled independently and identified hyper parameters of the distributions modeling these values separately. However, there are reasons to believe that these values are not independent. Furthermore, in our state detection algorithm we neglected the defrost cycles available in some of the units included in this study. Future work includes verifying whether the independence assumption is an acceptable one and developing an event detection algorithm that is capable of identifying the defrost cycle.

REFERENCES

Callaway, D. (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.

Callaway, D. (2011). "Can smaller loads be profitably engaged in power system services?" *Power and Energy Society General Meeting, 2011 IEEE*, 1–3 (July).

Callaway, D., Koch, S., and Mathieu, J. (2011a). “Modelling of control of aggregated heterogenous thermostatically controlled loads for ancillary services.” *Power Systems Computation Conference*.

Callaway, D. S. and Hiskens, I. A. (2011b). “Achieving controllability of electric loads.” *Proceedings of the IEEE*, 99(1), 184–199.

Faruqi, A.H.R., S.S., G., Bode, J., Mangasarian, P., Rohmund, I., Wikler, G., Glosh, D., and Yoshida, S. (2009). “A national assesesment of demand response potential.” *Federal Energy Regulatory Commission (FERC)*.

Gellings, C., Samotyj, M., and Howe, B. (2004). “The future’s smart delivery system [electric power supply].” *Power and Energy Magazine, IEEE*, 2(5), 40 – 48.

Kara, E. C., Berges, M., Krogh, B., and Kar, S. (2012). “Using smart devices for system-level management and control in the smart grid: A reinforcement learning framework.” *IEEE SmartGridComm’12*.

Kara, E. C. and Bergés, M. (2013a). “Demand Response in Buildings: Engaging Thermostatically Controlled Loads in the Power Grid.” In *Proceedings of the 2013 ASCE International Workshop on Computing in Civil Engineering*.

Kara, E. C., Kolter, Z., Bergés, M., Krogh B., Hug, G. and Yuksel T. (2013b). “A Moving Horizon State Estimator in the Control of Thermostatically Controlled Loads for Demand Response.” *IEEE SmartGridComm’13*.

Kassakian, J. and Schmalensee, R. (2011). “The future of the electric grid: An interdisciplinary MIT study.” Technical Report, *Massachusetts Institute of Technology.*

Kiliccote, S., Piette, M. A., and Ghatikar, G. (2011). “Smart buildings and demand response.” *AIP Conference Proceedings*, 1401(1), 328–338.

Mathieu, J. L. (2013). "Modeling, analysis, and control of demand response resources." PhD Thesis.

Mathieu, J. L., Koch, S., and Callaway, D. S. (2012). “State estimation and control of electric loads to manage real-time energy imbalance.” *Power Systems, IEEE Transactions on*, PP(99), 1.

Nelson, W. (1982), “Applied Life Data Analysis”, *Addison-Wesley*.

Trivedi, Kishor Shridharbhai. (1982) “Probability and statistics with reliability, queuing, and computer science applications.” *Vol. 2002. Englewood Cliffs: Prentice-Hall*

Kelle, P., Silver, E.A. (1990) “Safety stock reduction by order splitting.”*Naval Research Logistics*, 37 (5), pp. 725-743.

Kamgarpour, M., Ellen, C., Soudjani, S. E. Z., Gerwinn, S., Mathieu, J. L., Mullner, N., and Lygeros, J. (2013). “Modeling options for demand side participation of thermostatically controlled loads.” In *Bulk Power System Dynamics and Control-IX Optimization, Security and Control of the Emerging Power Grid (IREP), 2013 IREP Symposium* (pp. 1-15). IEEE.

Schwarz, G. (1978). "Estimating the dimension of a model." *The annals of statistics* 6, no. 2 461-464.

Zimmermann, J.P., Evans, M., Griggs, J., King, N., Harding L., Roberts, P. and Evans, C. (2012). “Household Electricity Survey: A study of domestic electrical product usage”, *Intertek Report R66141*.