Applied Energy 281 (2021) 116074

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy

Water-energy benchmarking and predictive modeling in multi-family residential and non-residential buildings

Matthew Frankel [a,b], Lu Xing [a], Connor Chewning [a], Lina Sela [a,*]

[a] Department of Civil, Architectural and Environmental Engineering, The University of Texas at Austin, United States
[b] Lyndon B. Jonhnson School of Public Affairs, The University of Texas at Austin, United States

HIGHLIGHTS

- Buildings with similar water-energy usage and intensity were identified and clustered.
- Analysis revealed heterogeneity in water and energy consumption patterns of buildings.
- Benchmarking buildings provided a measure of comparison for multi-utility management.
- Data-driven modeling revealed meaningful insights into urban water-energy nexus.

ARTICLE INFO

Keywords:
Water-energy urban modeling
Residential and non-residential buildings
Ensemble clustering
Water and energy intensity benchmarking

ABSTRACT

As the threat of climate change grows alongside a continual increase in urban population, the need to ensure access to water and energy resources becomes more crucial. In the context of the water-energy nexus in urban environments, this work addresses current gaps in understanding of coupled water and energy demand patterns and reveals apparent dissimilarities between utilization of water and energy resources for heterogeneous buildings. This study proposes a data-driven approach to identify fundamental water and energy demand profiles, cluster buildings into groups exhibiting similar water and energy use, and predict their demand. The clustering problem was cast as a two-stage cluster ensemble problem, in which several clustering methods with different settings were employed, and then the results obtained from partial view of the data were combined to achieve consensus among the partitionings. The influential drivers for water and energy consumption were identified, parametric and non-parametric prediction models were developed and compared, utilizing high and low temporal data resolution. The clustering analysis performed in this work revealed that water and energy consumption patterns of heterogeneous buildings are not exclusively characterized by general building characteristics. Analysis of the predictive models showed that an overall non-parametric model provides better predictions for water and energy compared with parametric models and that models with high and low data resolution provide comparable demand predictions. The results of this study highlight the value of data-driven modeling for revealing meaningful insights into usage patterns and benchmarking buildings' performance to provide a meaningful measure of comparison to facilitate multi-utility management. Overall, the methods outlined in this study provide another step towards building greater resiliency within urban areas in preparation for future changes in population and climate.

1. Introduction

Information about consumers' demand for water and energy is becoming increasingly available, detailed, and accurate due to the widespread utilization of high-resolution water and energy meters [1]. Meters collecting and transmitting consumption measurements on a daily, hourly or sub-hourly resolution, generate data, which have extensive uses for management of water and energy resources, including understanding patterns of consumer behavior, detecting abnormal events, and creating demand prediction models [2,3]. Insights into trends in consumer behavior derived from high-resolution water and electricity consumption data are used to inform planning, pricing

* Corresponding author.
E-mail address: linasela@utexas.edu (L. Sela).

https://doi.org/10.1016/j.apenergy.2020.116074
Received 25 April 2020; Received in revised form 8 October 2020; Accepted 16 October 2020

mechanisms, and conservation strategies [4,5,6,7]. With growing populations in urban areas, and arid conditions that are being exacerbated by climate change, water and energy demands must be sustained at a level that can be satisfied by the projected level of production.

In order to achieve the goal of decreased consumer water and energy demands, high-resolution consumption data is instrumental in the design and implementation of demand-side management policies (DSM), which allows policymakers to understand consumer behavior and design policies for incentivizing efficient resource consumption [8]. Policies could take the form of increasing technological efficiencies as well as promoting behavioral shifts in building occupants [9]. In addition, continuous monitoring of water demand has been applied to the detection of irregular usage patterns and leaks along with the modeling of consumer demand [10,11,12,13]. Similarly, electricity meter data is used for demand disaggregation, event detection as well as privacy hedging [14,15,16,17]. Furthermore, weekly and daily consumption patterns can be used for infrastructure planning and expansion. For water systems, timeseries data of consumer consumption are used to obtain accurate information about peak and average demands, which are key for infrastructure sizing (e.g., storage and conveyance capacity) [18,11].

Although typically addressed separately, the majority of previous works related to water and energy consumption data have primarily focused on: (1) extracting characteristic and common usage patterns across end-users using various timeseries clustering techniques (e.g., [19,20,14]); (2) identifying explanatory and influential variables characterizing end-uses and end-users behavior (e.g., [21,22,23,24,25]); and (3) creating demand prediction and event detection models using various parametric and non-parametric techniques (e.g., [26,27]). The explosion of previous works utilizing the available information from advanced meters is also evident from the several recently published review papers in the water and energy sectors [28,29,30,31,32,33].

There is an abundance of different clustering approaches that can be applied to analyze timeseries data of building water and energy consumption, e.g., *k*-means [34], *k*-medoids [35], agglomerative hierarchical [36], self-organizing map [37], spectral clustering [38], and various density-based methods [39,40,41]. The challenge in selecting the most appropriate clustering approach is mainly due to the fact that there are no ground-truth labels for the appropriate clusters. As such, only internal performance evaluation criteria, which are based on the information intrinsic to the data alone, can be used [42]. Often times, this leads to the application of multiple clustering approaches, requiring meticulous and non-trivial parameter tuning, often resulting with no clear advantage of one approach over the other. Ultimately, despite the large body of works, there are no standards for application of cluster analysis for pattern extraction from consumption data in terms of algorithm selection, parameter tuning, number of clusters, and data size [19,43,44].

For demand prediction, parametric and non-parametric models are commonly used [45,46]. Parametric models involve model selection and parameterization, such as a linear, exponential, or polynomial regression. Common non-parametric models include classification and regression trees, artificial neural networks, and random forest [47,48,49,50,51]. Parametric models are advantageous because of model simplicity and interpretability compared with non-parametric models [49]. In non-parametric models, the fundamental structure of the model is determined by the underlying data, which might not be captured by the parametric models. Not only are there many different types of non-parametric models available, but even among the same class of models, the choices made during model implementation, such as the model structure, hyper parameters, loss function, and cross-validation metric affect model performance and prediction accuracy. In addition, model selection depends on the specific application as well as the quality, resolution, type of data available, and the selected features. Similar to clustering, due to the high variability in drivers affecting model performance, there is no consensus on the most appropriate modeling approach for water and energy timeseries prediction. While some works demonstrated that neural networks outperformed multiple linear regression for predicting peak weekly water demand [52], others showed that multiple linear regression, decision trees, and neural network models were comparable for predicting weekly energy consumption [49]. Thus, suggesting that case-specific conditions dictate modeling choice and there is no “one size fits all” model.

The significance of the water-energy nexus has long been recognized and has been studied through multiple lenses, addressing considerations related to policy and water-energy production at regional and national scales [53,54,55,56]. Coupled analysis of water and energy consumption is advantageous, as it provides additional insights to support conservation strategies, design targeted pricing mechanisms, improve performance, or incentivize users to conserve resources [57,1,58]. Unlike previous examples, in the context of the water-energy nexus in urban environments, joint water and energy consumption profiles and demand patterns has not been fully explored.

Among the reasons for disjoint analysis of water and energy trends is the scarcity of spatiotemporally-resolved synchronous water and energy data, as multiple utility service providers are responsible for data collection, management, and ownership [1]. Noticeable exceptions include recent works by [59,57,60]. The authors in [59] investigated water and energy consumption usage patterns and explanatory variables for over 1000 residential users, where 18 unique water and energy usage patterns were identified, with factors including swimming pool or/and a hot tub explaining high water and energy use. The main determinants of water and electricity demand profiles were identified in order to enable targeted demand management recommendations for each user profile. In [60], a classification approach was proposed for identifying water usages of residential users by disaggregating coincident water and electricity data. Results showed improvement in water use classification of clothes washer, dishwasher, and shower events when both water and energy data were included in the analysis [60]. In [57,1], the authors propose an integrated data collection infrastructure, in which utilities collect and analyze building water, electricity, and gas consumption contemporaneously. The proposed benefits of the integrated system include the ability to segment end-users by consumption patterns, and enable real-time and customized conservation targets for individual utility customers.

In urban environments, with the exception of hot-water consumption [61], the drivers for water and energy consumption are fundamentally different, and therefore, efficiency in water consumption does not imply efficiency in energy consumption, and vice versa. Intuitively, building features such as effective insulation, strategically-placed windows, efficient heating, ventilation, and air conditioning systems would make a building more energy efficient, yet would have no effect on water consumption. Conversely, buildings, which exhibit low-flow plumbing fixtures, automatic water shut-offs at sinks, and leak detection capabilities would have no effect on energy consumption. Similar to the intuition on building features, occupant decisions, which have been shown to have an important impact on building energy efficiency, do not necessarily translate to water efficiency [62]. Therefore, given the global stresses in the availability of water and energy resources, and the need to decrease consumption, data-driven management strategies focusing on water and energy consumption are a necessary and effective tool to achieve efficiencies in both resources [57,58].

In light of the scarce prior research, gaps remain in the analysis and understanding of coupled water and energy demand patterns in the urban environment. First, much of the literature on characterizing, clustering, and predicting water consumption patterns revolves around residential users, which do not account for a large portion of water users. In Texas, for example, non-residential users (e.g., commercial, hospitals) account for over half of all water consumed [30,63]. Because of the high degree of heterogeneity among non-residential users, insights that apply specifically to single-family residences do not necessarily apply to non-

residential buildings and multi-family buildings. Second, only a few works have been proposed to characterize the joint water-energy consumption patterns [59,60]. However, the effectiveness of demand management strategies aiming at reducing resource consumption of the end-users strongly depends on the understanding of the drivers affecting water and energy consumption patterns. In this context, this study explores the following questions: (1) are there apparent dissimilarities between utilization of water and energy resources for heterogeneous buildings? (2) if there are dissimilarities in water and energy utilization, can these dissimilarities be captured by general building characteristics? (3) to what extent low-resolution consumption data can provide comparable prediction performance to utilizing high-resolution, continuous data? and (4) to what extent buildings studied in this work are representative of the national consumption benchmarks for water and energy [64–65]?

To address the research questions, this study proposes a data-driven approach to identify fundamental water and energy demand profiles, cluster buildings into groups exhibiting similar water and energy use, and predict their demand. To overcome the abundance of clustering methods and improve robustness, instead of choosing a single clustering method, several methods were employed, and their results were integrated to maximize the average mutual information [66]. Similarly, for prediction, parametric and non-parametric models were developed and their performance was compared and benchmarked against common performance standards. The proposed approach includes three main steps: (1) creating characteristic weekly water and energy consumption profiles, (2) identifying characteristic consumption patterns and cluster buildings exhibiting similar water- and energy-use shape and intensity, and (3) identifying influential drivers for water and energy consumption and compare prediction accuracy of parametric and non-parametric prediction models targeting different building categories and different resolutions of available consumption data. The applications of the proposed approach can be useful for building managers to assess efficiency of building operations, and for demand planners at water and energy utilities to predict future demand requirements.

2. Methods

The proposed approach consists of three main phases. First, characteristic weekly usage patterns for water and energy were identified. Second, the different buildings were clustered into building groups with similar water and energy consumption and intensity patterns. The clustering problem was cast as a two-stage cluster ensemble problem, in which several cluster methods with different hyper-parameters were employed, and the results from the partitions obtained from partial view of the individual metrics were combined to achieve consensus among the partitionings. Third, linear regression and regression tree models were fit to predict daily water and energy consumption and identify influential parameters in model predictions. Finally, clustering and prediction results were analyzed and discussed. Fig. 1 illustrates the main steps of the proposed approach and the details of each step are described next.

2.1. Data description

Daily water and energy usage data measured by continuously-recording meters installed at 70 buildings at the University of Texas at Austin (UT Austin) were utilized for this study. UT Austin is one of the largest universities in the U.S., serving approximately 51,000 students in Fall, 2018 [67]. The data used in this work were collected by the UT utilities to track building water and energy consumption and determine the degree to which buildings are achieving UT Austin's sustainability goal of reducing overall water and energy consumption by 20% in 2020, using 2009 as the basis for comparison [68]. Data available for buildings sporadically spans between 2009 and 2017, with all buildings reporting consumption levels from April 2014 to June 2017. The buildings studied herein, range in size from 20,372 to 69,275 m^2, have between 2 and 38 floors, and were originally built between 1926 and 2012. Based on the UT Austin designation, there are five general categories of building types: classroom and academic (CA), housing (H), office and administration (OA), public assembly and multipurpose (PA), and research laboratory (RL). A summary of the building attributes (e.g., number of buildings, area, year built, mean and standard deviation (s.d.) of the water and energy consumption) is presented in Table 1. Prior to the analysis, datasets were processed to remove erroneous and missing data, by removing any specific datapoints that were zero or negative. In addition, a moving filter was applied to remove datapoints that were more than two standard deviations away from the mean of the previous 25 days. Overall, the amount of data deemed unsuitable, and therefore removed, was negligible.

2.2. Characteristic pattern extraction

The first stage in the analysis entails extracting the characteristic consumption pattern for each building, which will subsequently be used to cluster the buildings. Several metrics can be used to characterize building efficiency, such as average annual consumption, annual water use intensity and energy use intensity [69,70,71]. The resolution at which a particular metric is assessed depends upon the desired outcomes of the analysis. In cases of long-range planning, monthly timescales are sufficient [26]. However, given that, the objective in this paper is to identify buildings that function similarly in terms of water and energy consumption, a more granular timescale is necessary. Hence, daily consumption data are used, and weekly patterns are selected as the characteristic feature of each building, where each weekly pattern was characterized using its intensity and variability. Water and energy intensity (i.e., WI and EI, respectively), defined as the daily water or energy consumption per building area, were used to obtain information regarding the magnitude of resource consumption. In order to obtain information regarding the variation of daily consumption over the course of a week, the normalized water and energy data (i.e., WN and EN, respectively) were used. For each building, the original

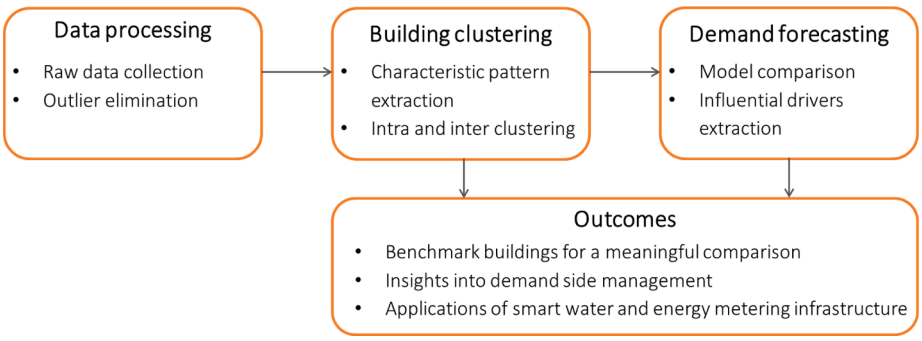

Fig. 1. Overview of the proposed approach.

Table 1
Summary of building characteristics and water/energy consumption by building designation.

Building Type	Classroom and Academic (CA)	Housing (H)	Office and Administration (OA)	Public Assembly and Multipurpose (PA)	Research Laboratory (RL)
Number of buildings	28	8	3	18	13
Area (m^2) [min; mean; max]	2419; 12526; 26,213	3683; 16029; 69275	5583; 7723; 9580	1893; 16783; 46789	4883; 18192; 39823
Year Built [min; mean; max]	1926; 1966; 2012	1933; 1953; 2000	1986; 1999; 2008	1930; 1975; 2010	1930; 1975; 2008
Number of Floors [min; mean; max]	2; 6; 9	4; 8; 28	3; 5; 7	2; 6; 11	5; 8; 15
Water Consumption (l/day) [mean; s.d.]	8873; 6088	73520; 47997	10233; 7100	16824; 10994	22901; 10404
Energy Consumption (kWh/day) [mean; s.d.]	3420; 465	4183; 679	1933; 269	5305; 776	11889; 861

consumption was scaled such that the normalized demand for water and energy ranged between 0 and 1, such that 1 represents the maximum daily consumption and 0 the minimum daily consumption of a building during the observation period. In summary, the normalized consumption constitutes an *internal index*, which is based on the information intrinsic to the building alone and indicates the temporary consumption level of a building compared to the entire observation period. Complementary, consumption intensity constitutes an *external index*, which indicates the consumption level of a specific building compared to the other buildings relevant to the specific case study. Consumption intensity additionally allows to benchmark building performance relative to national averages of different building types [64–65]. It follows that four metrics (WI, WN, EI, EN) were used to benchmark resource use of a building and to compare with other buildings.

Having established the metrics, each building can be represented using four characteristic patterns of intensity and shape (variability) of weekly water and energy consumption. The shape of a pattern was determined by comparing the weekday (Monday-Friday) consumption levels to the weekend (Saturday-Sunday) consumption levels. Pattern shape was classified into two categories, i.e., FLAT and CURVED, depending on the amount of variability between weekday and weekend consumption. One-sided Wilcoxon rank-sum statistical test was utilized to determine whether the difference between weekend and weekday consumption was statistically significant [72]. In this setting, the null hypothesis states that the difference between weekday and weekend consumption is statistically insignificant, i.e., FLAT pattern, and the alternative hypothesis states that the weekend consumption is significantly lower than the weekday consumption, i.e., CURVED pattern. The null hypothesis was rejected at small p (<0.05) values.

The method for classifying the magnitude of the normalized water and energy patterns differed slightly. Normalized water and energy patterns were classified as HIGH, MEDIUM, or LOW, based on if the median of the normalized weekly pattern was in the lower, middle, or upper tercile of the data, respectively. A similar process was used to classify weekly patterns of water and energy intensity; however, instead of the HIGH, MEDIUM, or LOW classification based on normalized terciles, the thresholds were based on the terciles of all water and energy intensity values across all buildings. The outcome of this process is a reduced representation of the raw timeseries water and energy data for each building.

After each weekly pattern of a building has been designated into one of six categories based on its classification of intensity and shape, a single, unique characteristic pattern of each building that will be used for further analysis was determined using the following steps. First, weekly patterns were selected from the category that contained the most patterns. For example, if 60% of the patterns of a particular building and metric were classified as HIGH-FLAT, then all the high-flat patterns would be considered to create the characteristic weekly pattern for that metric. Then, the characteristic weekly pattern of a building was determined as the median of all the selected patterns belonging to the same category. Continuing the previous example, for a particular building, the Monday value of the characteristic pattern would be the median of the Monday values of the HIGH-FLAT patterns. An example of the weekly pattern extraction process is demonstrated in the results section.

2.3. Water-energy clustering

The process of water-energy clustering was completed in two steps. First, buildings were *intra*-clustered based on similar usage patterns for each individual metric, resulting in four different *sub-clusterings* obtained from partial view of the data (i.e., individual metrics). In this step, instead of choosing and fine-tuning a single cluster method, several cluster methods with different hyper-parameters were employed, and then their results were integrated to achieve robustness [66]. Second, in order to integrate the intra-clustering partitions and identify common water and energy consumption patterns, *inter*-clustering (i.e., *meta-clustering*) was performed to group buildings exhibiting similar consumption patterns across all four metrics. Inter-clustering was cast as a cluster ensemble problem of combining the partitions obtained from partial view of the individual metrics. The following sections present the details of the proposed clustering procedure.

2.3.1. Robust intra-clustering

The goal of intra-clustering is to divide buildings into partitions based on their similarities in terms of a single individual metric, i.e., WI, EI, WN, or EN, such that buildings in the same cluster/partition exhibit a similar characteristic weekly pattern. For this task, many clustering algorithms can be applied, such as *k*-means [34], *k*-medoids [35], agglomerative hierarchical [36], self-organizing map [37], spectral clustering [38], and various density-based methods [39,40,41]. However, as mentioned previously, selecting the right clustering approach is challenging due to the variety of algorithms, hyper-parameter tuning, similarity/dissimilarly distance measures, as well as the intrinsic randomness in many of the clustering techniques making the results sensitive to the initialization of the algorithms. The process of algorithm selection and parameter tuning for evaluating the quality of clustering is not-trivial given that ground-truth labels, or even the number of clusters, are unknown. Additionally, the performance of clustering algorithms depends tremendously on the structure of the input data. Hence, clustering results can vary significantly with different algorithms and parameter settings, thus weakening the robustness and accuracy of the clustering results.

To improve the robustness of the results and avoid the task of algorithm selection and parameter fitting, the proposed approach relies on employing several clustering methods with different parameter settings and combines the results such that a consensus among the different methods is achieved [66]. Specifically, consider a *clusterer* as a single clustering algorithm with specific parameter settings, and a *partitioning* as clusterer output, i.e., a set of labels indicating the partitioning of objects to clusters. In this work, we applied seven different clusterers employing different algorithms, parameters, and number of clusters, as

will be described in detail in Section 3.2. *Cluster ensemble* is then defined as the problem of combining the multiple partitionings of objects obtained by different clusterers, without accessing the underlying data used to create the original partitions [66]. The main motivation behind cluster ensemble is that the combined partition exhibits improved accuracy and robustness by integrating the information exploited by different clusterers.

The cluster ensemble problem for partitioning buildings based on their similarities in terms of a single individual metric is applied as follows. Given r clusterers and n objects (i.e., buildings), let the *partition matrix* $\boldsymbol{P} = (\boldsymbol{p}_1, \boldsymbol{p}_2, \cdots, \boldsymbol{p}_r)^T$ denote the partitionings obtained from the r different clusterers, where $\boldsymbol{p}_i = (p_{i1}, p_{i2}, \cdots, p_{in})$ is a vector representing the partitioning results of the i^{th} clusterer. Thus, $\boldsymbol{P}$ is a $r \times n$ matrix, where p_{ij} represents the cluster label for the j^{th} object obtained from the i^{th} clusterer. A *consensus function* then maps the partition matrix ($\boldsymbol{P}$) to an integrated consensus partition vector ($\boldsymbol{c}$), i.e., $\boldsymbol{P}^{r\times n} \rightarrow \boldsymbol{c}^{1\times n}$, which is a combined partitioning that achieves the most consensus among all clusterers. Without the presence of a priori information about the ground-truth labels, the objective of the cluster ensemble problem is to identify the optimal consensus function, such that the consensus partitioning ($\boldsymbol{c}$) shares the most information with the individual partitionings, as presented in the partition matrix ($\boldsymbol{P}$).

To evaluate the quality of the consensus, it is necessary to define a metric that quantifies the amount of information shared between two different partitionings. In this study, average normalized mutual information (ANMI) is adopted to quantify the statistical information shared between the original parititioning matrix, $\boldsymbol{P}$, and the consensus partitioning, $\boldsymbol{c}$, [66]. Normalized mutual information is a symmetric information-theoretic metric, conveniently ranging from 0 to 1. Intuitively, when the amount of information shared between two clusterers is low, the normalized mutual information approaches 0, and, contrary approaches 1, when the amount of information shared between two clusterers is high. The ANMI then averages the amount of mutual information between each individual partitioning, $\boldsymbol{p}_i$, and the consensus paritioning, $\boldsymbol{c}$. The ANMI can be computed as follows:

$$ANMI(\boldsymbol{P}, \boldsymbol{c}) = \frac{1}{r}\sum_{i=1}^{r} \frac{I(\boldsymbol{p}_i, \boldsymbol{c})}{\sqrt{H(\boldsymbol{p}_i)H(\boldsymbol{c})}}$$

where $I(\boldsymbol{p}_i, \boldsymbol{c})$ denotes the mutual information between the partition obtained from the i^{th} clusterer and the integrated partition [73], and $H(\boldsymbol{p}_i), H(\boldsymbol{c})$ denote the entropy of the individual and consensus partitionings, respectively [74]. ANMI is bounded by 0, which indicates no mutual information between the individual and consensus partitionings, and 1, representing a perfect agreement between the individual and consensus partitionings. ANMI can be estimated by the sampled quantities provided by the partitionings. The estimation is elaborated in text S1 in the Supporting Information (SI).

Illustrative example: To illustrate the problem of cluster ensemble, consider a simple example where four clusterers ($r = 4$) are applied on a set of six objects ($n = 6$), i.e., buildings in the context of this paper. The following partitioning vectors ($\boldsymbol{p}_1, \boldsymbol{p}_2, \boldsymbol{p}_3, \boldsymbol{p}_4$) specify the partitionings obtained from four clusterers for six buildings:

$\boldsymbol{p}_1 = (1,1,2,2,3,3);$
$\boldsymbol{p}_2 = (2,2,3,3,1,1);$
$\boldsymbol{p}_3 = (1,1,1,2,2,2);$
$\boldsymbol{p}_4 = (1,1,2,2,2,2);$

The label vectors indicate that the first and second clusterers yield three clusters, while the third and fourth clusterers produce two clusters. Further inspection reveals that the partitions $\boldsymbol{p}_1$ and $\boldsymbol{p}_2$ are logically identical since the partition structure is invariant to the numbering of the labels. Additionally, $\boldsymbol{p}_3$ and $\boldsymbol{p}_4$ introduce some dispute regarding the third and the fourth buildings. Intuitively, the objective of the cluster ensemble problem is to find an integrated partitioning vector ($\boldsymbol{c}$), which shares as much information as possible with the individual partitionings ($\boldsymbol{p}_i$), as indicated by a high ANMI score. However, even in this simple example the solution of the best combined partition is not trivial due to the fact that the number of clusters and each cluster's interpretation vary significantly among the different clusterers.

In this study, the cluster-based similarity partitioning approach (CSPA) [66] is adopted to solve the cluster ensemble problem. CSPA involves a two-step approach, in which: (1) a combined similarity matrix is constructed that summarizes all partitionings, and (2) the objects are re-clustered using the combined similarity matrix. First, the original partitioning of each clusterer ($\boldsymbol{p}_i$) can be represented as a binary similarity matrix where two buildings have a similarity of 1 if they belong to the same cluster and 0 otherwise. Subsequently, the combined similarity matrix $\boldsymbol{S}^{n\times n}$ can be obtained by adding all the binary similarity matrices obtained from the r clusterers. The s_{ij} element of the combined similarity matrix represents the number of partitions in which buildings i and j belong to the same cluster. Fig. 2(a) illustrates the combined similarity matrix for the illustrative example. The elements s_{ij} in the combined similarity matrix can range between 0 and the number of clusterers ($r = 4$), where 0 (white) indicates that the corresponding buildings i and j are not clustered together in any of the partitionings, and 4 (black) indicates that buildings i and j are assigned to the same cluster based on all the 4 partitionings. For example, buildings 1 and 2 are assigned to the same cluster based on all four partitionings $\boldsymbol{p}_1, \boldsymbol{p}_2, \boldsymbol{p}_3, \boldsymbol{p}_4$, i.e., $s_{12} = s_{21} = 4$. On the other hand, buildings 2 and 3, belong to the same cluster based only one clusterer, hence, $s_{23} = s_{32} = 1$.

After the combined similarity matrix $\boldsymbol{S}$ is obtained, the next step is to re-cluster the buildings based on $\boldsymbol{S}$, without revisiting the data used to create the original partitions. Spectral clustering [38] is applied to re-cluster the buildings, where the optimal number of clusters was determined by varying the number of clusters and selecting the number that yields the highest ANMI score. Spectral clustering is a well-known partitioning method that relies on a similarity matrix between data points. The objective of the spectral clustering is to partition the data into k clusters, such that the cost of the weights in the similarity matrix between the clusters is minimized. The spectral clustering approach is very efficient for small number of clusters, which is the case in this setting. Fig. 2 (b) shows the values of ANMI as a function of the number of clusters, where $k = 3$ results in the higher ANMI. The corresponding consensus partitioning of the six buildings for the illustrative example is $\boldsymbol{c} = (1,1,2,2,3,3)$.

In summary, the intra-clustering procedure described above involves applying an ensemble of different clustering algorithms to each building and each metric, and finding a partitioning that achieves the best consensus among the different clustering algorithms. The intra-clustering procedure results in four integrated partitioning vectors, $\boldsymbol{c}^j = \left(c_1^j, c_2^j, \cdots, c_n^j\right)$, where $j \in (WI, EI, WN, EN)$, and c_i^j represents the label of the partitioning of building i based on metric j. The next step in the proposed approach is to group the buildings based on all four metrics.

2.3.2. Metric-integrated inter-clustering

Although individual intra-clustering is advantageous, the overall goal is to determine which buildings exhibit similar resource utilization patterns across water and energy intensity, and normalized water and energy consumption. Therefore, in the next step, buildings that behave similarly based on the all metrics are grouped into same meta-cluster. The main idea for grouping buildings into meta-clusters is based on consolidating and integrating the four partitioning vectors, $\boldsymbol{c}^{WI}, \boldsymbol{c}^{EI}, \boldsymbol{c}^{WN}$, and $\boldsymbol{c}^{EN}$, into a single partitioning, $\boldsymbol{c}^M = \left(c_1^M, c_2^M, \cdots, c_n^M\right)$.

As previously described, the integration of the partitions based on individual metrics and the determination of the meta-clusters can be cast as a cluster ensemble problem, in which the objective is to combine the multiple partitionings to achieve consensus among the different metrics, without accessing the underlying data used to create the original partitions [66]. This problem can again be solved using the CSPA for cluster

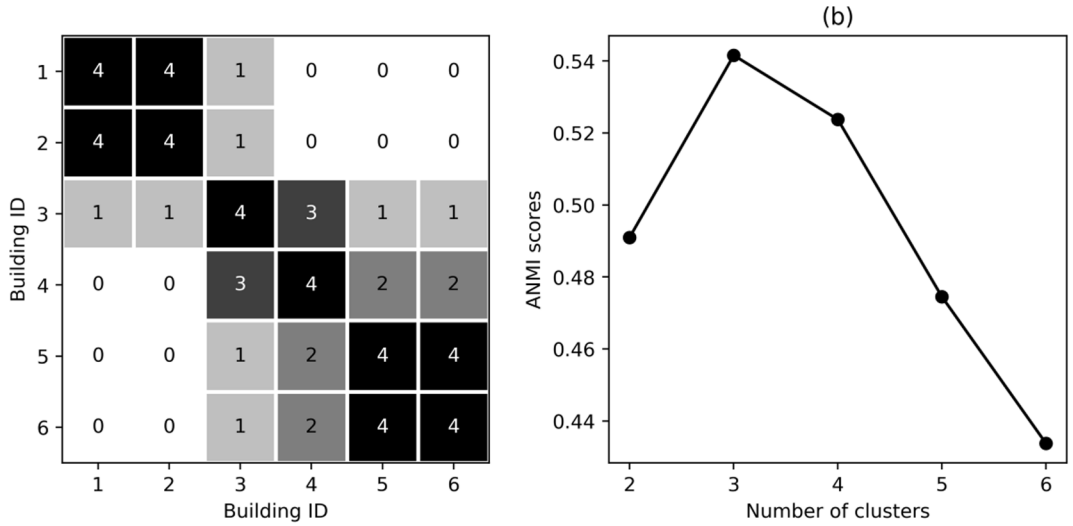

Fig. 2. CSPA for the illustrative example problem: (a) combined similarity matrix; (b) ANMI scores as a function of the number of clusters.

ensemble. In the context of inter-clustering, the partition matrix can be formulated that summarizes the four different partitions according to each metric, i.e., $\boldsymbol{P} = \left(\boldsymbol{c}^{WI}, \boldsymbol{c}^{EI}, \boldsymbol{c}^{WN}, \boldsymbol{c}^{EN}\right)^T$. The corresponding combined similarity matrix $\boldsymbol{S}^{n\times n}$ can be obtained from the four paritionings, where the s_{ij} element represents the number of metrics based on which buildings i and j belong to the same cluster. Thus, the elements in the similarity matrix are bounded by 0, indicating that the two buildings do not belong to the same cluster in any metric, and the number of metrics (4), which suggests that the building pair belongs to the same cluster in all four metrics.

Subsequently, with the similarity matrix as input, any clustering algorithm, such as the spectral clustering introduced in the previous section, can be applied to solve the inter-clustering problem. However, in this section, agglomerative hierarchical clustering (AHC) with complete linkage is chosen to perform inter-clustering on the buildings due to its capability of summarizing the entire hierarchical structure of the dataset. AHC starts by treating each building as a singleton cluster and then builds nested clusters by successively merging pairs of clusters until all buildings have been merged into a single cluster containing all buildings [75]. The outcome of AHC can be visualized using a dendrogram, i.e., a tree-like representation, which summarizes the nested clusters and the corresponding similarity levels. The advantage of using the AHC approach for finding meta-clusters is that it allows control of the granularity of the meta-clusters based on the decision maker's tolerance to the number of different intra-cluster labels. Furthermore, this approach avoids the need to use external evaluation indices, which are often not indicative of the decision-makers notion of good clustering [75]. The application of AHC for meta-clustering is illustrated as follows.

Consider four different partitionings of the buildings resulted from the inter-clustering process based on the four metrics, i.e., $\boldsymbol{c}^{WI}, \boldsymbol{c}^{EI}, \boldsymbol{c}^{WN}$, $\boldsymbol{c}^{EN}$, described in the previous section. For illustration purposes consider that $\boldsymbol{c}^{WI} = (1,1,2,2,3,3)$; $\boldsymbol{c}^{EI} = (2,2,3,3,1,1)$; $\boldsymbol{c}^{WN} = (1,1,1,2,2,2)$; $\boldsymbol{c}^{EN} = (1,1,2,2,2,2)$. Fig. 3 shows the corresponding dendrogram for this example, where the y-axis demonstrates the number of different intra-cluster labels, and the x-axis lists building labels (note that the buildings are ordered by similarity). Fig. 3 illustrates that that buildings 1 and 2, as well as buildings 5 and 6, belong to the same intra-cluster in all four partitions, i.e., the number of different intra-cluster labels equals zero, and thus are the first to be grouped into meta-clusters. Building 3 and 4 have mixed labels based on the four metrics, hence their grouping will depend on the level of tolerance of the decision maker. Three potential cuts are also shown in Fig. 3: Cut 1 implies zero tolerance for different intra-cluster labels, thus resulting in four meta-clusters,

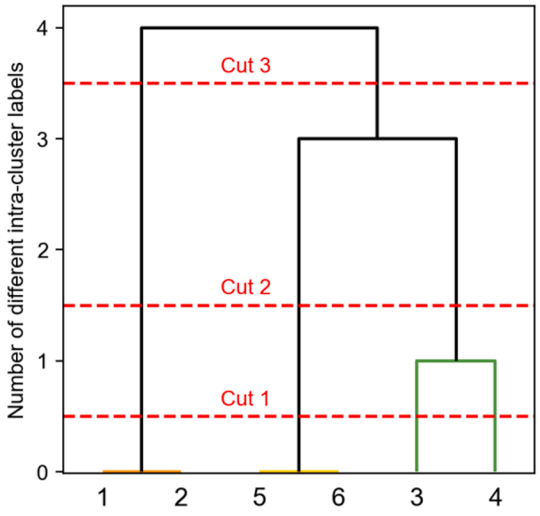

Fig. 3. Inter-cluster dendrogram for the illustrative example with different cuts.

$\{(1,2),(5,6),(3),(4)\}$; Cut 2 implies that at most one different intra-cluster label can be tolerated for buildings to be grouped into the same meta-cluster, thus yielding three meta-clusters, $\{(1,2),(5,6),(3,4)\}$; and Cut 3 increases the tolerance to three labels, which then decreases the number of meta-clusters to two, $\{(1,2),(3,4,5,6)\}$. This approach generates an explicit tradeoff between the number of meta-clusters and the level discord that the decision maker is willing to tolerate.

2.4. Prediction models

After the buildings have been clustered based on their characteristic consumption patterns, the clustering was leveraged to predict daily building water and energy consumption using varying prediction techniques. The following three questions were explored in regards to the creation and evaluation of prediction models: (1) whether meta-cluster-specific models will provide a better prediction than a single overall model; (2) if low-resolution consumption data (e.g., from monthly billing) can provide comparable prediction performance to utilizing high resolution, continuous data; and, (3) if a linear parametric model will provide comparable performance to a non-parametric model.

In order to determine whether an overall prediction model adequately describes usage patterns (i.e. one prediction model for all), or if multiple prediction models, segmented by clustered-buildings (i.e., distinct prediction model for each cluster), would increase the prediction accuracy, a linear regression model was created for each of the meta-clusters, as well as for all of the data. The prediction accuracy was then estimated and compared. An assessment of models based on low-resolution data verses high-resolution data, as raised in the second question, is important due to the two potential applications of prediction models, e.g., to detect anomalous behavior when measured water and energy consumption varies significantly from the predicted values [18], and to be used for planning and demand management [46]. In the first case, the specific building being monitored would require a meter to monitor day-to-day usage levels. In the second case, the buildings in question may not be equipped with a continuously-recording meter, and perhaps not yet built. Therefore, in all model formulations, two cases were considered, one assuming that the building had a meter continuously monitoring usage, in which the previous day's and week's usage data were used as inputs to the prediction models, and another, in which only monthly median consumption was considered, simulating the information of a monthly consumption bill. To address the third question, least absolute shrinkage and selection operator (LASSO) regression [76] and bootstrap-aggregated decision trees (TBAG) were used to predict water and energy usage, providing a comparison between a linear parametric model and a non-parametric model [77,78].

For all models, the response variable (y_i) is the daily water or energy demand at day t for a given building i, which is considered a function of a set of M predictors $\boldsymbol{x}_i = (x_{i,1}, \cdots, x_{i,M})$ that represent the characteristics of the building and the environmental conditions, such as building size and age, outdoor temperature and academic season. The considered features are summarized in Table 2. LASSO regression optimizes for model accuracy and sparsity [76] by minimizing the sum of the squared errors between the predicted and observed daily consumption and by minimizing the number of the estimation coefficients to prevent overfitting [79]. The objective function used by the LASSO algorithm is: $min_{\beta_0\beta}\left(\frac{1}{2N}\sum_{i=1}^{N}(y_i-\beta_0-\boldsymbol{x}_i^T\boldsymbol{\beta})^2+\lambda\sum_{j=1}^{M}|\beta_j|\right)$, where N is the number of observations, β_0 and $\boldsymbol{\beta}$ are regression coefficients, scalar and a vector of length M, respectively, and the parameter λ controls the tradeoff between model accuracy and sparsity. To tune λ, a range of values was generated by enumeration, and the value of λ that resulted in the minimum mean squared error was selected.

TBAG is an ensemble of R decision trees using bootstrap samples of the data, i.e., $T_1(X_1), T_2(X_2), \cdots, T_R(X_R)$, where X_i is a matrix containing the bootstrap sampling features. The results from all trees are then aggregated into a single model, i.e., $T_{BAG}(X) = \frac{1}{R}\sum_{i=1}^{R} T_i(X_i)$, which provides the final output as the unweighted average output of all trees.

Table 2
Input variables used for prediction models.

Model Input	Variable Type	Description
Previous day	Continuous	Daily consumption in the previous day
Previous week	Continuous	Daily consumption in the previous week
Monthly median	Continuous	Monthly median of consumption
Average temperature	Continuous	Daily average temperature [85]
Building area	Continuous	Total building area in square meters
Class session	Categorical	Whether class is in session or not
Day type	Categorical	Weekday or weekend
Season	Categorical	Season of the year
Use designation	Categorical	Building use designation (as originally assigned by UT facilities)
Building name	Categorical	Building label
Year built	Discrete	Year building originally built
Number of floors	Discrete	Number of floors in building

The advantage of regression trees compared to linear models is that instead of fitting the entire data set with a single model, closely related associated data are modeled separately. In addition, based on the successive partitioning of input data, regression tree models illuminate explanatory variables which have high importance on model outcomes [80]. TBAG has also been shown to reduce model over-fitting to the training data [81], resulting in better performance when tested on out-of-sample data. Non-parametric models typically comprise many hyper-parameters (e.g., number of trees and maximum tree depth), which makes tuning complicated. Herein, given the performance of the models, hyper-parameters remained constant.

In summary, for water and energy, two LASSO models were developed for individual building meta-clusters, and two for the entire data, with and without previous day and week consumption; and two TBAG models were developed for all the buildings with and without previous day and week consumption, the rest of the input parameters for all models were the same (i.e., six different models predicting the daily consumption (y_i) of each building).

2.4.1. Model evaluation

In order to assess the performance of each model and select the best-performing model, three different error metrics were employed: coefficient of variation of the root-mean-square error (CV_{RMSE}), median normalized absolute residual (MNAR), and normalized mean bias error (NMBE). CV_{RMSE} and NMBE are standard metrics used to assess prediction accuracy of building energy usage by the American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE), and have been employed in many previous works for assessing model performance [82,83]. The MNAR is an error metric, which was crafted for this specific context, as a more robust version of the NMBE, as described below.

The CV_{RMSE} is a measure of the overall variation between measured and modeled timeseries data. The CV_{RMSE} is calculated by taking the square root of the mean squared error (RMSE) between daily observed consumption (y_i) and model prediction ($\hat{y}_i$), and divided by the mean of the measured data ($\bar{y}$) and multiplied by 100 to be expressed as a non-dimensional quantity in (%). The CV_{RMSE} is defined as (where n denotes the number of observations) [82]:

$$CV_{RMSE} = \frac{1}{\bar{y}}\sqrt{\frac{\sum\left(y_i-\hat{y}_i\right)^2}{n}} \times 100(\%)$$

Values of the CV_{RMSE} can only be positive, where a lower value indicates a better model fit.

The NMBE is a measure of the mean of the error associated with each predicted daily value, calculated as the sum of the difference between error between daily observed consumption and model prediction devided by the sum of the daily observed consumption, and multiplied by 100 to be expressed as a percentage (%). The NMBE is defined as [82]:

$$NMBE = \frac{\sum(y_i-\hat{y}_i)}{\sum y_i} \times 100(\%)$$

The values of the NMBE can be negative or positive. A positive value of NMBE indicates that overall the model overpredicts the observed values, and a negative value of NMBE indicates that overall the model underpredicts the observed values. As noted in previous literature, the value of the NMBE is subject to the cancelation effect, whereby the sum of positive and negative model errors reduce the overall value of the NMBE [83]. Therefore, in this study, NMBE metric is used to ensure adherence to the ASHRAE guidelines for energy prediction, but not to evaluate or compare model performance.

The MNAR is introduced to provide a robust alternative to NMBE and compliment CV_{RMSE}. The MNAR quantifies the median absolute residual of each data measured and predicted daily value, defined as:

$$\text{MNAR} = median\left(\frac{|y_i - \widehat{y}_i|}{y_i}\right) \times 100(\%)$$

Values of the MNAR can only be positive, where a lower value indicates a better model fit. Unlike the NMBE, the MNAR is not subject to the cancelation effect because the absolute value is applied to the difference between observed and predicted value. In addition, both the NMBE and CV_{RMSE} rely on taking the mean of model errors, which may be sensitive to outliers. The MNAR is more robust to outliers by quantifying the median model error instead of the mean.

To ensure compliance with industry and research standards, the energy predictions from selected models were compared to ASHRAE guidelines. For monthly predictions, ASHRAE considers a model calibrated if the CV_{RMSE} is less than 15%, and the NMBE is within ± 5%. For hourly predictions, ASHRAE considers a model calibrated if the CV_{RMSE} is less than 30%, and the NMBE is within ± 10%. Following the example of [84], the guidelines for hourly data were applied to the daily prediction values in this paper, since there is no ASHRAE guideline for daily predictions. The results will demonstrate that the developed prediction models outperform the hourly AHSRAE standards. Because of the noted cancelation effect which plagues the NMBE, the NMBE was utilized to check model accuracy against ASHRAE guidelines, but not to select the best prediction model. In summary, the CV_{RMSE} and MNAR were used to assess and compare the accuracy of all models for predicting daily water and energy usage, and ASHRAE standards were used to benchmark performance of energy prediction models.

3. Application

The proposed clustering and demand prediction approaches were tested using the available water and energy consumption data of 70 buildings on UT Austin campus. For the clustering analysis, data was used for all buildings during the year of 2015. The year of 2015 was used for the clustering analysis because one year is generally representative of the building's behavior. For the prediction analysis, all of the available water usage was utilized, for which the availability varies between 2009 and 2017, where 70% of the data was randomly selected for model training and 30% for testing. Overall, data from 70 buildings were analyzed, with a total of 183,976 datapoints for energy and 187,064 datapoints for water usage.

3.1. Characteristic pattern extraction

Following the methods section, weekly patterns of daily consumptions were first analyzed for four metrics, WI, EI, WN, and EN. Then, each weekly pattern was classified into six categories, based on the shape and magnitude of the weekly pattern, i.e. FLAT or CURVED, and HIGH, MEDIUM, or LOW, respectively. Fig. 4 shows a one-year timeseries of water (top) and electricity (bottom) use intensity for Brackenridge Hall, a dormitory building, with each week color-coded based on the classification of the shape and magnitude of the pattern. Fig. 4 shows that for the majority of the year, daily water intensity was measured between 4 and 8 l/m^2, categorized as HIGH-FLAT, and HIGH-CURVED with the exception of mid-May to mid-August. Energy intensity remained relatively constant, with values between approximately 0.17 and 0.20 kWh/m^2, except for a decrease from about mid-May to mid-June, and an increase from August to September. It is hypothesized that the reduction in water intensity occurred due to the reduction in building occupancy, as many of the students vacate the dormitories for the summer. Similarly, large fluctuations in water intensity also occurred in mid-March and the end of November during the Spring and Thanksgiving breaks. Both water and energy intensity show a decrease during mid-March (which corresponds to the week of spring break and lower building occupancy). However, the major deviations in water consumption do not align with similar changes in energy consumption, an observation which will be further quantified and explored in the discussion of results.

The weekly classifications of water and energy use intensity are

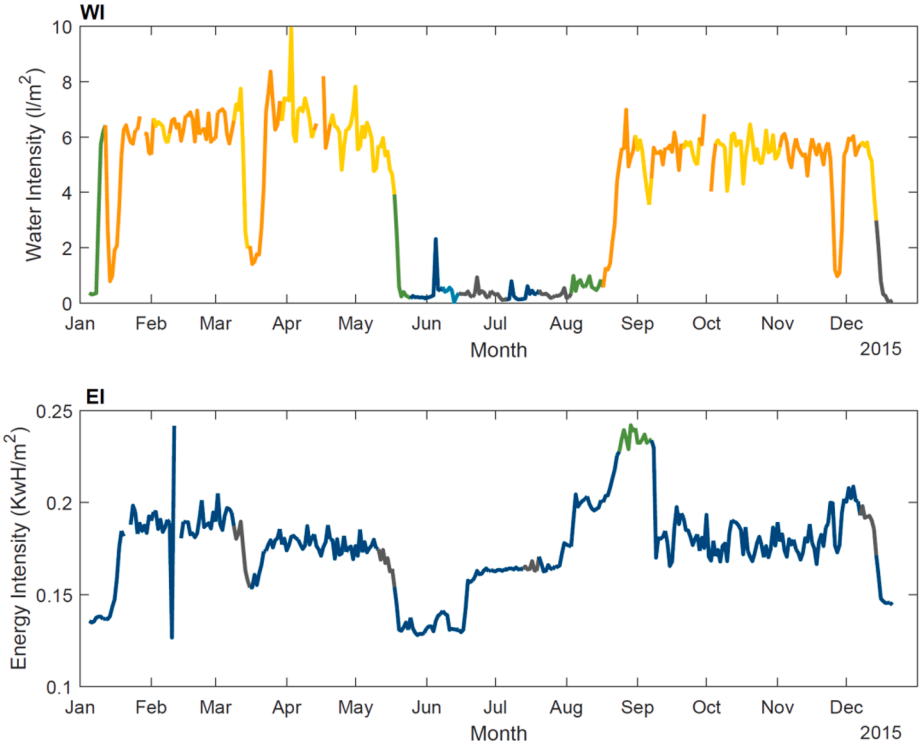

Fig. 4. Timeseries of water and energy intensity for Brackenridge Hall dormitory. Weeks are color-coded based on the classification of shape and magnitude: LOW-FLAT (blue), LOW-CURVED (gray), MEDIUM-FLAT (green), MEDIUM-CURVED (aqua), HIGH-FLAT (orange), HIGH-CURVED (yellow). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

shown in Fig. 5. In the presented example, a contrast between water intensity and energy intensity is observed. While water intensity shows greater variability over the course of a year, energy intensity is classified as LOW-FLAT for all but six weeks during 2015. The variation in usage classifications between water and intensity implies the differences between water and energy use drivers. In the case of the dormitory building examined in this example, both water and energy exhibit flat patterns, signifying that there is no significant variation between weekend and weekday consumption patterns. However, dormitory buildings were found to be higher users of water, but lower users of energy compared with other buildings on campus.

In order to extract the characteristic weekly patterns that will be used for clustering, the median daily values among all of the weeks with the most frequent pattern in each building were computed. For example, HIGH-FLAT patterns were used to compute the characteristic weekly WI pattern for the Brackenridge Hall, as shown in Fig. 6. Similarly, median daily values among all of the weeks with the LOW-FLAT patterns were computed and used as the characteristic weekly EI pattern.

3.2. Water-energy clustering

After the characteristic weekly patterns were extracted for each metric and building, the buildings were intra-clustered based on each individual metric (i.e., WI, EI, WN, EN), using the robust ensemble clustering method, as described in Section 2.3.1. For each metric, seven clusterers were first constructed: k-means with $k = 2$ and $k = 3$, k-medoids with $k = 2$ and $k = 3$, and agglomerative hierarchical clustering with ward, average, and complete linkage with $k = 3$. Euclidian distance was used to compute the similarity between weekly patterns in all clustering algorithms. An example demonstrating the partitionings for EN from these seven individual clusterers, i.e., the partition matrix $\boldsymbol{P}^{7\times 70}$, are summarized in Fig. 7. Each row represents the partitioning from one clusterer, and each column characterizes the cluster labels for one building obtained from different clusterers. Buildings belonging to the same cluster are shaded with the same color. Fig. 7 demonstrates that some buildings are consistently clustered into the same cluster by the different clusterers, while others are grouped differently among the clusterers. Similar results were obtained for all metrics. The next step is to integrate the results of the individual clusterers into a single partition that agrees the most between all of the clusterers and achieves the highest ANMI score.

Subsequently, spectral clustering was applied to integrate the individual partitionings into an ensemble partitioning. The ANMI score was evaluated for different number of clusters, as shown in Fig. 8, where (a) indicates that for WI and EI, the optimal number of clusters is 3, and (b) shows that the optimal ANMI is achieved with 2 clusters for NW and NE. It is observed that the optimal consensus for NW and NE are achieved with only 2 clusters, although five out of the seven clusterers were defined with $k = 3$. The reduction in number of clusters in consensus partitionings indicates that the weekly patterns of NW and NE cannot be meaningfully partitioned into three clusters, and ensemble cluster combined with ANMI score can automatically reveal the most appropriate number of clusters and the corresponding partitionings.

Finally, the results of intra-clustering of buildings and the characteristic patterns are shown in Fig. 9. Thick lines represent the medoid of a particular cluster, demonstrating the typical pattern of a cluster, and thin lines of similar color represent individual buildings that belong to the same cluster.

For the purpose of discussion for the remainder of this paper, the clusters in each metric are referred to in ascending order. For example, for WN, buildings in cluster WN1 (shown in orange in Fig. 9) maintain a low consumption throughout the week, buildings in cluster WN2 (shown in blue in Fig. 9) exhibit high consumption during the weekdays and lower consumption during the weekends, and metrics with 3 clusters, EI3 and WI3, are shown in green in Fig. 9.

Having clustered the buildings based on the individual metrics, i.e., WN, WI, EN, and EI, the next step entails grouping the buildings across all performance metrics into meta-clusters, $\boldsymbol{c}^M = \left(c_1^M, c_2^M, \cdots, c_n^M\right)$ using AHC, as described in Section 2.3.2. The resulting dendrogram is shown in Fig. 10, where the x-axis represents the buildings and the y-axis represents the number of different intra-cluster labels. Four possible paritionings, visualized by the corresponding cuts of the dendrogram, are attained: Cut 1 indicates that all intra-cluster labels achieve complete consensus, thus resulting in 21 meta-clusters; Cut 2 indicates that at most one different intra-cluster label can be tolerated for buildings grouped into the same meta-cluster, thus identifying 10 meta-clusters; Cut 3 increases the tolerance to two different intra-cluster labels, which then decreases the number of meta-clusters to 6; and Cut 4 further raises the tolerance to three different intra-cluster labels, yielding 3 meta-clusters. Either of these meta-clusters may be preferred by the decision maker depending on the specific task. Cut 2 with 10 meta-clusters, as denoted by different colors, is selected to perform the following analysis.

Selected results of the meta-clustering are illustrated in Fig. 11, demonstrating how buildings belonging to MC2 (highlighted in yellow) behave across all four metrics. Fig. 11 displays similar information as Fig. 9, except that the buildings in MC2 are highlighted to show their behavior among all four metrics. In this example, the majority of buildings in MC2 belong to the lowest cluster of metrics, i.e., WN1, WI1, EN1, and EI1. Similar figures depicting all other meta-clusters are found in Figures S1-S9 in the SI.

The results of the inter-clustering and the distribution of energy and water consumption patterns exhibited by buildings in each meta-cluster are summarized in Table 3. The values in Table 3 show the size of each meta-cluster and the fraction of buildings in a specific inter-cluster relative to the size of the meta-cluster, i.e., individual cluster purities [42]. For example, all 14 buildings in MC2 were originally classified into clusters WN1, EN1, and WI1. For EI, 9 of 14 buildings were classified to cluster EI1 and 5 of 14 buildings were classified to cluster EI2. The individual cluster purities can then be used to evaluate the overall purity of the meta-clustering by computing the weighted average of the individual purities [42]. The overall purity of the resulting 10 meta-clusters

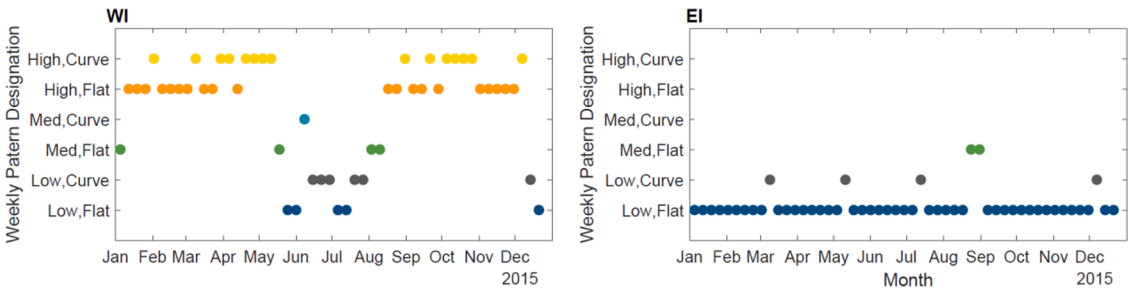

Fig. 5. Classification of weekly water (left) and energy (right) intensity patterns for Brackenridge Hall dormitory. Weeks are color-coded based on the classification of shape and magnitude: LOW-FLAT (blue), LOW-CURVED (gray), MEDIUM-FLAT (green), MEDIUM-CURVED (aqua), HIGH-FLAT (orange), HIGH-CURVED (yellow). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

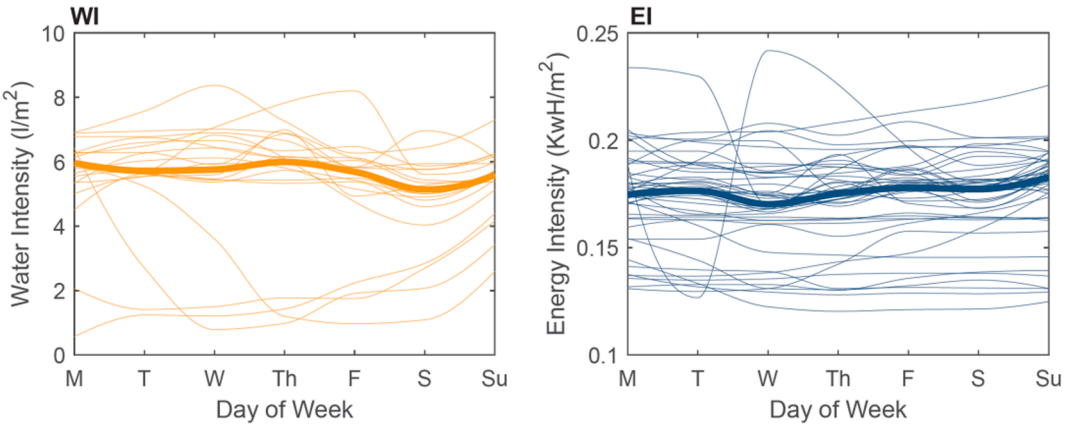

Fig. 6. Frequent (thin lines) and characteristic (thick line) weekly patterns for water intensity (left) and energy intensity (right) of Brackenridge Hall dormitory. Weeks are color-coded based on the classification of shape and magnitude (shown in Fig. 5): HIGH-FLAT (orange) and LOW-FLAT (blue). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

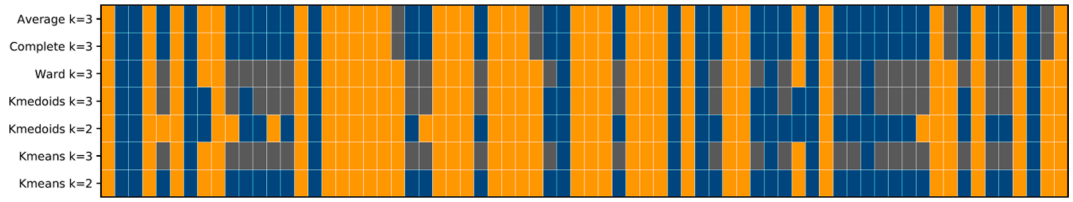

Fig. 7. Clustering results from seven clusterers for normalized energy. Buildings belonging to the same cluster are shaded with the same color. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

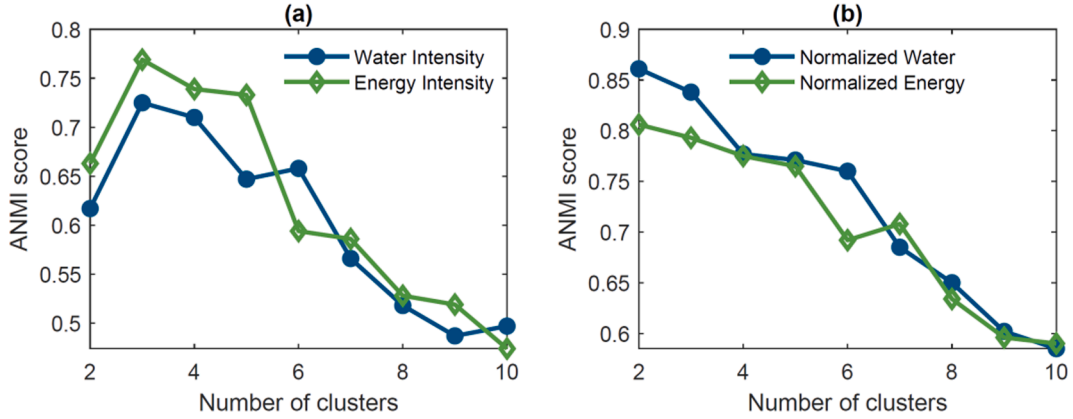

Fig. 8. ANMI score of robust intra-clustering at different number of ensemble clusters.

is 0.91, compared with 0.84 and 1.00, with 6 and 21 meta-clusters, respectively.

The distribution of energy and water consumption intensity and patterns exhibited by buildings in each meta-cluster are visualized in Fig. 12 (a) and (b). In Fig. 12 (a), the x- and y-axis are divided into three regions, each of which represents an intra-cluster of WI and EI, respectively. Similarly, the axes in Fig. 12 (b) are divided into two regions corresponding to the two intra-clusters of WN and EN. The colored markers represent the meta-clusters (with the number representing the meta-cluster label), the locations of the markers represent the corresponding intra-cluster partitioning, and the size of the markers reflect the size of the meta-clusters, i.e., the number of buildings in the meta-cluster. For example, meta-cluster 1, containing 22 buildings, is represented by one orange circle located in the lower-left phase in Fig. 12 (a), revealing that all buildings in meta-cluster 1 are classified into WI1 and EI1. Fig. 12 (b), on the other hand, shows that 6 out of the 22 buildings are located in the WN2 -EN1 quadrant, while the other 16 buildings belong to the WN2-EN2 quadrant, which indicates heterogeneity in EN in meta-cluster 1. Fig. 12 demonstrates that water and energy consumption trends, in terms of shape and intensity, vary significantly across buildings and have mixed correlation in terms of water and energy usage. For example, some buildings are high water and energy users, while others are high water users but low energy users. These mixed trends reveal apparent dissimilarities between utilization of water and energy resources in heterogeneous buildings.

3.3. *Water-energy demand prediction*

Next, LASSO, and TBAG were utilized in order to determine the most appropriate model for predicting daily consumption. As described in Section 2.4, LASSO models were developed for each meta-cluster as well as a single overall model including all the buildings; TBAG models that

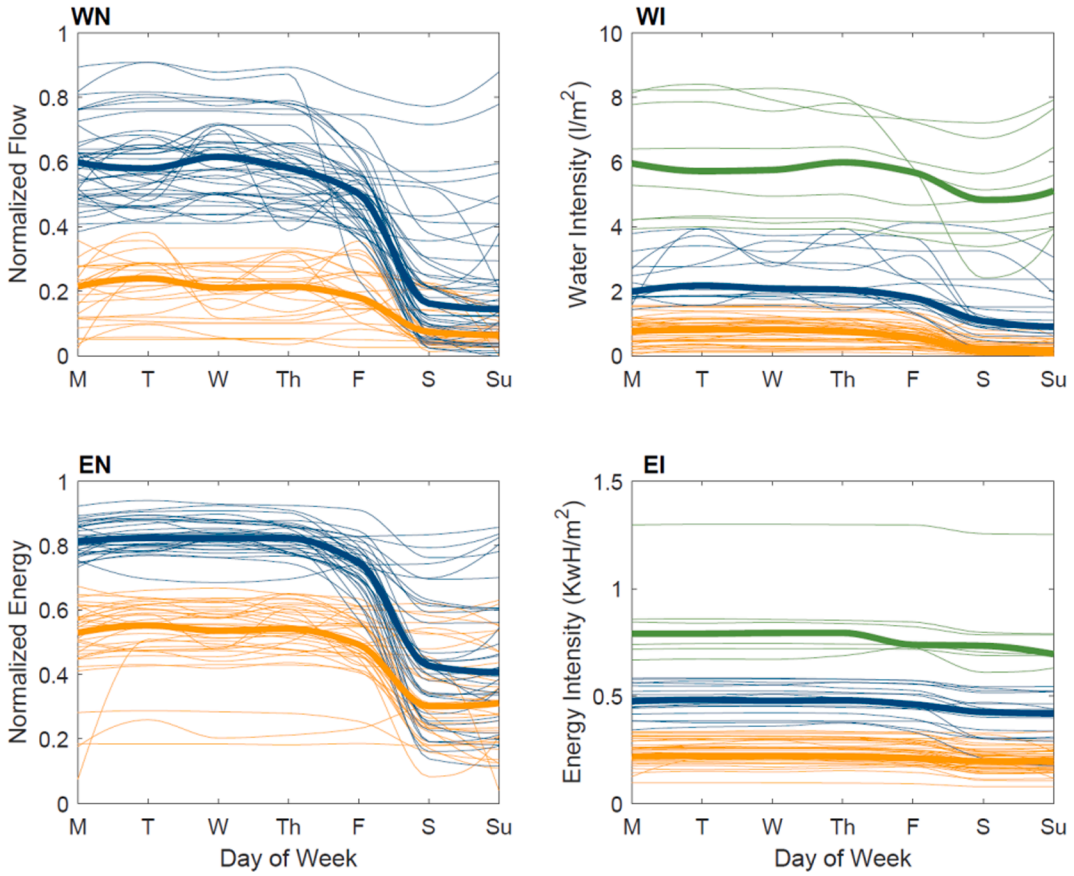

Fig. 9. Results of inter-clustering for all buildings and each metric. Thick lines represent the centroid of a particular cluster and thin lines of similar color represent individual building that were classified into the same cluster. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

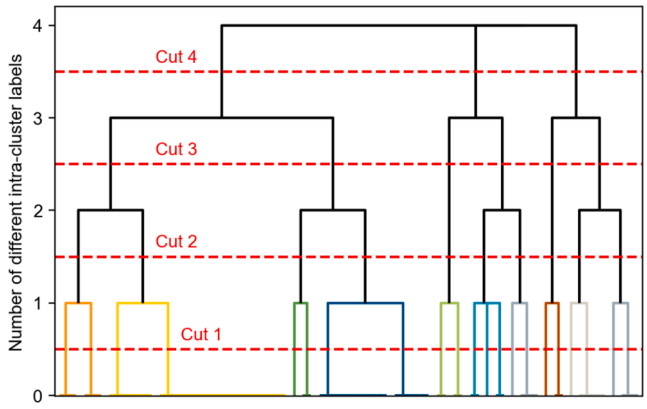

Fig. 10. Inter-clustering dendrogram with potential cuts.

included all the buildings were developed. For each model, two scenarios were created, with and without the inclusion of the water or energy usage from the previous day and previous week. The two scenarios enable the evaluation of the compromise in prediction accuracy as a function of data availability. As noted previously, the motivation behind creating models with and without the information about recent water and energy usage is that this daily and weekly usage may not be readily available for many of the existing buildings that employ mechanical meters, which typically provide only a monthly reading having to be read manually by utility personnel.

Model errors were calculated for each building, and the median model error among buildings in the same meta-cluster was calculated. A summary of the results of the prediction models is shown in Tables 4 and 5. Table 4 shows the median cross-validated root mean squared error (CV_{RMSE}) and median normalized absolute residual (MNAR) between the observed and predicted energy usage for each model for meta-cluster 2 for training and testing data. As seen in Table 4, the TBAG models outperformed the LASSO models for energy prediction in meta-cluster 2 with and without including energy usage in previous day and week. TBAG, with the inclusion of previous data, performs similarly for energy prediction across all other meta-clusters, with values of CV_{RMSE} and MNAR ranging between 3.9% and 10.8%, and 1.8% to 4.4%, respectively. The TBAG models also outperformed LASSO models for water prediction, however with higher modeling error compared with modeling error for energy prediction. Error values for water prediction, as seen in Table S2, of CV_{RMSE} and MNAR ranged between 23.3% and 40.0%, and 11.9% to 23.3%, respectively, among test datasets, for the TBAG model with the inclusion of previous data. A plausible explanation for the decreased model performance for water consumption compared to energy consumption is that there is more variability in daily water than energy consumption, leading to decreased model performance.

Table 5 summarizes the performance of all six models among all meta-clusters, by showing the rank of model performance for each meta-cluster, based on the number of times a model underperformed compared to all other models based on CV_{RMSE} and MNAR. Low rankings represent better performing models and high rankings represent worse performing models.

Based on Table 5 and Tables S1 and S2 in the SI, it is observed that the TBAG model with the inclusion lags is the most accurate model, with the lowest values of CV_{RMSE} and MNAR across all meta-clusters for both water and energy prediction. TBAG model also outperforms LASSO models even when previous day's and week's usage is not included as

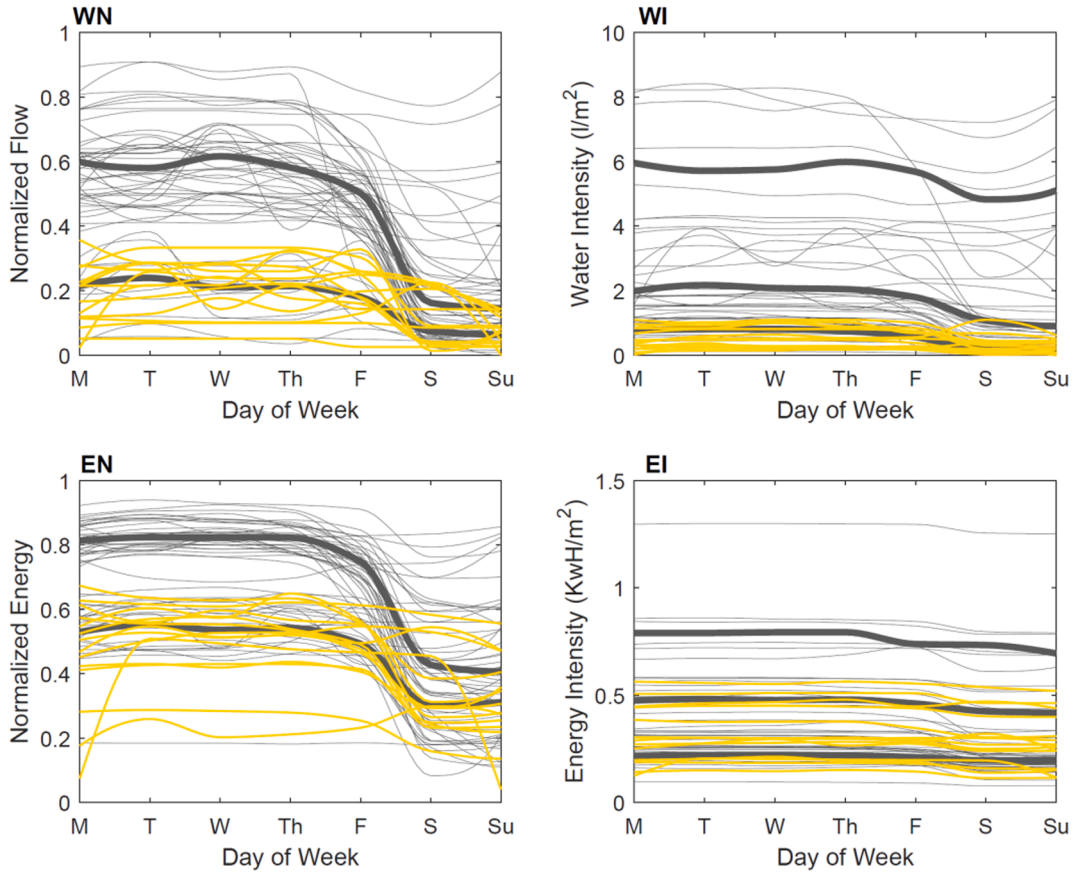

Fig. 11. Characteristic patterns of buildings belonging to MC2 are highlighted in yellow. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 3
Summary of meta-clusters

Metacluster	Number of Buildings	Proportion of Buildings Within Each Cluster									
		Normalized Water Consumption		Normalized Energy Consumption		Water Intensity			Energy Intensity		
		WN1	WN2	EN1	EN2	WI1	WI2	WI3	EI1	EI2	EI3
MC1	22	0.0	1.0	0.3	0.7	1.0	0.0	0.0	1.0	0.0	0.0
MC2	14	1.0	0.0	1.0	0.0	1.0	0.0	0.0	0.6	0.4	0.0
MC3	6	0.0	1.0	0.5	0.5	1.0	0.0	0.0	0.0	1.0	0.0
MC4	5	0.0	1.0	1.0	0.0	0.0	1.0	0.0	0.2	0.4	0.4
MC5	5	0.0	1.0	0.0	1.0	0.0	0.2	0.8	1.0	0.0	0.0
MC6	4	0.5	0.5	1.0	0.0	0.0	0.0	1.0	1.0	0.0	0.0
MC7	4	0.0	1.0	0.0	1.0	0.0	0.8[+]	0.3[+]	0.0	0.0	1.0
MC8	4	1.0	0.0	1.0	0.0	0.0	1.0	0.0	0.8[+]	0.0	0.3[+]
MC9	3	0.7	0.3	0.0	1.0	0.0	1.0	0.0	0.0	1.0	0.0
MC10	3	1.0	0.0	0.0	1.0	1.0	0.0	0.0	0.7	0.3	0.0

[+] Sum of fractions equals one, but not apparent with 2 significant digits.

input variables. When comparing the prediction accuracy of energy across all buildings, the median value of CV_{RMSE} for the TBAG model with previous information was 5.7% (for the test dataset) and increased to 7.4% when previous usage was not included. For water prediction, the median value of CV_{RMSE} for the TBAG model with previous information was 28.7%, and increased to 34.1% when previous usage was not included. Therefore, depending on the intended use of prediction models, models without an input from continuous-recording meters may be suitable. As seen from Table S2 in the SI, all energy prediction models meet the ASHRAE guidelines of CV_{RMSE} less than 30%. For TBAG models, NMBE values for each building range between −0.59% and 0.09% for models with previous data, and −0.72% and 0.36% without, thus, meeting the NMBE standards of ± 10% for energy prediction models.

4. Discussion

Next, the composition of building use designations within each meta-cluster is analyzed. Fig. 13 shows the distribution of building types (as assigned by UT utilities, i.e., CA, H, OA, PA, RL) among each meta-cluster.

Information presented in Figs. 12, 13 and S1-S9 in the SI are useful tools to identify buildings for which efficiencies could be realized. Meta-clusters 5 and 6, which contain all of the university housing buildings, exhibit normalized usage patterns with little variability between weekdays and weekends consumption and high water use intensity (as

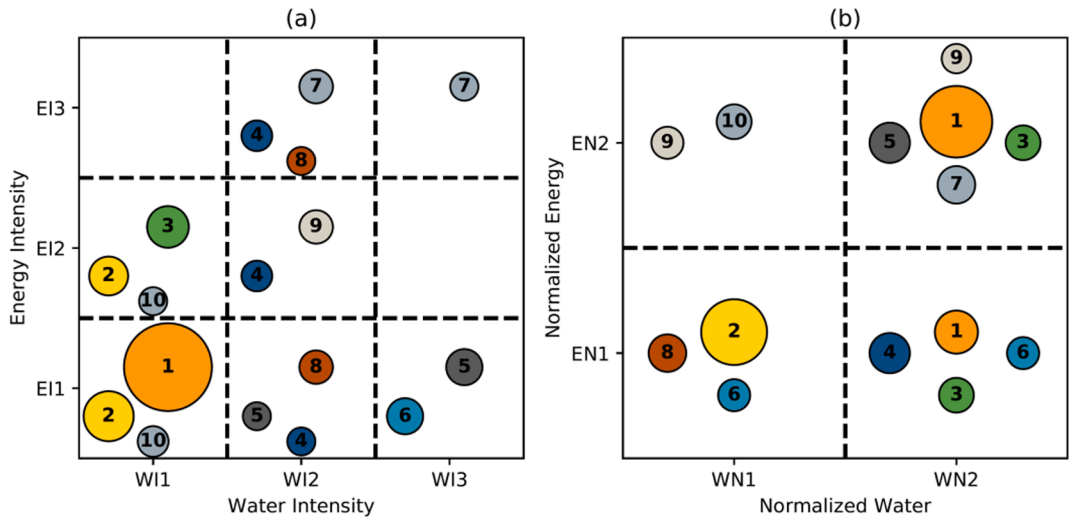

Fig. 12. Distribution of energy and water consumption intensity (a) and patterns (b). The size and the number of each marker represent the size and the label of each meta-cluster (Table 3), respectively.

Table 4
Summary of model performance for energy prediction in meta-cluster 2.

Data Used	Model Type	Training		Testing	
		CV_{RMSE} %	MNAR %	CV_{RMSE} %	MNAR %
Cluster-Specific	LASSO^{+}	7.5	4.2	7.2	4.3
	LASSO^{-}	12.7	6.2	10.0	6.6
Overall	LASSO^{+}	12.0	5.4	8.6	5.6
	TBAG^{+}	5.0	2.4	5.3	2.6
Overall	LASSO^{-}	15.0	6.9	11.7	7.3
	TBAG^{-}	6.4	3.1	6.8	3.8

+with and − without including usage of previous day and week.

Table 5
Aggregated ranking of each model, from best (rank = 1) to worst (rank = 6).

Data Used	Model Type	Water		Energy	
		CV_{RMSE} %	MNAR %	CV_{RMSE} %	MNAR %
Cluster-Specific	LASSO^{+}	3	3	2	3
	LASSO^{-}	5	4	5	5
Overall	LASSO^{+}	4	5	4	4
	TBAG^{+}	1	1	1	1
Overall	LASSO^{-}	6	6	6	6
	TBAG^{-}	2	2	3	2

+ with and − without including usage of previous day and week.

shown in Figures S4 and S5 in the SI). In this case, it can be inferred that that the main drivers in a residential setting (e.g., showers, toilets, faucets), differentiate multi-family residential buildings from the non-residential buildings, which is evident from the results as dormitory buildings do not appear in any of the other meta-clusters. Moreover, because buildings in the housing category exhibit high water use intensity, housing buildings would make ideal candidates for water efficiency upgrades and DSM policies. In terms of energy, dormitory buildings are among the lowest energy users, suggesting that dormitory buildings are not at high priority for energy efficiency upgrades.

Except for meta-cluster 6, all meta-clusters shown contain more than one building use designation, signifying that building use designation alone is not appropriate to characterize water and energy consumption patterns. To demonstrate this point, the usage patterns of four different buildings, each of which have the same building use designation, but different water and energy consumption patterns, were examined.

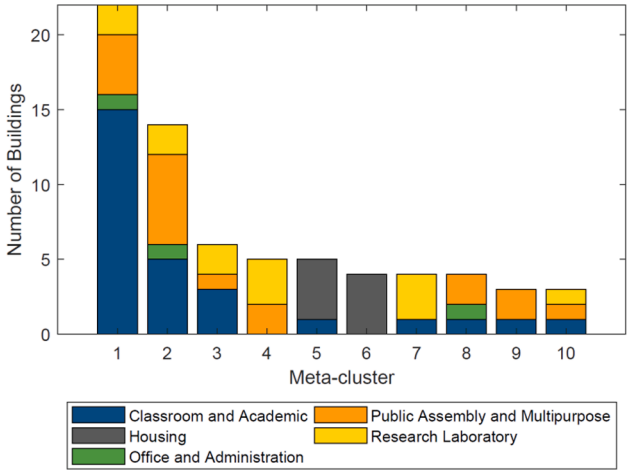

Fig. 13. Composition of each meta-cluster by building use designation.

Fig. 14 highlights the consumption patterns of four different classroom and academic (CA) buildings that were clustered into different meta-clusters. All buildings exhibit demand variability between weekday and weekend consumption. However, differences exist among the buildings in each metric. For example, when comparing the College of Nursing (blue in Fig. 14), to Sid Richardson Hall (green in Fig. 14), it is observed that the buildings exhibit comparable water use intensity, but the College of Nursing exhibits considerably higher energy intensity than Sid Richardson Hall. Additionally, it is observed that a reverse trend exists in the water and energy normalized consumption patterns, indicating that water consumption in Sid Richardson Hall is typically low throughout the year (compared to its overall annual consumption range), while the energy demand is typically high. The opposite is observed for the College of Nursing with characteristically high water consumption and low energy consumption throughout the year. Jester Hall, which contains many offices for university programs, as well as a large dining hall that serves as a major hub for student meals, is the highest consumer across all performance metrics (orange in Fig. 14). It is inferred that due to the water and energy demands for large amounts of food preparation, as well as the seasonal variation that mimics the water demand patterns observed in housing buildings (i.e., decreases in resource consumption during academic breaks), Jester Hall behaves differently from other CA buildings. Although this case is specific to a

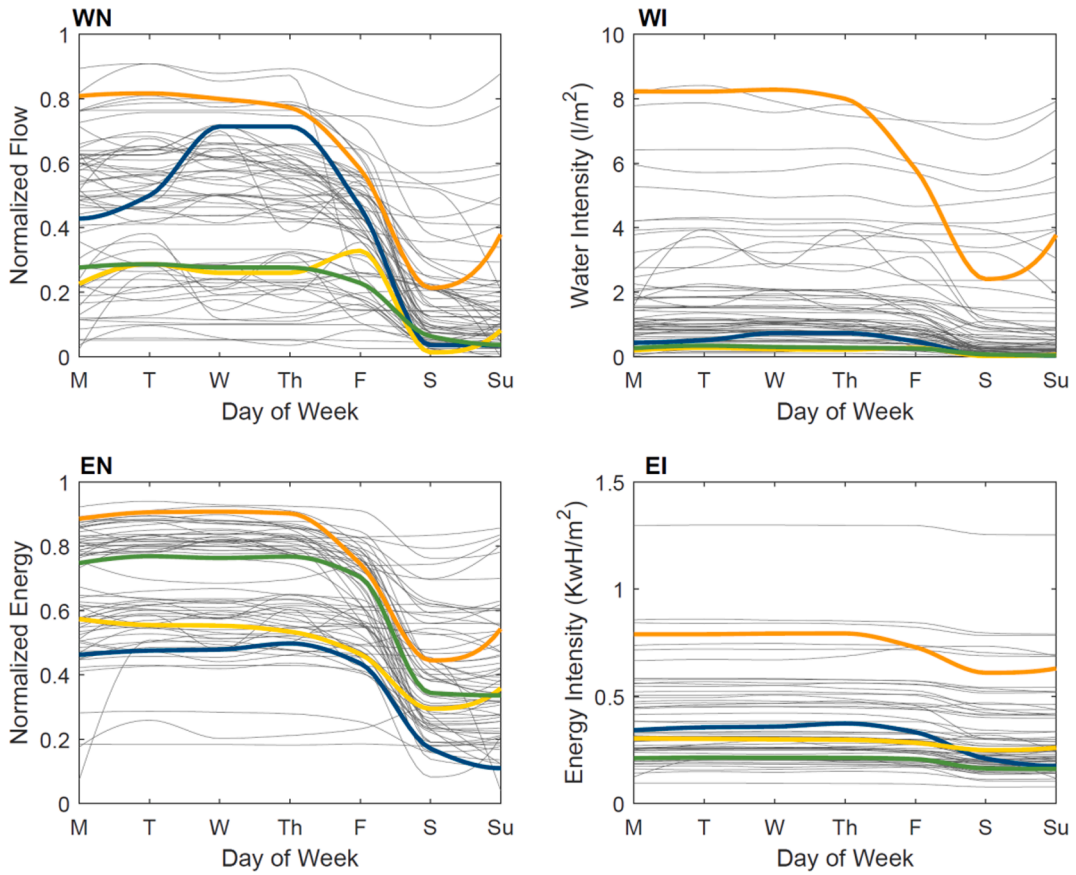

Fig. 14. Weekly patterns of four different CA buildings: College of Nursing (blue), Jester Hall (orange), Sid Richardson Hall (green), William Hearst Building (yellow). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

university setting, the observation of Jester Hall exhibiting high water and energy consumption is indicative of how a commercial office building with foodservice may behave.

While insights into building behavior are drawn from the clustering analysis, modeling of building performance yields insights into the effect of model inputs, and the importance of including a continuously-recording meter for demand prediction. Based on model performance, the relative effect of each independent variable on water and energy consumption prediction is analyzed. Because the TBAG model was the most accurate in predicting water and energy consumption, the discussion of important model variables is limited to TBAG models. The variable importance is quantified using the out-of-bag predictor importance, which is computed by permutating the observations of each independent variable, and determining the effect on the model results [78]. In essence, the out-of-bag predictor importance functions as a sensitivity test to determine how sensitive the overall model is to each variable. If a variable is important, then permutating that variable should have a relatively large effect on model performance, whereas permuting variables with little importance would have a small impact on model performance.

Fig. 15 shows the relative importance input variables, including building features (use designation, year built, number of floors, area) and temporal features (previous day and week consumption, temperature, monthly median consumption, class session, and day type) used in TBAG models for water and energy. Specifically, Fig. 15 displays the out-of-bag predictor index for each independent variable used as a model input for four cases: predicting water and energy, with and without the inclusion of previous consumption data. Overall, building features were less important compared to the temporal features for predicting consumption levels. It is observed that the type of day (i.e., weekday or weekend) is the most important explanatory variable for predicting

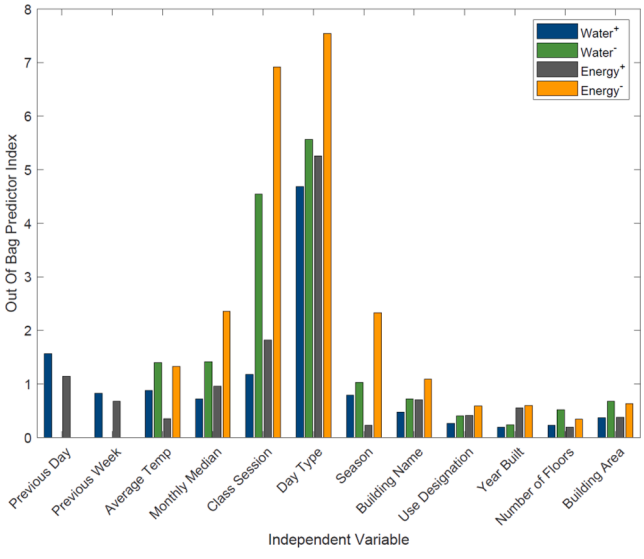

Fig. 15. Variable importance for predicting water and energy demand using the TBAG models. + with and − without including usage of previous day and week.

water demand regardless of the model. Class session, season, average daily temperature, and monthly median consumption are the consecutive important explanatory variables in the models that do not rely on previous observations from continuously-recording meters. These results are expected as these variables imply water consumption, e.g., higher water consumption is expected on a weekday during a semester

compared with a weekend during summer break. Previous day consumption, daily temperature, previous week and month consumption as well as class session and season were the consecutive explanatory variables in models that considered data available from continuously-recording meters. In general, observations made regarding the important variables for water consumption can be made for energy consumption. Noteworthy is that the importance of day type, monthly consumption, daily temperature, and season increase for energy consumption when previous data are not included.

For both water and energy predictions, results suggest that general building information, e.g., use designation, year built, number of floors, and area, are all among the least important variables when predicting consumption, which is consistent with the results shown previously in Fig. 13. The lower importance of variables related to general building information supports the assertion that this information should not be the sole basis of benchmarking, and other characteristics should be considered for categorizing buildings. Notably, the building features used in this study represent lumped building characteristics and not detailed information (e.g., number and type of water fixtures, efficiency of boilers, thermal transmission and resistance of buildings, etc.). Inclusion of such detailed information will improve prediction accuracy, but is significantly more intensive in terms of data collection, especially for large and multiple number of buildings. The variables that have a high influence on prediction models, such as day, class session, season, average daily temperature, and monthly consumption, provide information on the main drivers of consumption for a specific group of buildings. Characterizing how these variables affect demand and how specific building characteristics relate to these connections is crucial to understanding water and energy demand in urban setting. When specific building characteristics can be found, a better means for characterization and benchmarking can be developed without the use of daily or monthly water and energy data. Furthermore, planned buildings can be characterized into clusters and accounted for in DSM policies.

Finally, typical building water and energy consumption values published by the United States Energy Information Administration (USEIA) were reviewed to compare national consumption benchmarks to the building consumption values observed in this study [64–65]. Based on the USEIA surveys, mean daily water intensities among US buildings in different sectors range from 0.38 to 5.52 l/m^2. In this study, the mean daily water intensities of the different buildings range from 0.52 to 4.35 l/m^2, thus the buildings included in this study are representative of the large buildings surveyed by USEIA, in terms of water intensities. Among the different buildings, housing are the buildings with the highest water intensities, which are concentrated in meta-clusters 5 and 6, with water intensities comparable with the lodging sector, as reported by the USEIA. The rest of the buildings exhibit lower and mixed trends in terms of water intensity. Similar conclusions can be made regarding the applicability of buildings analyzed in this paper for energy intensity. The range of mean daily energy use intensities observed in this study is 0.22 to 0.87 kWh/m^2, compared to 0.19 to 0.91 kWh/m^2 reported by USEIA. Unlike with water intensity, housing have the lowest energy intensity (see Fig. 12), and the highest energy intensity buildings belong to meta-cluster 7 with a mix of research labs, classrooms and academics buildings. The energy intensity of the rest of buildings studied here was below 0.54 kWh/m^2.

5. Conclusions

Water and energy utilities are increasingly collecting vast amounts of data regarding user consumption, however the data must be analyzed in order to deliver insights and actionable information to the utility's managers. In the context of the water-energy nexus in urban environments, this work addressed current gaps in the literature revealing apparent dissimilarities between utilization of water and energy resources for heterogeneous buildings. The results of this study highlight the value of data-driven modeling for revealing meaningful insights and identifying emerging and/or unusual trends to facilitate multi-utility management.

Specifically, the clustering analysis performed in this work revealed that water and energy consumption patterns of heterogeneous buildings are not uniquely characterized by general building characteristics and that other specific building characteristics need to be explored. Analysis of the predictive models showed that an overall non-parametric model provides better predictions for water and energy compared with parametric models. Furthermore, while the non-parametric model performs best with high resolution, continuous data, the model also produces acceptable results with low-resolution consumption data, typically received from monthly billing data. The insights from this study have multiple implications. First, this study provides a new method to group buildings across water and energy consumption patterns. This information is valuable to benchmark buildings' performance and provide a meaningful measure of comparison as well as a measure for achieving targeted performance and identifying buildings that are candidates for water and/or energy efficiency upgrades. Furthermore, the data-driven prediction model could be used to forecast daily consumption levels for buildings that do not have continuous consumption data available and could facilitate water and energy utilities to update their demand forecasting models and detect abnormal technical (e.g., meters malfunctioning) and physical (e.g., leaks) events.

This paper provides a foundation upon which additional inquiries should be conducted. Although it was found that the buildings included in this study classify into one of 10 meta-clusters, exploring the drivers, which determine the classification of each building, was beyond the scope of this paper. Future work should further explore additional non-building features (e.g., occupancy, activity, gender) that would differentiate buildings from each other, following similar approaches applied to residential buildings [59,23,25]. Investigations into this aspect could provide means for classifying buildings without water and energy data and help predict changes in overall water and energy demand with future development. Another direction for future work should explore the effects of seasonality on meta-cluster composition. By stratifying the times for which buildings are classified (e.g., monthly instead of yearly), time-specific meta-clusters could inform resource consumption based on more granular timescales, thus generating dynamic clusters of buildings. Then by continuously analyzing buildings' classifications, temporal changes in building's water and energy behaviors can be detected and further investigated.

As the threat of climate change grows alongside a continual increase in urban population, the need to ensure access to water and energy resources becomes more crucial. Overall, the methods outlined in this study provide another step towards better water and energy management techniques, which can then be used to build greater resiliency within urban areas in preparation for future changes in population and climate.

CRediT authorship contribution statement

Matthew Frankel: Software, Writing - original draft, Visualization. **Lu Xing:** Methodology, Software, Writing - review & editing, Visualization. **Connor Chewning:** Software, Writing - review & editing. **Lina Sela:** Conceptualization, Methodology, Writing - original draft, Supervision, Funding acquisition.

Declaration of Competing Interest

None.

Acknowledgments

This work was supported by the University of Texas at Austin Green Fund and Startup Grants and by the National Science Foundation under Grant 1828974.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.apenergy.2020.116074.

References

[1] Stewart RA, Nguyen K, Beal C, Zhang H, Sahin O, Bertone E, et al. Integrated intelligent water-energy metering systems and informatics: Visioning a digital multi-utility service provider. Environ Modell Software 2018;105:94–117. https://doi.org/10.1016/j.envsoft.2018.03.006.

[2] Lewis C, Hendrix M. Smart Grid and AMI for Water Utilities. Retrieved from www.jstor.org/stable/jamewatworass.104.9.58. American Water Works Association 2012;104(9):58–61.

[3] Zhou K, Fu C, Yang S. Big data driven smart energy management: From big data to big insights. Renew Sustain Energy Rev 2016;56(2016):215–25. https://doi.org/10.1016/j.rser.2015.11.050.

[4] Alasseri R, Tripathi A, Joji Rao T, Sreekanth KJ. A review on implementation strategies for demand side management (DSM) in Kuwait through incentive-based demand response programs. Renew Sustain Energy Rev December 2015;77:617–35. https://doi.org/10.1016/j.rser.2017.04.023.

[5] DeOreo WB, Mayer P, Dziegielewski B, Kiefer J. Residential End Uses of Water, Version 2; 2016. Retrieved from https://www.circleofblue.org/wp-content/uploads/2016/04/WRF_REU2016.pdf.

[6] Jessoe K, Rapson D. Knowledge is (Less) power: Experimental evidence from residential energy use. Am Econ Rev 2014;104(4):1417–38. https://doi.org/10.1257/aer.104.4.1417.

[7] Olmstead SM, Stavins RN. Comparing price and nonprice approaches to urban water conservation. Water Resour Res 2009;45(4):1–10. https://doi.org/10.1029/2008WR007227.

[8] Yilmaz S, Weber S, Patel MK. Who is sensitive to DSM? Understanding the determinants of the shape of electricity load curves and demand shifting: Socio-demographic characteristics, appliance use and attitudes. Energy Policy 2019;133(August):110909. https://doi.org/10.1016/j.enpol.2019.110909.

[9] Petersen JE, Frantz CM, Shammin MR, Yanisch TM, Tincknell E, Myers N. Electricity and water conservation on college and university campuses in response to national competitions among dormitories: Quantifying relationships between behavior, conservation strategies and psychological metrics. PLoS ONE 2015;10(12):1–41. https://doi.org/10.1371/journal.pone.0144070.

[10] Britton TC, Stewart RA, O'Halloran KR. Smart metering: Enabler for rapid and effective post meter leakage identification and water loss management. J Cleaner Prod 2013;54:166–76. https://doi.org/10.1016/j.jclepro.2013.05.018.

[11] Gurung TR, Stewart RA, Beal CD, Sharma AK. Smart meter enabled informatics for economically efficient diversified water supply infrastructure planning. J Cleaner Prod 2017;163:S138–47. https://doi.org/10.1016/j.jclepro.2017.05.140.

[12] Nguyen KA, Stewart RA, Zhang H. An intelligent pattern recognition model to automate the categorisation of residential water end-use events. Environ Modell Software 2013;47:108–27. https://doi.org/10.1016/j.envsoft.2013.05.002.

[13] Patabendige S, Cardell-Oliver R, Wang R, Liu W. Detection and interpretation of anomalous water use for non-residential customers. Environ Modell Software 2018;100:291–301. https://doi.org/10.1016/j.envsoft.2017.11.028.

[14] Fonseca JA, Miller C, Schlueter A. Unsupervised load shape clustering for urban building performance assessment. Energy Proc 2017;122:229–34. https://doi.org/10.1016/j.egypro.2017.07.350.

[15] Rahayu D, Narayanaswamy B, Krishnaswamy S, Labbé C, Seetharam DP. Learning to Be Energy-Wise. Discriminat Methods Load Disaggr 2012;1–4. https://doi.org/10.1145/2208828.2208838.

[16] Shaw SR, Leeb SB, Norford LK, Cox RW. Nonintrusive load monitoring and diagnostics in power systems. IEEE Trans Instrum Meas 2008;57(7):1445–54. https://doi.org/10.1109/TIM.2008.917179.

[17] Zoha A, Gluhak A, Imran MA, Rajasegarar S. Non-intrusive Load Monitoring approaches for disaggregated energy sensing: A survey. Sensors (Switzerland) 2012;12(12):16838–66. https://doi.org/10.3390/s121216838.

[18] Candelieri A. Clustering and support vector regression for water demand forecasting and anomaly detection. Water (Switzerland) 2017;9(3). https://doi.org/10.3390/w9030224.

[19] Aghabozorgi S, Seyed Shirkhorshidi A, Ying Wah T. Time-series clustering - A decade review. Inform Syst 2015;53:16–38. https://doi.org/10.1016/j.is.2015.04.007.

[20] Avni N, Fishbain B, Shamir U. Water consumption patterns as a basis for water demand modeling. Water Resour Res 2015;2498–2514. https://doi.org/10.1002/2015WR017200.A.

[21] Beal CD, Stewart RA. Identifying residential water end uses underpinning peak day and peak hour demand. J Water Resour Plann Manage 2014;140(7):1–10. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000357.

[22] Heidarinejad M, Dahlhausen M, McMahon S, Pyke C, Srebric J. Cluster analysis of simulated energy use for LEED certified U.S. office buildings. Energy Build 2014;85:86–97. https://doi.org/10.1016/j.enbuild.2014.09.017.

[23] Kontokosta CE, Jain RK. Modeling the determinants of large-scale building water use: Implications for data-driven urban sustainability policy. Sustain Cities Soc 2015;18:44–55. https://doi.org/10.1016/j.scs.2015.05.007.

[24] Miller C, Meggers F. Mining electrical meter data to predict principal building use, performance class, and operations strategy for hundreds of non-residential buildings. Energy Build 2017;156:360–73. https://doi.org/10.1016/j.enbuild.2017.09.056.

[25] Rhodes JD, Cole WJ, Upshaw CR, Edgar TF, Webber ME. Clustering analysis of residential electricity demand profiles. Appl Energy 2014;135:461–71. https://doi.org/10.1016/j.apenergy.2014.08.111.

[26] Duerr I, Merrill HR, Wang C, Bai R, Boyer M, Dukes MD, et al. Forecasting urban household water demand with statistical and machine learning methods using large space-time data: A Comparative study. Environ Modell Software 2018;102:29–38. https://doi.org/10.1016/j.envsoft.2018.01.002.

[27] Zhao H, Magoulès F. A review on the prediction of building energy consumption. Renew Sustain Energy Rev 2012;16(6):3586–92. https://doi.org/10.1016/j.rser.2012.02.049.

[28] Boyle T, Giurco D, Mukheibir P, Liu A, Moy C, White S, et al. Intelligent metering for urban water: A review. Water (Switzerland) 2013;5(3):1052–81. https://doi.org/10.3390/w5031052.

[29] Carrie Armel K, Gupta A, Shrimali G, Albert A. Is disaggregation the holy grail of energy efficiency? The case of electricity. Energy Policy 2013;52:213–34. https://doi.org/10.1016/j.enpol.2012.08.062.

[30] Cominola A, Giuliani M, Piga D, Castelletti A, Rizzoli AE. Benefits and challenges of using smart meters for advancing residential water demand modeling and management: A review. Environ Modell Software 2015;72:198–214. https://doi.org/10.1016/j.envsoft.2015.07.012.

[31] Deb C, Zhang F, Yang J, Lee SE, Shah KW. A review on time series forecasting techniques for building energy consumption. Renew Sustain Energy Rev 2017;74(November):902–24. https://doi.org/10.1016/j.rser.2017.02.085.

[32] Hong T, Wang Z, Luo X, Zhang W. State-of-the-art on research and applications of machine learning in the building life cycle. Energy Build 2020;212:109831. https://doi.org/10.1016/j.enbuild.2020.109831.

[33] Miller C, Nagy Z, Schlueter A. A review of unsupervised statistical learning and visual analytics techniques applied to performance analysis of non-residential buildings. Renew Sustain Energy Rev 2018;81(May 2016): 1365–77. https://doi.org/10.1016/j.rser.2017.05.124.

[34] Jain AK. Data clustering: 50 years beyond K-means. Pattern Recogn Lett 2010. https://doi.org/10.1016/j.patrec.2009.09.011.

[35] Park HS, Jun CH. A simple and fast algorithm for K-medoids clustering. Expert Syst Appl 2009. https://doi.org/10.1016/j.eswa.2008.01.039.

[36] Contreras P, Murtagh F. Hierarchical clustering. Handbook Cluster Anal 2015. https://doi.org/10.1201/b19706.

[37] Vesanto J, Alhoniemi E. Clustering of the self-organizing map. IEEE Trans Neural Networks 2000. https://doi.org/10.1109/72.846731.

[38] Ng AY, Jordan MI, Weiss Y. On spectral clustering: Analysis and an algorithm. Advances in Neural Information Processing Systems; 2002.

[39] Ester M, Kriegel HP, Sander J, Xu X. A density-based algorithm for discovering clusters in large spatial databases with noise; 1996. Kdd. https://doi.org/10.1.1.71.1980.

[40] Rodriguez A, Laio A. Clustering by fast search and find of density peaks. Science 2014. https://doi.org/10.1126/science.1242072.

[41] McInnes L, Healy J, Astels S. hdbscan: Hierarchical density based clustering. J Open Source Software 2017. https://doi.org/10.21105/joss.00205.

[42] Rendón E, Abundez I, Arizmendi A, Quiroz EM. Internal versus External cluster validation indexes. Retrieved from Int J Comput Commun 2011;5(1):27–34. http://w.naun.org/multimedia/UPress/cc/20-463.pdf.

[43] Dolnicar S. A review of unquestioned standards in using cluster analysis for data-driven market segmentation. In: Faculty of Commerce-Papers; 2002.

[44] Saxena A, Prasad M, Gupta A, Bharill N, Patel OP, Tiwari A, et al. A review of clustering techniques and developments. Neurocomputing 2017. https://doi.org/10.1016/j.neucom.2017.06.053.

[45] Donkor EA, Mazzuchi TA, Soyer R, Robertson JA. Urban Water Demand Forecasting: Review of Methods and Models. J Water Resour Plann Manage 2013;139(June):554–64. https://doi.org/10.1061/(ASCE)WR.1943-5452.

[46] Suganthi L, Samuel AA. Energy models for demand forecasting - A review. Renew Sustain Energy Rev 2012;16(2):1223–40. https://doi.org/10.1016/j.rser.2011.08.014.

[47] Robinson C, Dilkina B, Zhang W, Guhathakurta S, Brown MA, Pendyala RM. Machine learning approaches for estimating commercial building energy consumption. 2017;208(May): 889–904. https://doi.org/10.1016/j.apenergy.2017.09.060.

[48] Tiwari MK, Adamowski JF. Medium-Term Urban Water Demand Forecasting with Limited Data Using an Ensemble Wavelet-Bootstrap Machine-Learning Approach. J Water Resour Plann Manage 2014;141(2):04014053. https://doi.org/10.1061/(asce)wr.1943-5452.0000454.

[49] Tso GKF, Yau KKW. Predicting electricity energy consumption: A comparison of regression analysis, decision tree and neural networks. Energy 2007;32(9):1761–8. https://doi.org/10.1016/j.energy.2006.11.010.

[50] Walker D, Creaco E, Vamvakeridou-Lyroudia L, Farmani R, Kapelan Z, Savić D. Forecasting domestic water consumption from smart meter readings using statistical methods and artificial neural networks. Proc Eng 2015;119(1):1419–28. https://doi.org/10.1016/j.proeng.2015.08.1002.

[51] Zeng A, Liu S, Yu Y. Comparative study of data driven methods in building electricity use prediction. Energy Build 2019;194:289–300. https://doi.org/10.1016/j.enbuild.2019.04.029.

[52] Adamowski J, Karapataki C. Comparison of Multivariate Regression and Artificial Neural Networks for Peak Urban Water-Demand Forecasting: Evaluation of Different ANN Learning Algorithms. J Hydrol Eng 2010;15(10):729–43. https://doi.org/10.1061/(asce)he.1943-5584.0000245.

[53] Frumhoff PC, Burkett V, Jackson RB, Newmark R, Overpeck J, Webber M. Vulnerabilities and opportunities at the nexus of electricity, water and climate. Environ Res Lett 2015;10(8). https://doi.org/10.1088/1748-9326/10/8/080201.
[54] Obringer R, Kumar R, Nateghi R. Analyzing the climate sensitivity of the coupled water-electricity demand nexus in the Midwestern United States. Appl Energy 2019;252(May):113466. https://doi.org/10.1016/j.apenergy.2019.113466.
[55] Stillwell AS, King CW, Webber ME, Duncan IJ, Hardberger A. The Energy-Water Nexus in Texas 2011;16(1).
[56] Vakilifard N, Anda M, Bahri AP, Ho G. The role of water-energy nexus in optimising water supply systems – Review of techniques and approaches. Renew Sustain Energy Rev 2018;82(June 2016), 1424–1432. https://doi.org/10.1016/j.rser.2017.05.125.
[57] Nguyen K, Stewart RA, Sahin O, Bertone E, Beal CD, Cominola A, et al. Digital Multi-Utility Data for Contemporaneous Water-Electricity-Gas End Use Categorization. In: Proceedings - 2019 3rd International Conference on Smart Grid and Smart Cities, ICSGSC; 2019. p. 45–50. https://doi.org/10.1109/ICSGSC.2019.00-20.
[58] United States Department of Energy, & University of California. Capturing the Benefits of Integrated Resource Management for Water & Electricity Utilities and their Partners; 2016. https://doi.org/10.1002/2014GL061055.
[59] Cominola A, Spang ES, Giuliani M, Castelletti A, Lund JR, Loge FJ. Segmentation analysis of residential water-electricity demand for customized demand-side management programs. J Cleaner Prod 2018;172:1607–19. https://doi.org/10.1016/j.jclepro.2017.10.203.
[60] Vitter JS, Webber ME. A non-intrusive approach for classifying residential water events using coincident electricity data. Environ Modell Software 2018;100:302–13. https://doi.org/10.1016/j.envsoft.2017.11.029.
[61] Beal CD, Bertone E, Stewart RA. Evaluating the energy and carbon reductions resulting from resource-efficient household stock. Energy Build 2012;55:422–32. https://doi.org/10.1016/j.enbuild.2012.08.004.
[62] Hong T, Lin HW. Occupant Behavior: Impact on Energy Use of Private Offices. In: ASim 2012- 1st Asia Conference of International Building Performance Simulation Association, (January); 2013. Retrieved from https://www.osti.gov/servlets/purl/1172115.
[63] Texas Water Development Board. Water Use of Texas Water Utilities 2019; 2019.
[64] U.S. Energy Information Administration. Select Results from the Energy Assessor Experiment in the 2012 Commercial Buildings Energy Consumption Survey; 2015. Retrieved from https://www.eia.gov/consumption/commercial/reports/2012/assessorexp/?src=⟨ Consumption Commercial Buildings Energy Consumption Survey (CBECS)-b3.
[65] U.S. Energy Information Administration. 2012 Commercial Buildings Energy Consumption Survey: Water Consumption in Large Buildings Summary; 2017. Retrieved from https://www.eia.gov/consumption/commercial/reports/2012/water/?src=⟨ Consumption Commercial Buildings Energy Consumption Survey (CBECS)-b2.
[66] Strehl A, Ghosh J. Cluster ensembles - A knowledge reuse framework for combining multiple partitions. J Mach Learn Res 2003. https://doi.org/10.1162/153244303321897735.
[67] The University of Texas at Austin. Facts & Figures. Retrieved from, https://www.utexas.edu/about/facts-and-figures; 2019.
[68] The University of Texas at Austin. Sustainability Master Plan. Retrieved from, https://capitalplanning.utexas.edu/sites/cpc.utexas.edu/files/Sustainability-MasterPlan-2016.pdf; 2016.
[69] Kontokosta CE, Tull C. A data-driven predictive model of city-scale energy use in buildings. Appl Energy 2017;197:303–17. https://doi.org/10.1016/j.apenergy.2017.04.005.
[70] Legesse G, Cordeiro MRC, Ominski KH, Beauchemin KA, Kroebel R, McGeough EJ, et al. Water use intensity of Canadian beef production in 1981 as compared to 2011. Sci Total Environ 2018;619–620:1030–9. https://doi.org/10.1016/j.scitotenv.2017.11.194.
[71] Zohrabian A, Sanders KT. Assessing the impact of drought on the emissions- and water-intensity of California's transitioning power sector. Energy Policy 2018;123(September):461–70. https://doi.org/10.1016/j.enpol.2018.09.014.
[72] Gibbons JD, Chakraborti S. Nonparametric Statistical Inference, Fourth Edition: Revised and Expanded (Fourth). Retrieved from 2003. https://books.google.com/books?id=kJbVO2G6VicC&pgis=1.
[73] Cover TM, Thomas JA. Entropy, Relative Entropy, and Mutual Information. Elements Inform Theory 2005. https://doi.org/10.1002/047174882x.ch2.
[74] Shannon CE. A Mathematical Theory of Communication. Bell Syst Tech J 1948. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x.
[75] Müllner D. Modern hierarchical, agglomerative clustering algorithms. ArXiv Preprint ArXiv:1109.2378; 2011. 1–29. https://doi.org/10.1109/LSP.2012.2188026.
[76] Tibshirani R. Regression Shrinkage and Selection via the Lasso 1996;58(1):267–88.
[77] Lai TL, Robbins H, Wei CZ. Strong consistency of least squares estimates in multiple regression models with random regressors. Metrika 1978;77(3):361–75. https://doi.org/10.1007/s00184-013-0443-y.
[78] Loh W. Fifty years of classification and regression trees. Int Stat Rev 2014;82(3):329–48. https://doi.org/10.1111/insr.12016.
[79] Ng AY. Feature selection, L1 vs L2 regularization, and rotatational invarience. Proceedings of the 21st International Conference on Machine Learning. 2004.
[80] Loh WY. Classification and regression trees. Classification and Regression Trees; January 2011. 1–358. https://doi.org/10.1201/9781315139470.
[81] Breiman L. Bagging Predictors. Mach Learn 1996;24(421):123–40. https://doi.org/10.1007/BF00058655.
[82] ASHRAE. Measurement of Energy, Demand, and Water Savings. In ASHRAE Guideline 14-2014 (Vol. 2014); 2014. Retrieved from www.ashrae.org/technology.
[83] Ruiz GR, Bandera CF. Validation of calibrated energy models: Common errors. Energies 2017;10(10). https://doi.org/10.3390/en10101587.
[84] Yang Z, Becerik-Gerber B. A model calibration framework for simultaneous multi-level building energy simulation. Appl Energy 2015;149:415–31. https://doi.org/10.1016/j.apenergy.2015.03.048.
[85] National Weather Service Forecast Office. Observed Weather Reports. Retrieved January 4, 2019, from, https://w2.weather.gov/climate/index.php?wfo=ewx; 2019.