Title

Quantifying the effect of multiple demand response actions on electricity demand and building services via surrogate modeling

Permalink

Authors

Luo, Na
Langevin, Jared
Chandra Putra, Handi

Publication Date

2021-09-01

DOI

10.26868/25222708.2021.30704

Building Technologies & Urban Systems Division
Energy Technologies Area
Lawrence Berkeley National Laboratory

Quantifying the effect of multiple demand response actions on electricity demand and building services via surrogate modeling

Na Luo, Jared Langevin, Handi Chandra-Putra

Lawrence Berkeley National Laboratory, Building Technology and Urban Systems Division, Berkeley, CA,

Energy Technologies Area
September 2021

This work was supported by the Assistant Secretary for Building Technologies Office of the U.S. Department of Energy under Contract No.DE-LC-000L048

Quantifying the effect of multiple demand response actions on electricity demand and building services via surrogate modeling

Na Luo[1], Jared Langevin[1], Handi Chandra-Putra[1]
[1]Lawrence Berkeley National Laboratory, Building Technology and Urban Systems Division, Berkeley, CA, USA

Abstract

The expansion of commercial building demand response (DR) as a demand-side management resource for the electric grid necessitates new decision support resources for customers to rapidly assess the benefit-risk tradeoffs of candidate load flexibility strategies. This work develops surrogate models of load flexibility impacts on office building electricity demand and indoor temperature. The surrogate models are fit to a large synthetic database generated via whole building simulations of multiple flexibility strategies under a variety of conditions; the models are translated to a Bayesian framework to allow straightforward communication of uncertainty and parameter updating given new evidence. The strong predictive performance of the models underscores their potential utility in guiding DR decision-making in office settings.

Key Innovations

- A comprehensive set of load flexibility strategies are determined and applied in EnergyPlus simulations as OpenStudio Measures.
- The simulated effects of the flexibility strategies on building demand and indoor temperature are compiled into a large synthetic database covering medium offices across climates and vintages in the summer season.
- Multiple regression surrogate models are fit to the synthetic data to predict changes in electricity demand and indoor temperature based on flexibility strategy characteristics and contextual conditions.
- The surrogate regression models are translated to a Bayesian framework to facilitate communication of prediction uncertainty and future updating of model parameters.
- The fit and predictive performance of the surrogate regression models are evaluated using quantitative metrics and qualitative graphical checks.

Practical Implications

The surrogate models provide a simple and computationally tractable tool for commercial building operators to forecast the potential benefits (e.g, demand reductions and associated economic incentives) and risks (e.g., temperature increases) of participating in DR events; the model implementation facilitates understanding of prediction uncertainties and updating given new field data.

Introduction

Commercial demand response (DR), which encompasses a set of time-dependent utility program activities and tariffs that seek to reduce electricity demand or shift demand across time periods (Pinson and Madsen (2014)), is likely to play an expanded role in the coming years as a demand-side management tool that facilitates grid reliability and resilience in the face of day-to-day stresses, increased variable renewable energy penetration, and emergency events (Surampudy et al. (2019)). On the customer side, the increasing role of commercial building DR as a grid flexibility resource necessitates new decision making resources that enable rapid, real-time assessments of the benefit-risk tradeoffs across candidate load flexibility strategies, in terms of potential economic gains and changes to core building services (e.g., HVAC, lighting, plug power) (Smith and Controls (2010)).

Previous reviews of load flexibility strategies for commercial DR group strategies into four areas, namely HVAC, lighting, miscellaneous equipment, and measures that work across components or end uses (Motegi et al. (2007)). Adjustments to commercial HVAC and lighting schedules are considered as particularly attractive strategies, due to the substantial portion of total loads that these strategies can affect and their comparatively lower risks to occupant comfort than strategies that make centralized adjustments to air distribution or cooling systems, for example. Outside of HVAC and lighting, commercial flexibility strategies may also target miscellaneous electric equipment (e.g., computers, fountain pumps, industrial process loads etc.) to reduce demand without influencing the basic activities of occupants. Non-component-specific measures, which coordinate across the aforementioned types of DR strategies based on conditions such as outdoor temperature and electricity price, constitute more sophisticated ways of managing a building's demand dynamically, but also require a higher degree of programming.

Recent studies have attempted to represent commercial load flexibility strategies in dynamic building energy simulations (Bossmann and Eser (2016); Chen et al. (2018)). Such studies use physics-based energy modeling tools such as EnergyPlus, DOE-2, and TRNSYS to investigate changes in building demand, thermal dynamics, and changes to other building services from flexible operations during DR events. Since the underlying modeling tools are able to capture both whole building and sub-system/component changes under DR at temporal increments that are suitable for grid planners (e.g., 15 minute–to –hourly changes in HVAC, lighting, and electric equipment schedules), these tools are well-suited to represent the various types of load flexibility strategies identified above. Accordingly, other tools such as Demand Response Quick Assessment Tool (DRQAT) (Center (2015)), and EnergyPLAN (Neves et al. (2015)) have drawn from physics-based building simulation capabilities specifically for the purpose of assessing tradeoffs across candidate flexibility strategies.

Data-driven models offer an alternative to physics-based simulations for characterizing a building's response to dynamic load adjustments (Wang and Chen (2019)). For example, Yin et al. (2016) developed a surrogate modeling framework for estimating the theoretical demand flexibility in both commercial and residential buildings, in which regression models were trained to large EnergyPlus datasets on the simulated demand impacts of thermostat adjustment strategies. Kara et al. (2014) similarly used a data-driven surrogate modeling approach to quantify the flexibility of residential thermostatically controlled loads for demand response. Amara et al. (2015) conducted a review comparison among different types of data-driven models and approaches, identifying a hybrid physics-statistical modeling framework as most effective for managing building energy use. Nevertheless, surrogate models and other data-driven approaches such as machine learning require substantial measured or simulated data for model training – including building energy use, indoor and outdoor environmental variables, and system control data (Korolija et al. (2013)). Accordingly, data-driven approaches remain a promising but uncommon basis for informing commercial load flexibility strategies in practice (Kathirgamanathan et al. (2021)).

Overall, existing literature suggests the following key challenges to quantitative assessment of commercial building load flexibility strategies:

- existing models of load flexibility are not easily adapted to specific building instances; given the lack of modeling and computational resources, it is not feasible for building owners, operators, or consultants to build unique physics-based models of the potential impacts of flexibility strategies in specific building applications,
- few studies represent multiple load flexibility strategies at once, despite the potential for such flexibility packages to yield deeper demand reductions while distributing risks across building services, and
- few existing studies communicate the uncertainties in their predictions, imparting false confidence about flexibility impacts.

To address such limitations, this study develops surrogate models of key load flexibility strategies for office settings that are robust, comprehensive, and adaptable to new information. Specifically, we use EnergyPlus to generate synthetic data on the potential impacts of an array of commercial load flexibility strategies and packages in a variety of operational contexts; train a series of multiple regression models of building demand and service changes under these strategies; and translate the regression-based surrogate models to a Bayesian framework to facilitate communication of model prediction uncertainty and parameter updating given new evidence. The developed models are intended to serve as simple tools that commercial DR participants can use to prospectively assess the relative benefits and risks of candidate load flexibility response strategies under a particular set of conditions (e.g., weather, occupancy, incentives).

Methods

Surrogate models of commercial building demand and services under load flexibility are developed as follows: 1) define commercial DR contexts (climate, building type, building vintage) and develop candidate load flexibility Measures in OpenStudio/EnergyPlus; 2) simulate load flexibility Measures across climate zones, building types, and vintages of interest; 3) compile results into synthetic database covering simulated electricity demand and indoor temperature outcomes under the various measure settings; 4) fit a series of multiple regression models of building demand and indoor temperature under load flexibility using the synthetic database; translate the regression models to a Bayesian inference framework to address model uncertainty and new parameter updating; and 5) assess model fit and predictive performance using quantitative metrics and qualitative graphical checks. Figure 1 summarizes the methodological steps used in this study, which are further enumerated below.

Definition of DR contexts and load flexibility strategies

DR contexts for load flexibility strategies were defined based on the prototypical building types and vintages published for 19 representative weather locations by the U.S. DOE (2020). In this study, we focus on the medium office (MediumOfficeDetailed) prototype across the 1980-2004 and 2010 vintages, which we use to represent older and newer buildings, respectively. We simulate across a range of different climate zones – 2A, 3C, 4A, and 6B – to capture the influence of variation in weather as well as location-

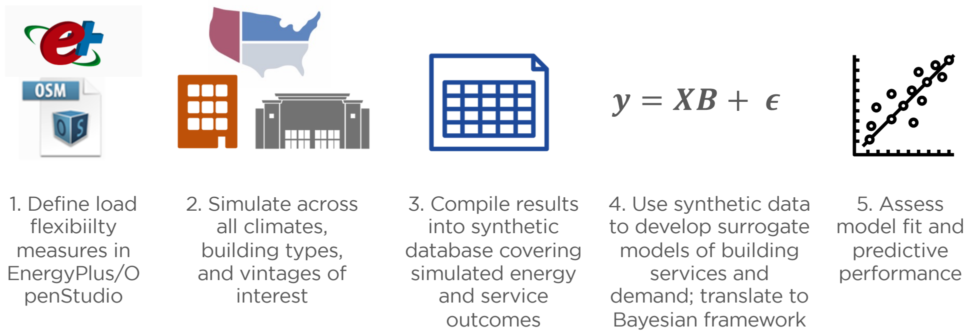

Figure 1: Summary of methodological steps.

specific building design codes. In all, we simulate 8 unique contexts (2 vintages * 4 climate zones).

Table 1 summarizes the load flexibility strategies represented, along with key characteristics. We represent four individual flexibility strategies, which are applied during the DR event hours of 3–7 PM: 1) lighting dimming, 2) plug load reduction through low-priority device switching, 3) global temperature adjustment (GTA), and 4) GTA + pre-cooling. In the simulations, we apply each strategy to a baseline daily schedule across summer days, excluding weekends. The adjusted schedule is written in .csv format and used in the simulation using OpenStudio Software Development Kit with EnergyPlus engine. For both lighting and MELs strategies, the reductions are generated using a continuous uniform distribution bounded from 0 to 100%. Adjustment settings for GTA and pre-cooling are also generated using a discrete uniform distribution; GTA cooling set point increases during the DR period are sampled between the range of 1°F and 6°F, while pre-cooling set point decreases are sampled between the range of 1°F and 4°F. Pre-cooling is represented for half of the simulated summer days to ensure a balance in the dataset between days with and without pre-cooling applied; pre-cooling application ranges between 1 and 8 hours directly preceding the DR event period, with durations drawn from a discrete uniform distribution.

Whole building simulation of load flexibility strategies

Load flexibility strategies and contexts are simulated using the EnergyPlus engine (Strand et al. (2000)) and OpenStudio software development kit Guglielmetti et al. (2011). Flexibility strategies are developed as OpenStudio Measures, small programs that adjust an EnergyPlus model across various dimensions including the building and system characteristics and operational and occupancy schedules. Measures developed to represent the aforementioned flexibility strategies modify the fractional schedules for lighting and office equipment power, as well as the cooling temperature setpoint schedule during specific times of day – in this case, before, during, and after the 3–7 PM DR event window described in the previous section. These Measure definitions have been published on a GitHub[1] repository.

The full set of Measures and DR contexts were generated via batch simulations that leverage the OpenStudio Workflow (OSW). Using the OSW, the MediumOfficeDetailed prototype model was seeded as the OpenStudio Model, to which flexibility Measures and an additional set of Reporting Measures were applied. Flexibility Measure arguments include the magnitudes and durations of each candidate strategy, while Reporting Measure arguments contain the type of reporting variable desired and the time resolution with which that variable should be reported. Given these arguments, the OSW executes the EnergyPlus engine and generates the required reporting files across the various DR contexts. All simulations use an hourly time step.

Development of synthetic database

After conducting the batch simulations, hourly results are compiled into a synthetic database, with a particular focus on electricity demand and indoor temperature outputs and the potential predictors of these outputs. Regarding indoor temperature results, the following adjustment to the raw temperature data generated for each of the 28 MediumOfficeDetailed prototype zones to yield a single indoor temperature variable to use in the dataset:

$$T_{ave} = \frac{\sum_{i=1}^{n}(T_i \times Occ_i)}{\sum_{i=1}^{n} Occ_i} \quad (1)$$

Where T_{ave} represents the occupant-weighted averaged indoor temperature; i represents the conditioned zone number, while n represents the total number of conditioned zones; T_i represents the indoor temperature within zone i; and Occ_i represents the occupant numbers within zone i.

Raw simulation results are further filtered in the following ways:

[1] `https://github.com/jtlangevin/flex-bldgs/tree/master/measures`

Table 1: Load flexibility strategy assumptions.

Category	Measure	Low	Magnitude of adjustment (Uniformly distributed)	High	Duration of adjustment
HVAC	Global cooling temp. adjustment (GTA)	+1°F	to	+6°F	3-7 PM
	GTA	+1°F	to	+6°F	3-7PM
	+ pre-cooling	-1°F	to	-4°F	1 to 8-hour ahead (Uniformly distributed)
Lighting Plug Loads	Dimming Low-priority device switching	-0%	to	-100%	3-7 PM

- we restrict to summer data only (Jun–Sep) given the focus on cooling adjustments in the set of flexibility strategies examined,
- we restrict the data to weekdays only, given significant building demand during weekday hours and high variability in weekday occupancy schedules during the chosen DR event window (3–7PM), and
- within each day, we restrict hourly data points to those that fall in the DR event window, as well as any hours preceding the window in which pre-cooling was simulated and one hour following the event, to capture potential rebounds in demand as the HVAC system recovers from adjustments to the cooling set point.

As mentioned, the synthetic database also reflects potential predictors of changes to building demand and indoor temperature under load flexibility, including: outdoor environmental conditions (e.g., outdoor temperature and humidity), changes in operational schedules consistent with the various flexibility strategies, and time lags of these variables. The synthetic data are published on the aforementioned GitHub[2].

Surrogate model development

To develop the surrogate models of building electricity demand and indoor temperature under load flexibility strategies, we fit multiple linear regressions to the synthetic data, prioritizing model simplicity and interpretability. Model predictors are specified based on domain expertise about the variables that are likely to have a real-world relationship to the building demand and indoor temperature outputs of interest. Interaction terms are included under the expectation that the influence of certain model predictors on the output is conditioned on the values of other predictors in the model, such as the moderating effect of weather and occupancy on the impacts of changing a zone cooling setpoint.

To predict changes to building demand, we fit separate surrogate models for strategies that drive changes in demand through changes in thermal loads (e.g., global temperature adjustment and pre-cooling) and strategies that do not primarily influence demand through changing thermal loads (e.g., dimming lights and reducing plug load power). We also fit a separate demand model for changes to demand during the pre-cooling period, given the different thermal load dynamics of this period compared to the DR event window (e.g., pre-cooling *increases* in thermal load vs. the *decreases* during the DR event window).

To predict changes to indoor temperature, we further restricted the synthetic training dataset to exclude data points that reflect only lighting or plug load changes, as such changes were observed to have only small effects on zone temperature. Furthermore, the temperature model is trained on synthetic data with the one hour rebound period removed, as the indoor temperature tends to move quickly back to the thermostat set point during this period, leaving no temperature changes for the model to predict.

Table 2 summarizes the ultimate set of surrogate models that was developed, and further shows the outputs and set of predictor variables chosen for each model. The surrogate models for demand yield demand shed intensity (W/ft^2), while the surrogate model for indoor temperature yields the change in indoor temperature *relative to the baseline setpoint* (ºF). Predictor variables are generally of four types: 1) outdoor environmental conditions and occupancy, 2) load flexibility strategy characteristics and their single time step lags (e.g., changes in cooling setpoint, lighting and plug load schedules), 3) time duration indicators (e.g., hours since flexible operations have started and ended), and 4) interactive terms that capture the moderating effect of one predictor variable on another. Note that in the models of thermally-driven changes in demand and indoor temperature during the DR event window, inputs concerning the magnitude and duration of any pre-cooling before the event are included to account for the effects of pre-cooling during the DR event window.

Translation of models to Bayesian framework

In practice, the input/output relationships that are mapped by the surrogate models may differ between prototypical and real offices. To mitigate this issue, uncertainty in surrogate model predictions should be represented in a straightforward manner and the models must be flexible to updating given new evidence collected in real building environments. Both of these criteria are well supported by translation of the surrogate models to a Bayesian inference frame-

[2] `https://github.com/jtlangevin/flex-bldgs/tree/master/data`

Table 2: Summary of demand and temperature surrogate model inputs and outputs. Shown are model types, outputs, and predictors, with interactive predictor variables highlighted in red.

	Whole Building Demand (DR, Non-thermal related)	**Whole Building Demand (DR, Thermal related)**	**Whole Building Demand (Pre-cool)**	**Indoor Temperature (DR)**
Output	Demand shed intensity (W/ft^2)	Demand shed intensity (W/ft^2))	Demand shed intensity (W/ft^2)	Indoor temperature change (^{o}F)
Input	Lighting dimming (%)	Outdoor temperature (^{o}F)	Outdoor temperature (^{o}F)	Outdoor temperature (^{o}F)
	Plug loads reduction (%)	Outdoor humidity	Outdoor humidity	Outdoor humidity
		Occupancy fraction	Occupancy fraction	Occupancy fraction
		Cooling set pt. change (^{o}F)	Cooling set pt. change (^{o}F)	Cooling set pt. change (^{o}F)
		Lighting dimming (%)	Hours since pre-cool started	Cooling set pt. lag (^{o}F)
		Plug loads reduction (%)	*Cooling change * Outdoor temp.*	Hours since DR started
		Cooling set pt. lag (^{o}F)	*Cooling change * Occ. fraction*	Pre-cool set pt. change (^{o}F)
		Hours since DR started	*Cooling change * Since Precool started*	Pre-cool duration
		Hours since DR ended		*Cooling change * Outdoor temp.*
		Pre-cool set pt. change (^{o}F)		*Cooling change * Occ. fraction*
		*Cooling change * Outdoor temp.*		*Cooling change * Since DR started*
		*Cooling change * Occ. fraction*		*Pre-cool change * Pre-cool duration*
		*Cooling change * Since DR started*		

work.

Bayesian inference treats model parameters as random variables, deriving the *posterior* probability of model parameter values as a function of the *likelihood* of observed data given the parameters and *prior* parameter probabilities:

$$p(\theta|X, \alpha) \propto p(X|\theta)p(\theta|\alpha) \tag{2}$$

Where θ is a vector of model parameters, X is the observed data, and α is a vector of hyperparameters; $p(\theta|X, \alpha)$ is the posterior parameter probability given the observed data and hyperparameter values; $p(X|\theta)$ is the likelihood of observed data given the parameter values; and $p(\theta)$ is the prior parameter probability given the hyperparameter values.

Following from equation 2, the *posterior predictive distribution* of new data points $\hat{x}$ is generated by marginalizing over the posterior parameter distributions:

$$p(\hat{x}|X, a) = \int p(\hat{x}|\theta)p(\theta|X, \alpha)\, d\theta \tag{3}$$

The range of model predictions generated by equation 3 satisfies the requirement of communicating uncertainty in model outcomes, while the ability to weigh prior expectations about model parameters against new evidence as in equation 2 supports model updating with data collected in the field. To implement the models in this framework, we use PyMC3 (Salvatier et al. (2016)), a Python package for probabilistic programming. Models are initialized using the previously described synthetic data to populate the likelihood function and assuming diffuse parameter prior distributions ($\theta \sim N(0, 10)$) to reflect our lack of *a priori* beliefs about parameter values. The posterior parameter distributions estimated through this process are stored on GitHub[3] and may serve as informative prior distributions in subsequent rounds of model updating.

[3] `https://github.com/jtlangevin/flex-bldgs`

Model evaluation

Model fit and predictive performance were assessed via the following:

1. *R-squared (R^2)* is a goodness–of–fit metric; R^2 indicates the proportion of the variability in the response data about its mean that is explained by the modeled independent variable(s).
2. *Absolute Relative Error (ARE)* is a predictive accuracy metric; ARE subtracts observed from predicted values and normalizes each difference by the observed values (Yin et al. (2016)).
3. *Mean Absolute Deviation Percentage (MADP)* is a second measure of accuracy; MADP is the ratio of the sum of deviations between observed and model-predicted values to the sum of observed values (Lucas Segarra et al. (2019)).
4. *Variance inflation factor (VIF)* is a measure of variable multicollinearity; VIF is the ratio of overall model variance to the variance of model including only a given predictor variable (Craney and Surles (2002)). VIF values higher than 10 are assumed to suggest the need for model adjustments to improve the stability of variable coefficient estimates.
5. *Posterior predictive checks (PPCs)*, which are relevant specifically to the Bayesian model implementation, graphically compare real and simulated data for systematic discrepancies (Gelman et al. (2004)).

Results

Electricity demand models

Figure 2 shows the cumulative distribution of the absolute relative error (ARE) of predictions from the model of non-thermally-driven changes in building demand during the DR event window, for new and old office vintages, respectively. The red dashed line establishes a threshold for predictions that fall within 20% of the ground truth values from the synthetic database. In this case, 91% and 89% of the predictions meet this accuracy threshold for the new and old office vintages, respectively, while MADP values are 9–10% and R^2 values are 0.93 and 0.94, respectively.

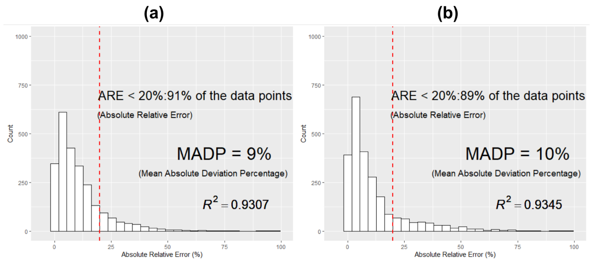

Figure 2: Histogram of the absolute relative error in model of non-thermally-driven changes in demand during the DR event window for newer (a) and older (b) medium office vintages.

Figure 3 shows the same results as Figure 2 for the model of thermally-driven changes in building demand during the DR event window. Here, 90% of the predictions meet the accuracy threshold for both new and old office vintages, while MADP values for both vintages are 8% and R^2 values are 0.96 and 0.97, respectively.

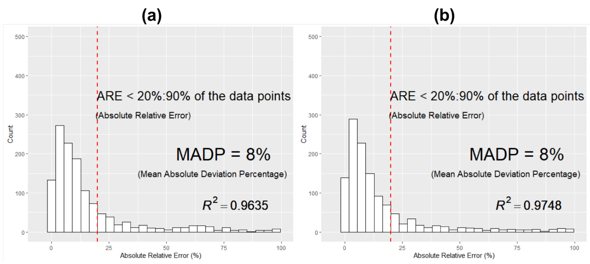

Figure 3: Histogram of the absolute relative error in model of thermally-driven changes in demand during the DR event window for newer (a) and older (b) medium office vintages.

Finally, Figure 4 shows the same results as Figures 2–3 for the model of changes in demand from implementing pre-cooling. For this model, 86% and 87% of the predictions meet the accuracy threshold for new and old office vintages, respectively, while MADP values are 13% and 12%, respectively, and R^2 values are 0.86 and 0.90, respectively.

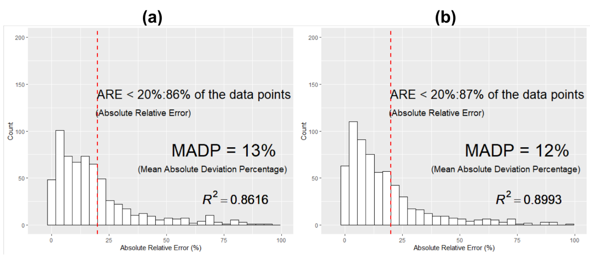

Figure 4: Histogram of the absolute relative error in the model of changes in demand during the pre-cooling period for newer (a) and older (b) medium office vintages.

Indoor temperature model

Figure 5 shows the same information as Figures 2–4 for the model of changes in indoor temperature during the DR event window. 77% and 79% of predictions fall within 20% of ground truth values in new and old buildings respectively, MADP values are 16% and 15%, respectively, and R^2 values are 0.78 and 0.82, respectively.

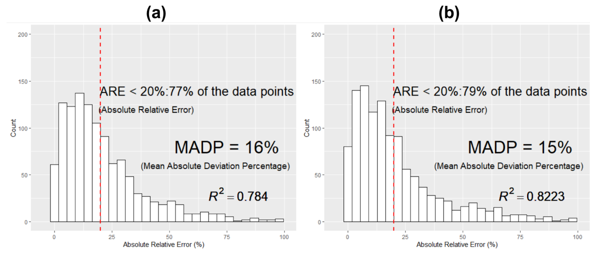

Figure 5: Histogram of the absolute relative error in the model of changes to indoor temperature during the DR event window for newer (a) and older (b) medium office vintages.

Bayesian model implementation

All models were successfully re-estimated in a Bayesian framework and subjected to a series of posterior predictive checks (PPCs) as described in the Methods. Figures 6a–c and 6d–f demonstrate PPC results for the model of thermally-driven demand changes during the DR event window in new and old office vintages, respectively. Checks of parameter posterior distributions (6a, 6d) demonstrate that under the absence of informative prior expectations about parameter values, parameter posterior distributions are centered on the frequentist point estimates that underpin model results in the previous sections. Furthermore, checks on the posterior distribution of model outputs (6b–c, 6e–f) reveal no systematic discrepancies between simulated and observed data (this finding holds for the models not shown in Figure 6 as well).

Discussion

Model assessment results generally show strong predictive performance for the developed surrogate models, which yield MADP values of less than 16% and R^2 values above 0.78 across the various model types and DR contexts explored. Demand models display particularly high accuracy, yielding MADP values less than 10% and R^2 values above 0.93 across newer and older vintages. This level of predictive accuracy compares favorably with previous studies (Yin et al. (2016)) and suggests the models would be practically useful in guiding decision-making regarding load flexibility strategies for offices under a wide variety of conditions.

While these preliminary findings are encouraging, we note the following important limitations at this stage of the work:

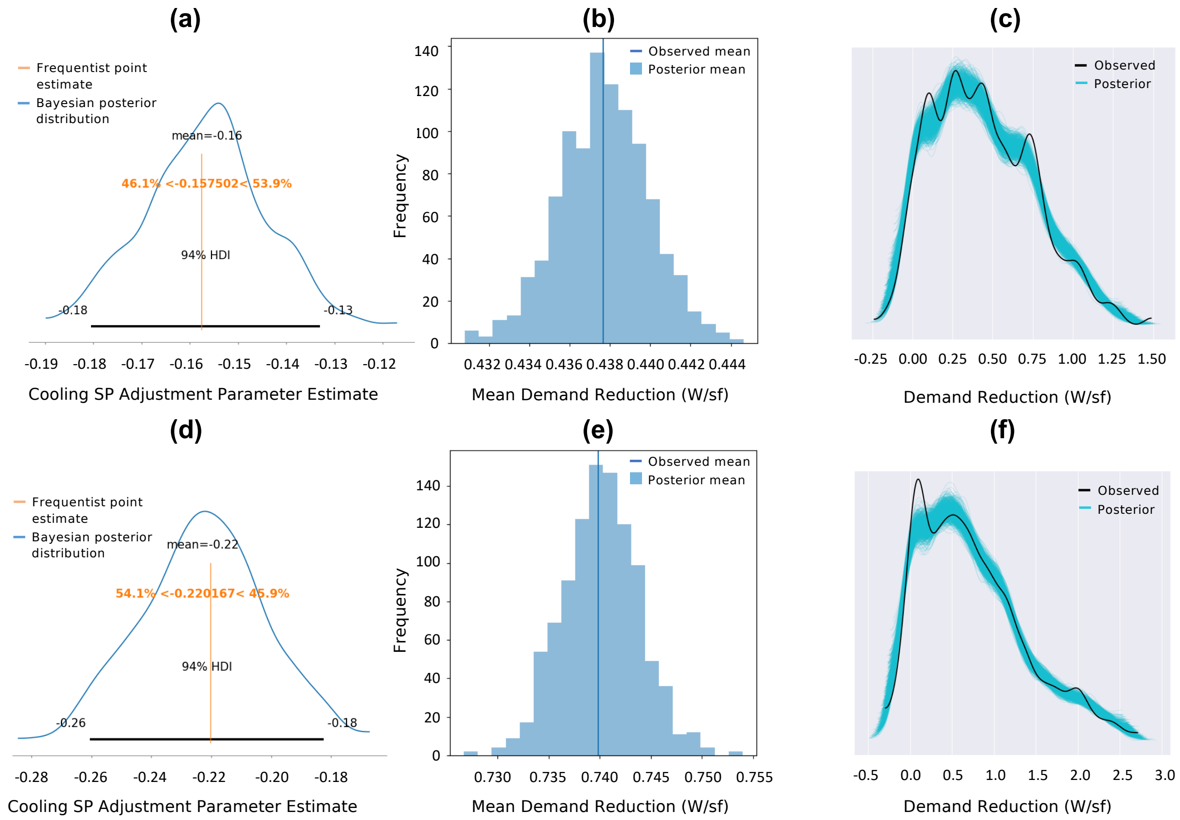

Figure 6: Bayesian estimation and assessment of thermally-driven demand change model for newer (a–c) and older (d–f) medium office vintages. (a, d) Example kernel density distribution of cooling set point change parameter coefficent compared to frequentist estimate. (b, c, e, f) Posterior predictive checks against observed mean demand change and the distribution of observed demand change, including histogram (b, c) and kernel density estimation (e, f) of demand reduction output as observed in the underlying synthetic dataset and predicted by the Bayesian models.

- we investigate a limited set of load flexibility strategies and contexts, focusing on strategies that adjust control schedules rather than more fundamental modifications of equipment settings, in medium offices in summer only; DR programs are likely to target a wider variety of commercial building types – particularly larger ones, where DR participation may be more cost-effective for customers – potentially across both cooling and heating seasons,
- our models are limited to producing decision-making insights at the hourly temporal resolution, in line with the resolution that outcomes are assessed for many traditional DR programs and in wholesale energy markets; the models do not provide insights regarding faster DR services such as frequency regulation or load modulation, for which sub-hourly predictions are required,
- while we predict the likely magnitude of impacts from load flexibility strategies on demand and indoor temperature, we do not assign operator valuations of these impacts; in practice, operators may weigh changes to demand (and associated economic benefits) more heavily than changes to temperature, or vice versa, and
- we do not extend our modeling to other building services such as control of humidity, illuminance, and CO_2, which are not as readily measured or modeled by the physics-based simulations that underpin our synthetic database of load flexibility impacts.

Future work will address some of these limitations by investigating commercial load flexibility strategies in the winter; expanding synthetic data to cover additional building types – retail and large offices, and incorporating operator load adjustment preference data generated from discrete choice experiments to attach valuations to predicted changes in building demand and services.

Conclusion

In this study, we used a surrogate modeling approach to predict changes in office building electricity demand and indoor temperature under candidate load flexibility strategies, including adjustments to HVAC, lighting, and plug load schedules. The surrogate models were fit to a large synthetic database generated via whole building simulations of the strategies under a variety of conditions; models were translated to a Bayesian framework to allow straightforward communication of uncertainty and parameter updating given new evidence. The surrogate models showed strong predictive performance, yielding overall prediction errors of less than 16% and R^2 values above 0.83, underscoring their potential utility in guiding DR decision-making in office settings.

Acknowledgment

This work was supported by Building Technologies Office of the U.S. Department of Energy under Contract No.DE-LC-000L048. We appreciate technical input and feedback from Mary Ann Piette, Tianzhen Hong, Margaret Taylor, and Jeff Deason of Lawrence Berkeley National Laboratory.

References

Amara, F., K. Agbossou, A. Cardenas, Y. Dubé, and S. Kelouwani (2015). Comparison and simulation of building thermal models for effective energy management. *Smart Grid and renewable energy 6(04)*, 95.

Bossmann, T. and E. J. Eser (2016). Model-based assessment of demand-response measures—a comprehensive literature review. *Renewable and Sustainable Energy Reviews 57*, 1637–1656.

Center, D. R. R. (2015). Demand response quick assessment tool (drqat) for buildings. *LBNL technical report*.

Chen, Y., P. Xu, J. Gu, F. Schmidt, and W. Li (2018). Measures to improve energy demand flexibility in buildings for demand response (dr): A review. *Energy and Buildings 177*, 125–139.

Craney, T. A. and J. G. Surles (2002). Model-dependent variance inflation factor cutoff values. *Quality Engineering 14(3)*, 391–403.

DOE (2020). Commercial prototype building models. https://www.energycodes.gov/development/commercial/prototype_models (Accessed 01/30/21).

Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin (2004). *Bayesian Data Analysis* (2nd ed. ed.). Chapman and Hall/CRC.

Guglielmetti, R., D. Macumber, and N. Long (2011). Openstudio: an open source integrated analysis platform. *National Renewable Energy Lab.(NREL) technical report 5500-51836.*

Kara, E. C., M. D. Tabone, J. S. MacDonald, D. S. Callaway, and S. Kiliccote (2014). Quantifying flexibility of residential thermostatically controlled loads for demand response: a data-driven approach. In *In Proceedings of the 1st ACM conference on embedded systems for energy-efficient buildings.*

Kathirgamanathan, A., M. De Rosa, E. Mangina, and D. P. Finn (2021). Data-driven predictive control for unlocking building energy flexibility: A review. *Renewable and Sustainable Energy Reviews 135*, 110120.

Korolija, I., Y. Zhang, L. Marjanovic-Halburd, and V. I. Hanby (2013). Regression models for predicting uk office building energy consumption from heating and cooling demands. *Energy and Buildings 59*, 214–227.

Lucas Segarra, E., H. Du, G. Ramos Ruiz, and C. Fernández Bandera (2019). Methodology for the quantification of the impact of weather forecasts in predictive simulation models. *Energies 12*(7).

Motegi, N., M. A. Piette, D. S. Watson, S. Kiliccote, and P. Xu (2007). Introduction to commercial building control strategies and techniques for demand response. *Lawrence Berkeley National Lab.(LBNL) technical report 59975.*

Neves, D., A. Pina, and C. A. Silva (2015). Demand response modeling: A comparison between tools. *Applied Energy 146*, 288–297.

Pinson, P. and H. Madsen (2014). Benefits and challenges of electrical demand response: A critical review. *Renewable and Sustainable Energy Reviews 39*, 686–699.

Salvatier, J., T. V. Wiecki, and C. Fonnesbeck (2016). Probabilistic programming in python using pymc3. *PeerJ Computer Science 2*, e55.

Smith, K. and J. Controls (2010). Scaling demand response through interoperability in commercial buildings. In *In Grid Interop forum.*

Strand, R. K., D. B. Crawley, C. Pedersen, R. Liesen, L. Lawrie, and F. Winkelmann (2000). Energyplus: A new-generation energy analysis and load calculation engine for building design. In *In Proc. Association of Collegiate Schools of Architecture Technology Conference.*

Surampudy, M., B. Chew, K. M., and G. T. (2019). 2019 utility demand response market snapshot. *SEPA report*.

Wang, Z. and Y. Chen (2019). Data-driven modeling of building thermal dynamics: Methodology and state of the art. *Energy and Buildings 203*, 109405.

Yin, R., E. C. Kara, Y. Li, N. DeForest, K. Wang, T. Yong, and M. Stadler (2016). Quantifying flexibility of commercial and residential loads for demand response using setpoint changes. *Applied Energy 177*, 149–164.