Pattern Analysis of Dynamic Grid Incentives and the Implications on Optimal Control of Building Thermal Energy Storage

Lily X. Li
The Pennsylvania State University, xml5187@psu.edu

Gregory S. Pavlak

Pattern Analysis of Dynamic Grid Incentives and the Implications on Optimal Control of Building Thermal Energy Storage

Lily X. Li[1], Gregory S. Pavlak[1*]

[1] The Pennsylvania State University,
University Park, PA, U.S.A
Contact Information (xml5187@psu.edu, gxp93@psu.edu)

* Corresponding Author

ABSTRACT

Building thermal energy storage has been utilized for decades for various objectives, such as reducing peak electrical demand, reducing building operating expenses, and increasing the efficiency of systems when charged from waste heat or free cooling. As building thermal storage control strategies become more dynamic, optimization of building performance often considers multiple objectives that aim to improve building performance in energy, economic, environmental, and grid support categories. The dynamics of the incentive signal used for one objective, as well as its relation to signals from other objectives—for instance, whether the signals are "in sync" or are "conflicting"—heavily influence the tradeoffs that may exist among performance objectives. To better understand the degree of alignment that may exist between grid incentive signals, we apply unsupervised learning to a novel grid data set that includes hourly signals for energy price and marginal carbon emissions. Clustering algorithms identify common patterns in the dynamic signals. Overall, Hierarchical Clustering demonstrated the best performance, evaluated by DB index and Silhouette score. While the algorithms did not find distinctive patterns among the carbon signals, they did identify 7 to 9 patterns within the January and July pricing signals. The highly fluctuating nature of the carbon emission signals could lead to a diverse range of tradeoffs between building energy cost and carbon emission reduction objectives, if the signals were used as the basis for a building control optimization problem. This finding iterates the importance of understanding incentive signal dynamics, in both individual and collective contexts, and the implications for development of new control technologies for grid-interactive buildings.

1. INTRODUCTION

Building thermal storage has been utilized for decades to facilitate building performance improvements across various operational objectives. In published works from the last decade, building thermal storage has been shown to excel at reducing building operating expenses and utility costs (Henze et al., 2004; Greensfelder et al., 2011), shifting or shedding peak electric load (Klaassen et al., 2002; Pavlak et al., 2014), and increasing the efficiency of systems when charged from waste heat or free cooling. More recently, studies by Niu et al. (2019) and Delgado et al. (2020) suggest that building thermal storage is also effective in providing grid flexibility and compensating for the power fluctuations created by variable renewable energy generation. In addition, our recent work has shown that building flexibility can reduce the required size of battery storage in building-level microgrids (Swaminathan et al., 2020), and that there are potential benefits to coordinating the operations of a portfolio of thermal energy storage resources to manage price volatility and uncertainty (Yu and Pavlak, 2021).

As building thermal storage control strategies become more dynamic, optimization of building performance often considers multiple objectives that aim to improve building performance in energy, economic, environmental, and grid support categories. The result of these optimizations, which often utilize several incentive signals as the model input, came to be directly and heavily influenced by these signal profiles. For example, an optimal setpoint control strategy that minimizes the building's utility bill will avoid consuming electricity at hours with higher electricity prices and shift the building's power or energy demand to hours with lower prices. However, depending on whether or not other incentive signals, such as carbon emission rates, share a similar pattern and fluctuates to a high or low point at similar hours, the optimal control strategy that minimizes cost may or may not end up increasing the carbon emission due to the fluctuation of the electricity demand profile. "Similarity" in this case, refers to the "degree of alignment" between two signals at a fixed time. Two sets of signals are considered more "in sync", when they rise and fall at similar times of the day; and are considered "conflicting" when one increases while the other decreases at similar times of the day. In other words, the dynamics of the incentive signal used for one objective, as well as the relation to signals from other objectives—whether the signals are "in sync" or are "conflicting"—heavily influence the tradeoffs that may exist among the various performance objectives.

While the advent of high temporal and spatial resolution incentive signals, often at hourly or sub-hourly level, can promote more tailored optimal control strategies that meet local objectives, it also magnifies the numerous possibilities of control strategies and complicates the tradeoff dynamic between different aspects of building performance. With signals fluctuating hour to hour, creating a broader range of patterns over daily, seasonal, and annual periods, the common workflow where the optimizer only considers incentive signals or scenarios from a "typical day" profile or carries out optimization over a short period of planning horizon will be inadequate in capturing the dynamic control outcomes and tradeoffs that exist among them. Therefore, a thorough understanding of the signal dynamics, in an individual and collective context, becomes an important step in developing more effective optimal control strategies.

Given this context, this work is among the first to examine the dynamics of forward looking incentive signals (end-use pricing signal, average emission rate and marginal emission rate) generated at an hourly level across thirty years for various balancing areas in the United States. The results of this study shed light on the impact that signal dynamics and relationships may have on the tradeoffs between building MPC objectives, as well as the implication on more effective MPC workflows. Section 2 provides additional context and reviews relevant literature. Section 3 describes the unsupervised clustering process, including data description, processing, parametric study and best cluster selection. Results of best clusters of three sets of incentive signals and discussion on their implications to the building performance optimization are presented in Sections 4. Limitations, conclusions and future work are provided in Section 5.

2. RELATED WORK

Defined as the incremental change in carbon emissions incurred by an incremental change in the total electricity generated (or demanded), marginal emission factors (MEFs) have been considered more prominently over the past several years to analyze the environmental impacts of a wide variety of actions that would alter the consumption or supply patterns of electricity. While average emission factors (AEFs) quantify the total carbon emitted from generating electricity, the MEF, which focuses on the marginal generator or technology, directly considers the response of the power system to a change in consumption or generation. Studies by Hawkes (2010) and Siler-Evans et al. (2012) found that due to the often higher intensity of carbon emission of the generators at the margin, MEF is often higher than AEF and more adequate in capturing the dynamic change of carbon emissions.

Many studies have leveraged the advantage of MEFs to investigate the environmental impact of a wide range of objectives and programs. Smith and Hittinger (2019) used marginal emission factors to improve estimates of environmental benefits from appliance efficiency upgrades. Amoroso et al. (2018) utilized the marginal emission factors to examine different sizing options and operating patterns of central air conditioning (AC) systems in the context of its economic and environmental impact. Marginal emission factors were also adopted to evaluate emission benefits of building design implementations, such as increased insulation for new homes (Levy et al., 2016), and the long-term environmental performance of buildings designed to high performance standards (Collinge et al., 2018).

The penetration of renewable energy generation diversified the generation mix, creating unprecedented opportunities for researchers to examine the electricity supply-demand relationship under a new light. In this context, MPC and other smart technologies that provide demand-side flexibility, among many other objectives, adopted the concept of MEF when considering the environmental impact and implications of its technology. Using a simple conceptual model of demand response MEF, a study by McKenna and Darby (2017) argues that demand response technology and smart appliances create large carbon saving potential, especially when coupled with long-term structural change. As MEF signals become available at a high spatial and temporal resolution, studies have noticed the highly volatile nature of the profile and the implication on the environmental impact resulting from MPC load shifting. Zivin et al. (2014) evaluated the environmental impact of electric vehicles (EVs) as a load shifting technology using MEF and found substantial spatial and temporal heterogeneity of marginal emission. Depending on the location and time of the day, charging EVs could lead to drastically different environmental and economic costs. Callaway et al. (2018) also emphasized the importance of location in evaluating the environmental impact of renewable energy technology and efficiency-induced demand reductions, due to the highly fluctuating MEF profile.

While much interest has been given to MEF, studies have not taken the next step of incorporating high resolution MEF as an incentive signal in building MPC. Existing optimization research in building performance, single objective and multi-objective, has been largely focused on the energy, economic and thermal comfort categories. For instance, studies by Yang and Wang (2012) and Missaoui et al. (2014) both focused on a bi-objective optimization of reducing energy cost and maximizing thermal comfort. A few studies considered carbon emissions and energy costs with real-time pricing signals however, did not have high resolution carbon emission data to incorporate marginal effects. Vogler-Finck et al. (2018) used average emission factor as their environmental incentive signal in their single objective MPC optimization. Péan et al. (2019) calculated marginal emission factor for its MPC study from an empirical approach, establishing a regression model of average marginal emission rate as a function of load and the proportion of renewable energy resources, given historic data for Spain. Although the regression model offered improvements over average emission factors, it was still limited in its spatial applicability and could not adapt to an ever-evolving generation mix. The most relevant study was conducted by Knudsen and Petersen (2016), who developed a building MPC framework that considered both real-time pricing signal and carbon intensity (again, using AEF as the incentive signal) in a multi-objective optimization setting. This study was also among the very few that attempted to quantify the tradeoff dynamics between the environmental and economic objective by altering the "added weight" coefficient for the two objectives. The authors alluded to the importance of correlation between the two incentive signals; when they observed a stronger tradeoff between the two objectives the incentive signals were less correlated.

The findings of earlier works have directly motivated the study presented in this paper to more closely analyze grid signal patterns in attempt to better understand the tradeoff dynamics that may exist among building MPC grid incentive signals. To our knowledge, this is the first study to apply unsupervised clustering algorithms to examine patterns of hourly real-time pricing signals and average and marginal emission signals using a detailed dataset with modeled hourly grid operational data that spans a thirty-year period for all balancing areas in the United States for various future grid scenarios. To illustrate the approach and potential implications, we focus on a subset of the data that includes two months of data for one location. This paper explores whether or not recurring patterns exist within the individual grid incentive signals and discusses how the signals may lead to tradeoffs within optimal building control applications.

3. METHODOLOGY

The approach followed in this work is summarized in Figure 1. Descriptions of the data set and assumptions are provided in Section 3.1, data cleaning and preprocessing are included in Section 3.2, and parametric studies for each clustering algorithm and best cluster selection criteria are documented in Section 3.3.

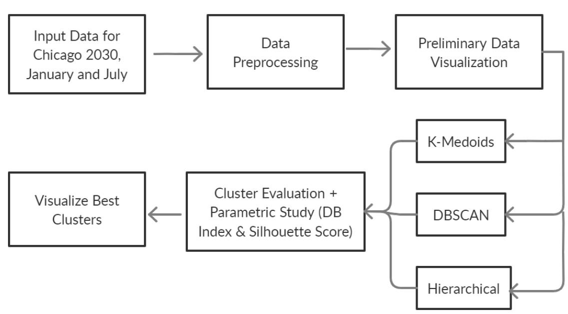

Figure 1: General workflow for data clustering and analysis.

3.1 The Data Set

Data used for this study, including real-time pricing and average and marginal emission factors, were generated by a novel grid operation model, Cambium, developed by the National Renewable Energy Laboratory (NREL) (Gagnon et al., 2020). Built to expand the metrics reported in NREL's Standard Scenarios (Cole et al., 2020), an annually released set of projections of how the U.S. electric sector could evolve across a suite of different potential futures, Cambium assembles structured data sets of simulated hourly cost and operational data for modeled futures of the U.S. electric sector with metrics designed to be useful for long-term decision-making. For each scenario, which outlines a possible realization of a future grid generation mix, Cambium models hourly grid operations for every balancing area in the United States, at two-year intervals until 2050. Spanning over thirty years at unprecedented spatial and temporal resolution, Cambium provides users with the most detailed outlook of plausible electric grid futures. For this study, we selected predictions for the Mid Case in the general Chicago balancing area (p80) in the future year 2030. The Mid Case serves as a baseline among the Standard Scenarios, where the total generation grows steadily over time and is provided primarily by a mix of new natural gas combined cycle (NGCC), PV, and wind generation.

A brief explanation of the three signal types used in this work is provided later in this section. For more detail regarding the Cambium modeling approach and assumptions, readers should refer to the full Cambium documentation (Gagnon et al., 2020). The pricing signal used in this study refers to the total end use marginal cost ($/MWh) that accounts for energy, capacity, operating reserve, and portfolio costs. Costs for distribution capacity, administrative and general expenses, and other electric sector expenses are not included. The average emission factor refers to the average emission rate of the generation induced by a region's end-use load in kilograms of CO_2 emissions per megawatt-hour of end-use load. The marginal emission factor refers to the short-run marginal emission rate for end-use load, which is the emission rate of the marginal generator that serves the marginal increase in load. Cambium identifies the marginal generator or the energy-constrained generator on the margin (storage technology like battery) through a five-step procedure, including identifying balancing areas that share a marginal generator, evaluating non-energy-constrained generators and energy-constrained generators, as well as reevaluating non-energy-constrained generators, with relaxed filters.

3.2 Data Preprocessing

Extreme pricing spikes (mostly within the $1000/MWh to $2000/MWh range) occurring at up to four hours in July are not realistic and do not reflect how pricing signals would be in any real-life electricity market. For this reason, days that contain these extreme pricing spikes were eliminated by applying an upper limit filter of $800/MWh. Any days with a pricing signal higher than the limit will have all 24 data within the day removed. This process eliminated four days from the 365-day data set.

The original pricing and emission signal had different magnitudes and ranges, thus, to compare their temporal dynamics and study their interrelationship on the same scale both signal sets were normalized. The `MinMaxScaler` function from the Scikit-learn library (version 0.24.1) (Pedregosa et al., 2011) was applied in Python Jupyter Notebook to each signal separately over the entire year to normalize the signal between 0 and 1, where 0 represents the lowest value of the signal in the entire year and 1 represents the highest.

For this study, January and July were selected to represent a typical winter and a summer scenario. Compilations of the signals over a 24-hour period for the entire selected month are provided in Figures 2, 3 and 4 to establish a pre-clustering baseline and serve as a visual aid to inform comparisons with cluster patterns.

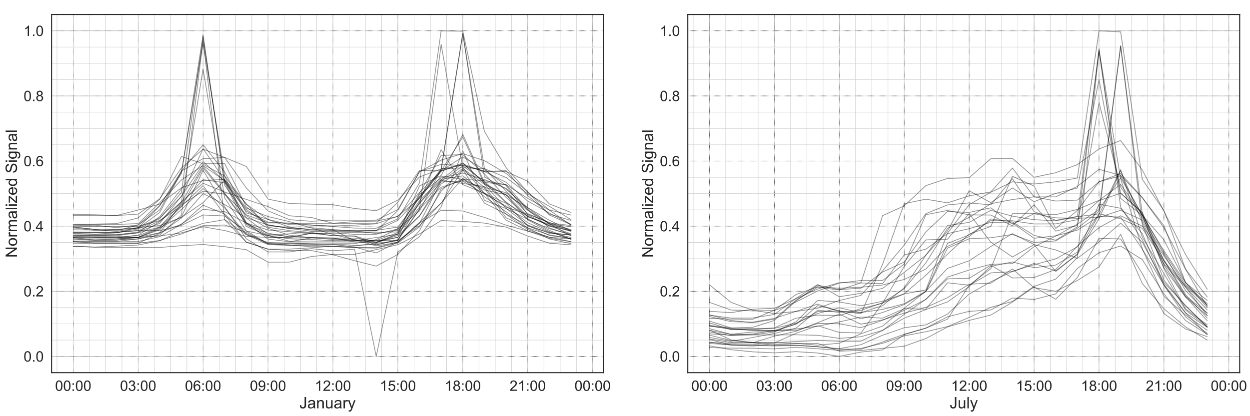

Figure 2: Pricing signal hourly distribution in January and July, 2030.

3.3 Parametric Study and Best Cluster Selection

K-Medoids, DBSCAN and Agglomerative Hierarchical Clustering (HC), with Ward's method and the Euclidean distance metric, were applied to each signal set in the selected months and compared in terms of the "goodness" of the clusters. Three sets of signals per month, for two months, results in six total cases. One best-performing algorithm was selected per case. DB Index and Silhouette Coefficient were used to evaluate results from the clustering procedure and determine the optimal parameter(s) for each algorithm and the best algorithm for each case. Both methods, regularly

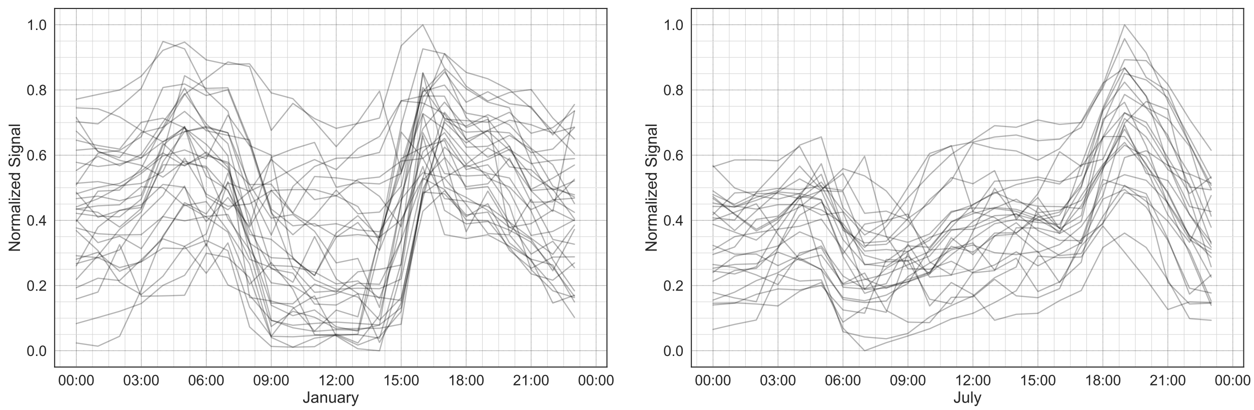

Figure 3: Average emission factor signal hourly distribution in January and July, 2030.

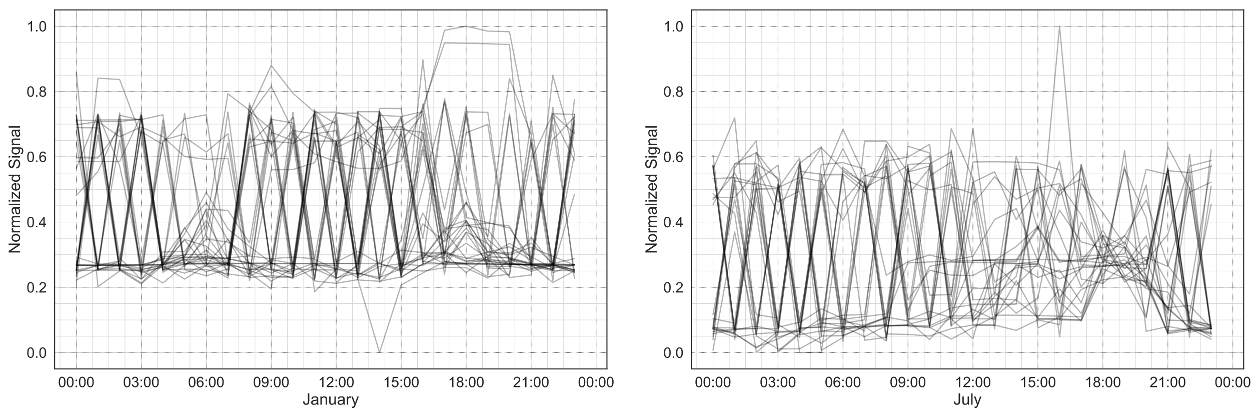

Figure 4: Marginal emission factor signal hourly distribution in January and July, 2030.

used in internal evaluation, establish indices that contrast between high similarity within a cluster and low similarity between clusters. DB Index, or Davies–Bouldin index, is defined as the maximum of the ratios between each intra-cluster distance and inter-cluster distance. A low DB Index indicates a more distinctive clustering process. Similarly, the silhouette value measures how similar data are within its cluster compared to other clusters. A high coefficient indicates a good cluster while low values indicates outliers.

Best clusters were selected based on the results of the parametric study. The parametric study evaluates DB index and Silhouette values for a given range of the key parameter(s) of a method. For K-Medoids and HC, the key parameter evaluated is the number of clusters ranging from 1 to the maximum days within the month. For DBSCAN, the key parameters are epsilon ("eps"), which defines the radius of neighborhood around a point x, and minSample, which is the minimum number of data points within each cluster and defines the minimum size of each cluster. To execute the parametric study, first, key parameters of each algorithm were examined with respect to their cluster verification results using DB index and Silhouette value. Then, parametric results were compared between algorithms to determine a best-performing algorithm for each case. For DBSCAN, heat maps that summarize number of clusters found and number of outliers were compiled first, given a range of eps (0.2 to 1.2) and minSample (1 to max days of the month). A list of combinations of eps and minSample were compiled for all combinations that produced total clusters larger than one and number of outliers fewer than the total days in the month. The list of combinations were then used to repeat the parametric process with DB index and Silhouette score. In many cases, a "perfect" performing set of parameters with the lowest DB index and the highest Silhouette score did not exist. More often, there was a short list of potential parameter values that produced a reasonably good combination of DB index and Silhouette score. Clusters from the short list were generated, visualized and compared. A winner was selected based on visual agreement and the discretion from the authors. Once the algorithm for each case was selected, best clusters were generated and visualized.

4. SIMULATION RESULTS

4.1 Baseline Visualization

Preliminary data assessment involves visualization of the original signals to establish a starting point and visual aid for comparisons of clusters. Figures 2 – 4 reveal the seasonal variations of the three signals and provide a visual comparison of the level of "neatness" among the three. The pricing signal, although having seasonal variation of pattern, displays a clean organization: distinctive patterns can be identified by eye, such as the two-spike profile in the winter, and the one-spike profile in the summer. Such patterns repeat throughout the respective month, with only small hourly variations regarding the exact hour in which the spike(s) occur. In comparison, the marginal carbon emission signal is less structured. Visually, it is difficult to identify any distinctive daily profiles that repeat throughout the month. Moreover, there is a highly fluctuating nature of the signal, where the signal varies from hour to hour in potentially large magnitudes, resulting in the sharp, long spikes that often take place within two hours. Where and how much the signal will fluctuate fully depends on which type of generator is switched onto the margin at that hour, a decision based on cross-regional coordination of the power plants and economic dispatch model that does not have a long-term, intrinsic pattern. The average emission signal is somewhat organized. As expected, the graphs suggest a correlation of average carbon emission and pricing signal, such that the AEF is high when the pricing is high and vice versa. Compared to the pricing signal, the AEF profile is nevertheless more disorganized.

4.2 Parametric Study and Best Cluster Selection

Following the process of selecting a best algorithm and key parameters in Section 3.3, Table 1 summarizes the best algorithm for each case and "goodness" of their performance evaluated by DB index and Silhouette score. As predicted, a "perfect" parameter that guarantees the lowest DB index and the highest Silhouette score within an algorithm rarely existed. For example, in row two of Table 1, which represents the best algorithm (HC) for the pricing signal in July, the winning "number of clusters" selected was 9. Although the DB index (0.75) and Silhouette score (0.3) were not the best within their own category, they presented a favorable combination when evaluated together. Such observations were prevalent throughout the selection process of this study, illustrating a common challenge that arises in evaluating unsupervised learning methods. Overall, Hierarchical Clustering was the best-performing among the three algorithms.

Table 1: Best algorithm for pricing, marginal emission factor, and average emission factor by month.

Signal Type	Month	Best Algorithm	DB Score	Silhouette Score	# Clusters
Pricing	January	DBSCAN	0.51	0.33	7
Pricing	July	Hierarchical	0.75	0.3	9
Carbon (avg)	January	Hierarchical	0.9	0.24	10
Carbon (avg)	July	DBSCAN	0.39	0.12	19
Carbon (m)	January	Hierarchical	0.7	0.09	21
Carbon (m)	July	Hierarchical	0.95	0.107	14

As expected, algorithms found better clusters for the pricing signal, followed by the AEF and MEF. This was demonstrated by the pricing results having the lowest DB index and the highest Silhouette score, followed by AEF and then MEF, which had the highest DB index and lowest Silhouette score. There were also more clusters found for the carbon signals than for the pricing signal. Both findings were not surprising and were consistent with the observations from Section 4.1. Overall, the best clusters found for both AEF and MEF were less helpful, either having numerous visually obvious misfits within every cluster, or having too many clusters that only had one daily profile. This was more so for MEF for which the algorithms struggled to find any plausible patterns and clusters.

4.3 Cluster Visualization

Based on parametric study for the pricing signal, best clusters for the pricing signal in January and July are shown in the first column of Figure 5 and 6. The average emission factor profiles that correspond to the pricing clusters are shown in the middle column, and marginal emission factor profiles are in the right column. Thus, Figure 5 and Figure 6 allow visual comparison of the average and marginal carbon signals, when grouped by similar pricing signals. The best clusters found by the algorithm visually agree with our observations of distinct clusters from the baseline graphs. These clusters sufficiently identified the slight variations within the pricing signal. Within each cluster, the daily profiles adhere to the general shape of the cluster well, with some degrees of over-generalization, such as the second row of Figure 5.

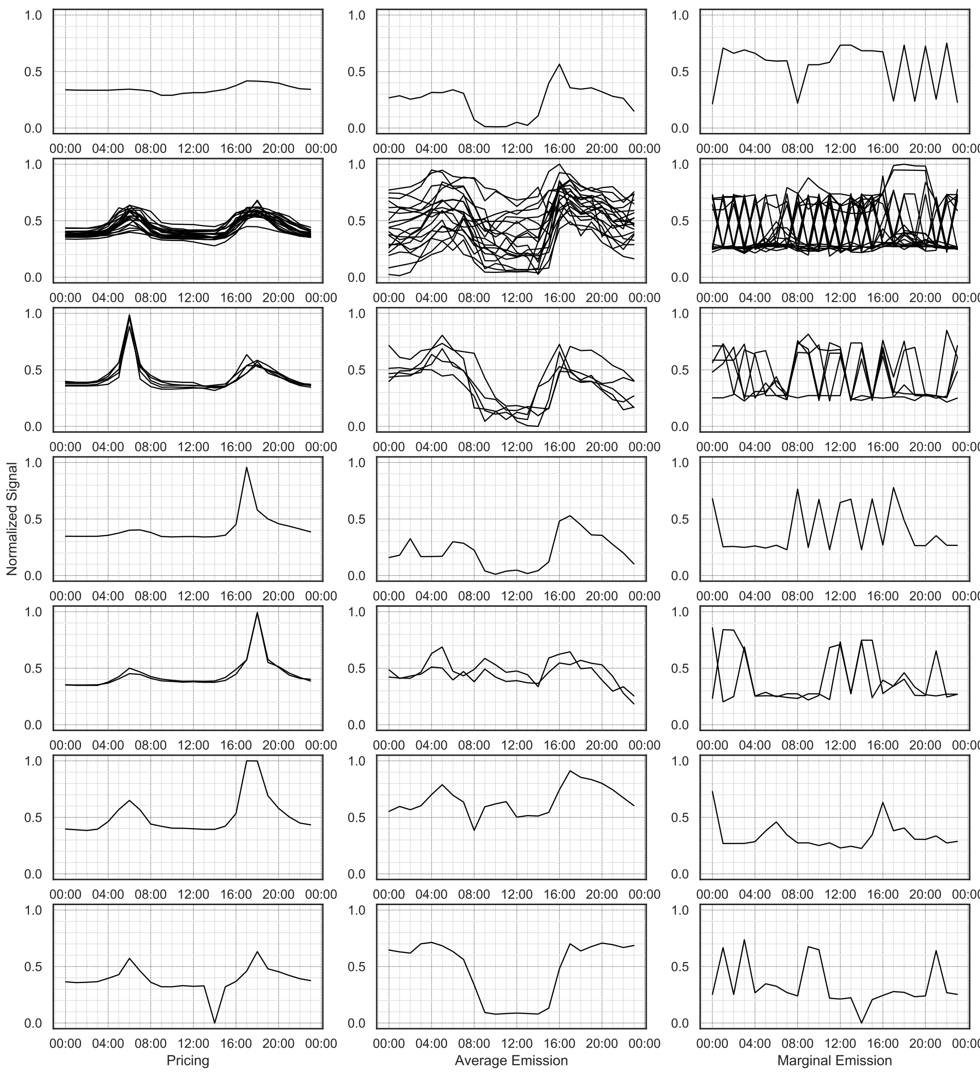

Figure 5: Best clusters for the pricing signal (left column) and corresponding profiles of average emission rate (middle column) and marginal emission rate (right column), for January, 2030.

For every cluster of the pricing signal, there exists a wide and diverse range of signal dynamics from the AEF and MEF. For example, for the two days that have almost identical pricing signals on the fifth row of Figure 5, with the largest deviation between two days around 0.05, the corresponding two AEF profiles fluctuate much more on a day-to-day basis, with the largest hourly difference of 0.2 at 5:00 A.M. The level of fluctuation is even higher for the MEF of the same pricing cluster, with the largest hourly deviation of 0.65 at 1:00 A.M. Further, both AEF and MEF do not share a similar signal synchrony with the pricing signal. While the pricing cluster has a small spike around 6:00 A.M. and a large spike around 5:00 P.M., the AEF slightly fluctuated around 0.5 all day. The MEF displayed a "conflicting" pattern with the pricing signal having spiked at hours where the pricing signal stayed low, and stayed low at hours where the pricing signal spiked. This means, a single-objective MPC that considers these two days within the pricing cluster would have very similar strategies to minimize the energy cost, but that strategy would result in drastically different total carbon emission and would most definitely lead to a tradeoff between energy cost and carbon emission, if the MEF signal were used as the indicator of carbon emission. This example illustrates how the dynamic nature of the emissions profile could lead to unpredictable environmental impacts associated with cost optimal controls.

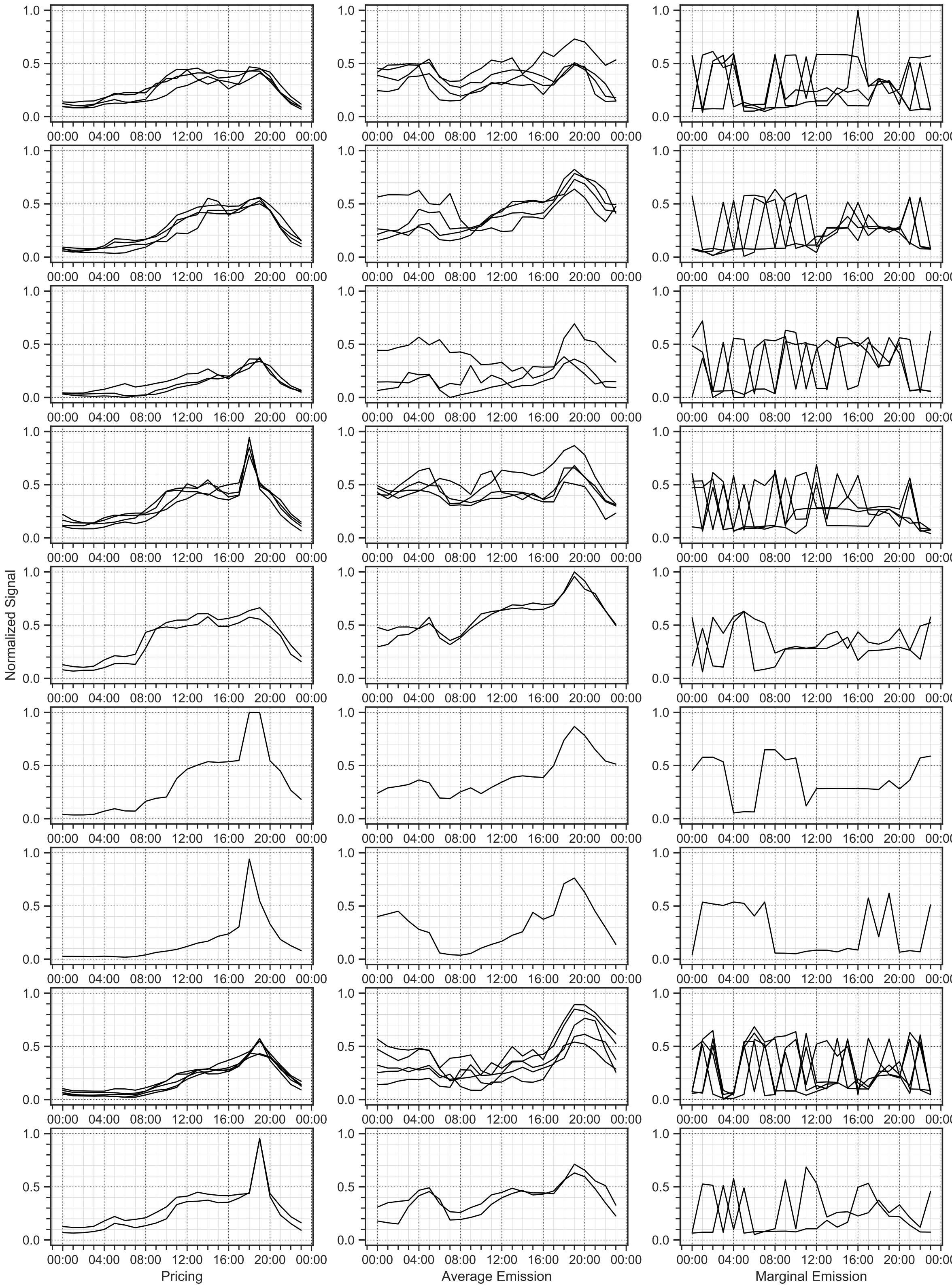

Figure 6: Best clusters for the pricing signal (left column) and corresponding profiles of average emission rate (middle column) and marginal emission rate (right column), for July, 2030.

This study also has implications on how the type of emission signals—AEF or MEF—used in MPC research could affect the tradeoff dynamic between different objectives. Consider the sixth row of Figure 6. The pricing signal and the AEF profile share a good level of agreement and synchrony, with both staying low before 10:00 A.M., increasing from 10:00 A.M. to 4:00 P.M. and hitting a spike at 7:00 P.M. However, the MEF profile of that day is "conflicting" with this pattern, having a spike at 2:00 A.M. and 7:00 A.M., when both the pricing signal and the AEF signal are staying at a constant low level. This means that due to the correlation between the AEF, an optimal MPC strategy that minimizes energy cost is also likely to reduce the carbon emission if AEF is used to quantify the emission rate. However, due to the conflicting pattern of MEF with the pricing signal, if MEF is used, the same strategy that minimizes energy cost would create a tradeoff and be likely to increase the total carbon emission compared to the base case.

Both findings pose suggestions to the current framework of building MPC optimization. Studies that focus on single objective optimizations are encouraged to leverage the dynamics of other incentive signals and the degrees of signal synchrony to estimate potential tradeoff dynamics that could exist among the objectives. Studies that relied on average emission factor or average marginal emission factor also need to recognize the potential of significant hidden tradeoffs that could be revealed by a more detailed data set such as Cambium.

5. CONCLUSIONS

In this study, unsupervised clustering algorithms, K-Medoids, DBSCAN and Hierarchical Clustering were applied to hourly pricing signal, average and marginal emission signal generated by a forward-looking grid model. Overall, Hierarchical Clustering demonstrated the best performance, evaluated by DB index and Silhouette score. While the algorithms did not find distinctive patterns of the carbon signals, it identified 7 and 9 patterns within the January and July pricing signal. The highly fluctuating nature of the carbon emission signals could lead to a diverse range of tradeoffs between energy cost and carbon emission in a building MPC optimization. This finding iterates the importance of understanding incentive signal dynamics, in an individual and collective context, and its impact on the tradeoffs between different objectives in any building MPC optimization research. The findings in this study suggest that, as more high resolution carbon and pricing signals like Cambium became available, using "typical day", AEF or other seasonal or average emission data set will fall short in recognizing the fluctuating and "random" nature of some of these signals that happens from day to day or hour to hour, and cannot fully address the complex tradeoff dynamic as a result of a highly and fast changing reality that is our electric grid.

REFERENCES

Amoroso, B., Hittinger, E., and McConky, K. (2018). Keeping your cool–a multi-stakeholder look at ac sizing. *Building and Environment*, 131:306–329.

Callaway, D. S., Fowlie, M., and McCormick, G. (2018). Location, location, location: The variable value of renewable energy and demand-side efficiency resources. *Journal of the Association of Environmental and Resource Economists*, 5(1):39–75.

Cole, W. J., Corcoran, S., Gates, N., Mai, T. T., and Das, P. (2020). 2020 Standard Scenarios Report: A U.S. Electric Sector Outlook. Technical report, National Renewable Energy Lab.(NREL), Golden, CO (United States).

Collinge, W. O., Rickenbacker, H. J., Landis, A. E., Thiel, C. L., and Bilec, M. M. (2018). Dynamic life cycle assessments of a conventional green building and a net zero energy building: exploration of static, dynamic, attributional, and consequential electricity grid models. *Environmental science & technology*, 52(19):11429–11438.

Delgado, M. G., Ramos, J. S., Domínguez, S. Á., Ríos, J. A. T., and Cabeza, L. F. (2020). Building thermal storage technology: Compensating renewable energy fluctuations. *Journal of Energy Storage*, 27:101147.

Gagnon, P., Frazier, W., Elaine, H., and Cole, W. J. (2020). Cambium documentation: Version 2020. Technical report, National Renewable Energy Lab.(NREL), Golden, CO (United States).

Greensfelder, E. M., Henze, G. P., and Felsmann, C. (2011). An investigation of optimal control of passive building thermal storage with real time pricing. *Journal of Building Performance Simulation*, 4(2):91–104.

Hawkes, A. D. (2010). Estimating marginal CO_2 emissions rates for national electricity systems. *Energy Policy*, 38(10):5977–5987.

Henze, G. P., Felsmann, C., and Knabe, G. (2004). Evaluation of optimal control for active and passive building thermal storage. *International Journal of Thermal Sciences*, 43(2):173–183.

Klaassen, C., House, J. M., and Center, I. E. (2002). Demonstration of load shifting and peak load reduction with control of building thermal mass. *Teaming for Efficiency: Commercial buildings: technologies, design, performance analysis, and building industry trends*, 3:55.

Knudsen, M. D. and Petersen, S. (2016). Demand response potential of model predictive control of space heating based on price and carbon dioxide intensity signals. *Energy and Buildings*, 125:196–204.

Levy, J. I., Woo, M. K., and Tambouret, Y. (2016). Energy savings and emissions reductions associated with increased insulation for new homes in the united states. *Building and Environment*, 96:72–79.

McKenna, E. and Darby, S. J. (2017). How much could domestic demand response technologies reduce CO_2 emissions? *European Council for an Energy-Efficiency Economy*.

Missaoui, R., Joumaa, H., Ploix, S., and Bacha, S. (2014). Managing energy smart homes according to energy prices: analysis of a building energy management system. *Energy and Buildings*, 71:155–167.

Niu, J., Tian, Z., Lu, Y., and Zhao, H. (2019). Flexible dispatch of a building energy system using building thermal storage and battery energy storage. *Applied energy*, 243:274–287.

Pavlak, G. S., Henze, G. P., and Cushing, V. J. (2014). Optimizing commercial building participation in energy and ancillary service markets. *Energy and Buildings*, 81:115–126.

Péan, T., Costa-Castelló, R., and Salom, J. (2019). Price and carbon-based energy flexibility of residential heating and cooling loads using model predictive control. *Sustainable Cities and Society*, 50:101579.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.

Siler-Evans, K., Azevedo, I. L., and Morgan, M. G. (2012). Marginal emissions factors for the U.S. electricity system. *Environmental science & technology*, 46(9):4742–4748.

Smith, C. N. and Hittinger, E. (2019). Using marginal emission factors to improve estimates of emission benefits from appliance efficiency upgrades. *Energy Efficiency*, 12(3):585–600.

Swaminathan, S., Pavlak, G. S., and Freihaut, J. (2020). Sizing and dispatch of an islanded microgrid with energy flexible buildings. *Applied Energy*, 276:115355.

Vogler-Finck, P. J. C., Wisniewski, R., and Popovski, P. (2018). Reducing the carbon footprint of house heating through model predictive control–a simulation study in danish conditions. *Sustainable Cities and Society*, 42:558–573.

Yang, R. and Wang, L. (2012). Multi-objective optimization for decision-making of energy and comfort management in building automation and control. *Sustainable Cities and Society*, 2(1):1–7.

Yu, M. G. and Pavlak, G. S. (2021). Assessing the performance of uncertainty-aware transactive controls for building thermal energy storage systems. *Applied Energy*, 282:116103.

Zivin, J. S. G., Kotchen, M. J., and Mansur, E. T. (2014). Spatial and temporal heterogeneity of marginal emissions: Implications for electric cars and other electricity-shifting policies. *Journal of Economic Behavior & Organization*, 107:248–268.

ACKNOWLEDGEMENT

The authors thank Pieter Gagnon at The National Renewable Energy Laboratory (NREL) for his generous discussion on the Cambium dataset.