Diss. ETH No. XXXXX

Screening Meter Data: Characterization of Temporal Energy Data from Large Groups of Non-Residential Buildings

A thesis submitted to attain the degree of

DOCTOR OF SCIENCES of ETH ZURICH

(Dr. sc. ETH Zurich)

presented by

CLAYTON C. MILLER

Masters of Science (MSc.) Building, National University of Singapore

Masters of Architectural Engineering (MAE), Univ. of Nebraska

born on 22 January 1984

citizen of the United States of America

accepted on the recommendation of

Prof. Dr. Arno Schlueter, examiner

Prof. Dr. Stefan Mueller Arisona, co-examiner

2016

Abstract

This study focuses on the screening of characteristic data from the ever-expanding sources of raw, temporal sensor data from commercial buildings. A two-step framework is presented that extracts statistical, model-based, and pattern-based behavior from two real-world collected data sets. The collection is 507 commercial buildings extracted from various case studies and online data sources from around the world. The second are advanced metering infrastructure (AMI) data from 1,600 buildings. The goal of the framework is to reduce the expert intervention needed to utilize measured raw data in order to extract information such as building use type, performance class, and operational behavior. The first step is feature extraction and it utilizes a library of temporal data mining techniques to filter various phenomenon from the raw data. This step transforms quantitative raw data into qualitative categories that can be interpreted easily heat map visualization. In the second step, or the investigation, a supervised learning technique is tested in the ability to assign impact scores to the most important features from the first step. The efficacy of estimating variable causality of the characterized performance is tested to determine scalability amongst a heterogeneous sample of buildings. In the first set of case studies, characterization as compared to a baseline was three times more accurate in characterizing primary buildng use type, almost twice for performance class, and over four times for building operations type. For the AMI data, characterizing the standard industry class was improved by 27% and predicting the success of energy savings measures was improved by 18%. Qualitative insight from several campus case study interviews are discussed as well. The usefulness of the approaches was discussed in the context of campus building operations.

Kurzfassung

Acknowledgements

Contents

List of Figures

1 Introduction

The built and urban environments have a significant impact on resource consumption and greenhouse gas emissions in the world. The United States is the world's second largest energy consumer, and buildings there account for 41% of energy consumed[1]. The most extensive meta-analysis thus far of non-residential existing buildings showed a median opportunity of 16% energy savings potential by using cost-effective measures to remedy performance deficiencies (Mills 2011). Simply stated, roughly 6% of the energy consumed in the U.S. could be easily mitigated - a figure that would eventually grow to an annual energy savings potential of $30 billion and 340 megatons of C02 by the year 2030. Beyond saving energy, money and mitigating carbon, the impact of building performance improvement also extends to the health, comfort and satisfaction of the people who use buildings.

It is mysterious that these performance improvements are not rapidly being identified and implemented on a massive scale across the world's building stock given the incentives and amount of research focused on building optimization in the fields of Architecture, Engineering and Computer Science. A comprehensive study of building performance analysis was completed by the California Commissioning Collaborative (CACx) to characterize the technology, market, and research landscape in the United States. Three of the key tasks in this project focused on establishing the state of the art (Effinger *et al.* 2010), characterizing available tools and the barriers to adoption (Ulickey *et al.* 2010), and establishing standard performance metrics (Greensfelder *et al.* 2010). These reports were accomplished through investigation of the available tools and technologies on the market as well as discussions and surveys with building operators and engineers. The common theme amongst the interviews and case studies was the *lack of time and expertise* on the part of the dedicated operations professionals. The findings showed that installation time and cost was driven by the need for an engineer to develop a full understanding of the building and systems. These barriers reduce the implementation of performance improvements.

[1] As of 2014, according to: `http://www.eia.gov/`

In another study, Ruparathna et al. created a contemporary review of building performance analysis techniques for commercial and institutional buildings (Ruparathna *et al.* 2016). This review was comprehensive in capturing approaches related to technical, organizational, and behavioral changes. The majority of publications considered fall within the domains of automated fault detection and diagnostics, retrofit analysis, building benchmarking, and energy auditing. These traditional techniques focus on one building or a small, related collection of buildings, such as a campus. Many require complex characteristic data about each building, such as it geometric dimensions, building materials, the age and type of mechanical systems, and other metadata, to execute the process. Once again, such detailed techniques rely on metadata that often doesn't exist in the field, thus contributing to the barriers listed above.

Another issue facing the building industry is the characterization of the commercial building stock for benchmarking, intervention targeting, and general understanding of the way modern buildings are being utilized and operated. The Commercial Building Energy Consumption Survey (CBECS) is the primary means of collecting a characteristic data about the global commercial building stock in the United States. This survey is conducted every four years, the latest in 2012 in which information on over 6,700 building around the U.S. was collected for characterization. A large amount of meta-data was collected about each building from categories such as size, vintage, geographic region, and principal activity. This data collection was done through the efforts about 250 interviewers across the country under the supervision of 17 field supervisors, three regional field managers, and a field director. This manpower was utilized over the course of over three years to characterize and document the commercial building stock.

From these studies, it becomes apparent that the biggest barrier to achieving performance improvement in buildings is scalability. Architecture is a discipline founded with aesthetic creativity as a core tenet. Frank Lloyd Wright once stated, "The mother art is architecture. Without an architecture of our own, we have no soul of our civilization." Designers rightfully strive for artistic and meaningful creations; this phenomenon results in buildings with not only distinctive aesthetics but also unique energy systems design, installation practices and different levels of organization within the data-creating components. In this dissertation, I show that an emerging mass of data from the built environment can facilitate better characterization of buildings by through automation of meta-data extraction. These data are temporal sensor measurements from performance measurement systems.

1.1 Growth of Raw Temporal Data Sources in the Built Environment

As entities of analysis, buildings are less on the level of a typical mass-produced manufactured device in which each unit is the same in its components and functionality; and more on the level of customers of business, entities that are similar and yet have numerous nuances. Conventional mechanistic or model-based approaches, typically borrowed from manufacturing, have been the status quo in building performance research. As previously discussed, scalability amongst the heterogeneous building stock is a significant barrier to these approaches. More appropriate means of analysis lies in statistical learning techniques more often found in the medical, pharmaceutical and customer acquisition domains. These methods rely on extracting information and correlating patterns from large empirical data sets. *The strength of these techniques is in their robustness and automation of implementation - concepts explicitly necessary to meet the challenges outlined.*

This type of research on buildings would have been tough even a few years ago. The creation and consolidation of measured sensor sources from the built environment and its occupants is occurring on an unprecedented scale. The Green Button Ecosystem now enables the easy extraction of performance data from over 60 million buildings[2]. Advanced metering infrastructure (AMI), or smart meters, have been installed on over 58.5 million buildings in the US alone[3]. A recent press release from the White House summarizes the impact of utilities and cities in unlocking these data (The White House 2016). It announces that 18 utilities, serving more than 2.6 million customers, will provide detailed energy data by 2017. This study also suggests that such accessibility will enable improvement of energy performance in buildings by 20% by 2020. A vast majority of these raw data being generated are sub-hourly temporal data from meters and sensors.

To understand the exponential magnitude of this source data growth in the building industry, one can estimate the amount of measurements being generated by these sensors. The United States context has public data available to create a set of assumptions to roughly quantify this growth. Before the widespread use of digital building automation systems, buildings were controlled either manually or using pneumatic controls and building electrical use was measured and reported monthly. According to the Commercial Building Energy Consumption Survey, there were over 4.5 million commercial buildings in the United States in 1996. The theoretical amount of data from monthly electrical meters for all of these buildings for one year would be 54 million measurements. In about 2007,

[2]According to: http://www.greenbuttondata.org/

[3]As of 2014, according to: `http://www.eia.gov/tools/faqs/faq.cfm?id=108&t=3`

electrical meters with the capability to capture and store data at 15-minute frequencies were introduced into the market, and 7 million were installed on all building types [4]. If one assumes that the proportion of these meters that are commercial is similar to today[5], that will result in approximately 784,000 buildings creating 27.4 billion measurements per year. By 2014, AMI meters have been installed on 6.53 million commercial buildings resulting in 228 billion measurements per year. The exponential magnitudes of growth of these data can be seen in Figure 1.1. This discussion ignores the concept of accessibility which has also vastly improved due to the technology.

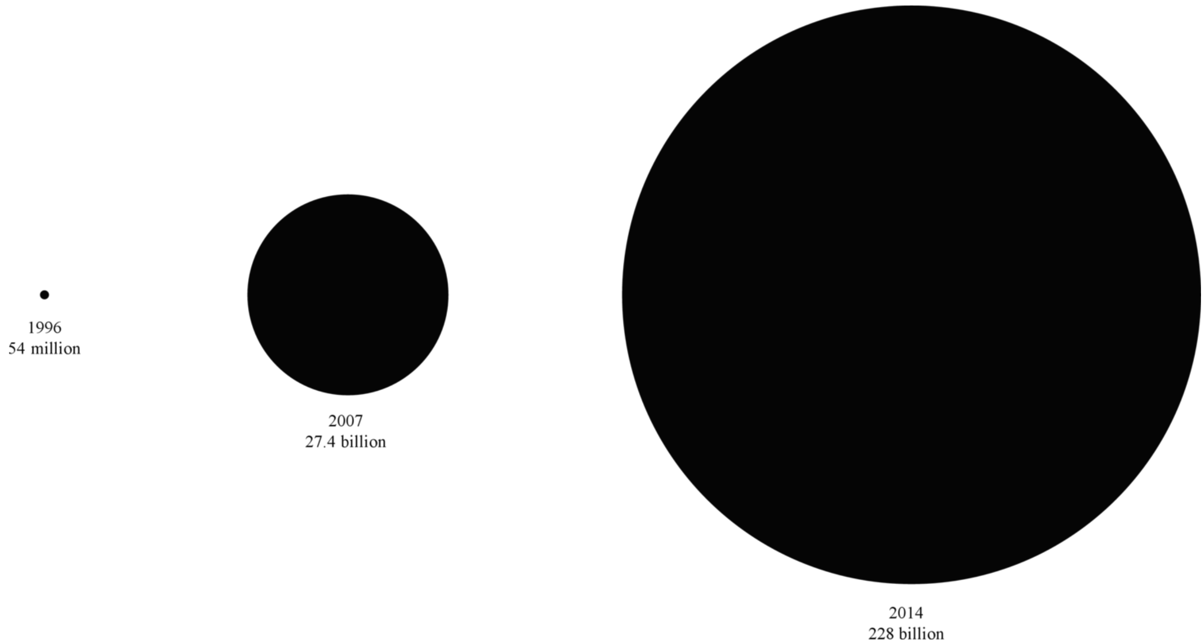

Figure 1.1: Theoretical growth of measurement data from electrical meters in commercial buildings in the USA in the last 20 years

The analysis of the performance of buildings and the characterization of the building stock are necessary and, as discussed, quite tedious challenges in the building industry. Thus, a critical opportunity for the building industry is how these techniques can utilize the aforementioned explosion of detailed, temporal sources.

- *If one has access to raw data from hundreds, or even thousands, of buildings, how can analysis be scaled in a robust way?*
- *How can these data be used to inform the larger research community about the phenomenon occurring in the actual building stock?*

[4] `http://www.edisonfoundation.net/iei/Documents/IEI_SmartMeterUpdate_0914.pdf`
[5] About 11.2% according to: `http://www.eia.gov/tools/faqs/faq.cfm?id=108&t=3`

- *What characteristic data about buildings can be inferred from these sources?*

Non-residential buildings are the focus of this analysis as they are unique and complex in their energy-consuming systems. This decision was designed to limit the scope to a subset of the building industry that is under-researched as compared to residential buildings.

1.2 A Framework for Automated Characterization of Large Numbers of Non-Residential Buildings

This thesis develops a framework to investigate which characteristics of whole building electrical meter data are most indicative of various meta-data about buildings amongst large collections of commercial buildings. This structure is designed to *screen* electrical meter data for insight on the path towards deeper data analysis. The screening nature of the process is motivated by the scalability challenges previously outlined. An initial component in the methodology was a series of case study interviews and data collection processes to survey field data from numerous buildings around the world. Two phases were then applied to the collected data. The first was to use a library of temporal feature extraction techniques for the purpose of retrieving various behavior from whole building electrical sensor data in a relatively fast and unsupervised fashion. The second process utilizes these features in classification models to determine the accuracy of predicting various meta-data about each building. The classification aspect of the process is designed primarily to establish the importance of the input variables in their ability to characterize various behavior. Several meta-data are targeted to test this framework such as building use type, performance class, and operational strategy. These objectives were chosen as they represent steps in the direction of benchmarking, diagnostics, retrofit analysis, and other types of building performance analysis techniques.

1.3 Research Questions

The primary question addressed through this research is:

- How accurately can the meta-data about a building be characterized through the analysis of raw hourly or sub-hourly, whole building electrical meter data?

This question is dissected into several more specific parts:

- Which temporal features are most accurate in classifying the primary use-type, performance class, and operational strategy of a building?
- Can temporal features be used to better benchmark buildings by signifying how *well a building fits within its designated use-type class*?
- Can temporal features be used to forecast whether an energy savings intervention measure will be successful or not?
- What are the most appropriate parameter settings for various generalized temporal feature extraction techniques as applied to this context?
- Is it effective or possible to implement such features across data from tens of thousands of buildings?

1.4 Objectives

The objectives of this research are as follows:

1. Consolidate and curate a set of feature extraction techniques from various research domains that automatically extract characteristic information from raw, temporal data
2. Extend these feature sets to include pattern recognition approaches that capture more information through characterizing usage patterns
3. Deploy these features on a test data set of 507 buildings to quantify the ability to characterize building use type, in-class performance, and operations types
4. Deploy a subset of features on a data set of approximately 1,600 buildings to test the ability to predict whether an energy-savings measure implementation will be a success

1.5 Organization of the Thesis

The remainder of this thesis is organized as follows. The research context of contemporary statistical learning and visual analytics techniques as applied to building performance is reviewed in Section 2. This section has a special focus on unsupervised learning techniques as they are a strong basis for many of the temporal features extracted. Section 3 provides an overview of the two steps in the framework as well as the process of collecting data and insight from a series of case studies from around the world. Data from over 1200 buildings was collected on-site or through various open web portals and 507 were selected for further analysis. Sections 4-6 provide in-depth overviews of each category of the temporal mining techniques implemented on the case study buildings, including explanatory visualizations of the range of values across the tested time range. Section 7 discusses the use of these features for the characterization of objectives such as predicting building use type, performance class, and operations type. Section 8 focuses on the use of a subset of temporal features in the industry classification and prediction of energy savings measures of close to 10,000 buildings with AMI data available. Finally, Section 9 provides concluding remarks to understand the overall results of the thesis and future directions to pursue using the outlined techniques.

2 Research Context: Statistical Learning and Visual Analytics of Building Data

This section gives an extensive overview of the techniques developed to extract automatically information from raw data to meet the scalability challenge. This content is developed as a publication submitted to the Renewable and Sustainable Energy Reviews Journal (Miller *et al.* Submitted for publication). The domains and range of techniques reviewed go beyond the scope of this dissertation. It considers a range of applications and objectives beyond the presented framework and research questions. The purpose of this effort is to set a wider context for understanding and discuss broader challenges and opportunities.

Researchers from several domains have developed methods of extracting insight from raw data from the built environment. Often these methods fall into the category of statistical learning, often from unsupervised learning. Methods from this sub-domain of machine learning are advantageous due to their ability to characterize measured or simulated performance data quickly with less analyst intervention, meta-data, and ground truth labeled data. In this section, a review of previous work in analytics methods is covered by the categories of smart meter analytics, portfolio analytics, operations and control, and anomaly detection for buildings.

2.1 Previous Reviews of Data Analytics in Buildings

Various reviews have been completed that overlap with this section. Most of them are designed to focus on a single core domain of research; the main two areas are building operations analysis and smart grid optimization. One of the earliest reviews of artificial intelligence techniques for buildings was completed in 2003 by Krarti and covered both supervised and unsupervised methods (Krarti 2003). Dounis updated this work and focused on outlining specific techniques in detail (Dounis 2010). Reddy's seminal book about a large variety of analysis techniques for energy engineers includes chapters on clustering and unsupervised methods specifically (Reddy 2011). Lee et al. describe a variety of retrofit analysis toolkits which incorporate unsupervised and visual analytics approaches in a practical sense (Lee *et al.* 2015). Ioannidis et al. created a large ontology of data mining and visual analytics for building performance analysis, however with a strong focus on the techniques and not examples of works using them (Ioannidis *et al.* 2015). From the utility and power grid side, Morias et al. created a general overview of various data mining techniques as focused on power distribution systems (Morais *et al.* 2009). Chicco covered clustering methods specifically focused on load profiling tasks (Chicco 2012). Zhou et al. included the concept of customer load classification (Zhou *et al.* 2013).

2.2 Overview of Publications

The work for this section was created through a selection of unsupervised analytics categories outlined by authoritative sources from the machine learning community (Hastie *et al.* 2009; James *et al.* 2013; Duda *et al.* 2012; Mirkin 2012). The groups selected are clustering, novelty detection, motif and discord detection, and rule extraction. The field of visual analytics was added to these groups to cover the presentation layer of many of these types of techniques. An initial search of publications was then selected for inclusion through a Google Scholar search of the combination of the method categories and the terms "building energy", "building performance analysis", and "building energy analysis". From this initial list of publications, a set of application categories and sub-categories was developed as seen in Figure 2.1. A more detailed search of each application class was then completed to account for the unique analytics techniques used in those domains. Only publications with a majority of the focus on utilization of unsupervised techniques and with a focus only on non-residential buildings are reviewed. Only works completed since 2005 are included to discuss only the most contemporary work and due to the relatively recent development of most of the techniques examined. A cutoff date of April 1, 2016, is applied for inclusions of publications in this review.

2.2.1 Research Sectors

Figure 2.2 illustrates the breakdown of publications based on the year published since 2005. They are further divided into four broad research domains: building energy analysis,

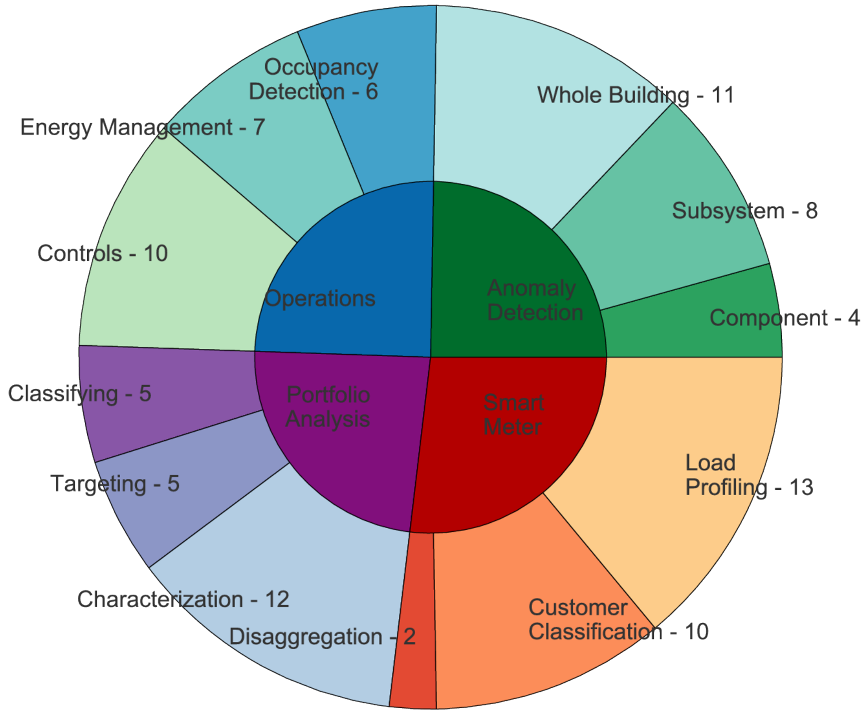

Figure 2.1: Categories and sub-categories (including number of publications) of building performance analysis applications of statistical learning and visual analytics

building simulation, computer science and electrical engineering. These research field categories were subjectively determined for each paper through evaluating a combination of which university department the authors were from and in which publication the study was published. Building energy analysis pertains to researchers who predominantly focus on measured data analysis from buildings while simulation experts research forward modeling and simulation of building and urban systems. Both fields of study most often exist within architecture or mechanical engineering departments. Electrical engineering and computer science are two well-established domains and exist in their departments. It is noticed that there is a gradual increase in the number of publications over the last ten years with electrical engineering and building energy analysis being the most common in the first few years and computer science and building simulation picking up since 2008.

2.2.2 Publications Venues

This section analyzes the prevalence of certain publication venues within this section. Figure 2.3 illustrates the breakdown of the publication venues represented. The Energy and Buildings Journal from the building energy analysis domain dominates this list with

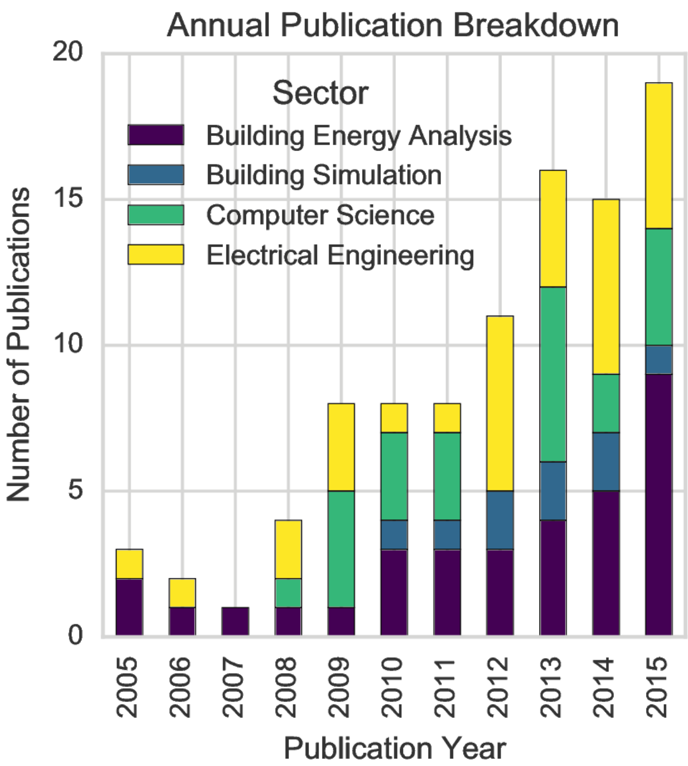

Figure 2.2: Breakdown of publications by year published and research domain

17 articles. Building simulation and energy analysis research domains publish most often in this journal as well as Applied Energy and Energy Efficiency. Several IEEE conferences and journals are also dominant as most of the papers from the electrical engineering domain are in these venues.

2.3 Smart Meter Analytics

Advanced Metering Infrastructure (AMI), also known as smart meter systems, is a network of energy meters, most often focused on the electrical power measurement of a whole building. These systems are implemented and utilized by electrical utility providers. Conventional metering infrastructure only facilitates monthly data collection for billing purposes, while the new AMI framework allows for sub-hourly electrical demand readings. These data are primarily used for demand characterization and billing, however, many additional uses are being discovered. A wide-range of studies have been completed in recent years to focus on a range of issues related to automatically extracting information from these data using unsupervised techniques. In this section, three sub-categories of

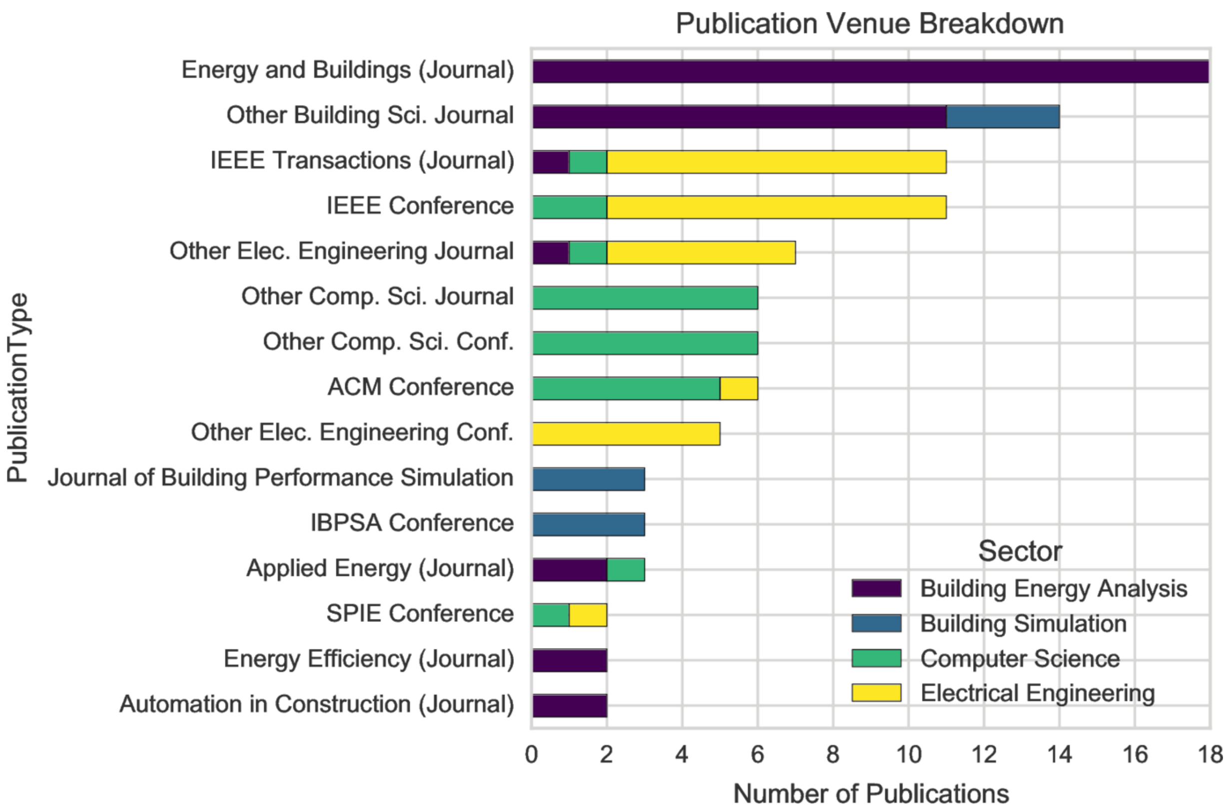

Figure 2.3: Breakdown of publications by publication type and research domain

application are discussed: load profiling, account classification, and disaggregation.

2.3.1 Load Profiling

Load profiling is the process of grouping temporal subsequences of measured energy data for the purpose of characterizing the typical behavior of an individual customer. It involves time-series clustering and feature extraction. Chicco et al. provide an original example in our review of this process using support vector machine clustering (Chicco & Ilie 2009). Gullo et al. and Räsänen et al. took the process further by introducing a framework of various clustering procedures that were implemented on case studies (Gullo *et al.* 2009; Räsänen & Kolehmainen 2009). Ramos et al., Iglesias et al., and Panapakidis et al. tested various conventional and new clustering methods and similarity metrics to determine those most applicable to electrical load profiling (Iglesias & Kastner 2013; Ramos *et al.* 2012; Panapakidis *et al.* 2015). Chicco et al. explored new clustering techniques based on ant colony grouping while Pan et al. discovered the use of kernel PCA for the same purpose (Chicco *et al.* 2013; Pan *et al.* 2015). Several groups of researchers such as Lavin and

Klabjan and Green et al. have found efficient use in using the core K-Means clustering algorithm for load profiling (Lavin & Klabjan 2014; Green *et al.* 2014). Shahzadeh et al. discussed the use of profiling as applied to forecast accuracy of temporal data (Shahzadeh *et al.* 2015). Two studies diverge from the standard profile development using clustering paradigm. The first is by De Silva et al. who uses Incremental Summarization and Pattern Characterization (ISPC) instead of clustering to find load profiles (De Silva *et al.* 2011). The other is the visual analytics-based approach of creating a smart meter analytics dashboard by Nezhad et al. to set up and inspect typical load profiles (Jarrah Nezhad *et al.* 2014).

2.3.2 Customer Classification

Automated account classification is the next sub-category that utilizes unsupervised learning techniques within the smart meter domain. These methods often employ load profile clustering as a first step but differentiate themselves in using those features to classify accounts, or buildings, that fit within various categories. Therefore, account classification is a type of manual semi-supervised analysis utilizing load profiling as a basis. The study by Figueiredo et al. harnessed K-Means and a labeled sample from accounts in Portugal to showcase this concept (Figueiredo *et al.* 2005). Verdu et al. and Räsänen et al. applied self-organizing maps (SOM) to accomplish a similar study that classifies accounts according to the applicability of several demand response scenarios (Verdu *et al.* 2006; Räsänen *et al.* 2008). Vale et al. give an overview of a general data mining framework focused on characterizing customers (Vale *et al.* 2009). Florita et al. diverge from the use of measured data by creating a massive amount of simulation data of load profiles to quantify energy storage applications for the power grid (Florita *et al.* 2012). Fagiani et al. use Markov Model novelty detection to automatically classify customers who potentially have leakage or waste issues (Fagiani *et al.* 2015). Cakmak et al. and Liu et al. test new visual analytics techniques within more holistic analysis framework for analyzing customers (Çakmak *et al.* 2014; Liu *et al.* 2015). Borgeson used various clustering and occupancy detection techniques to analyze a large AMI data set from California (Borgeson 2013). Bidoki et al. tested various clustering techniques to evaluate applicability for customer classification (Bidoki *et al.* 2010). A recent study in Korea develops a new clustering method for segmenting customers to analyze demand response incentives (Jang *et al.* 2016).

2.3.3 Disaggregation

The last area of smart meter data analysis is the field of meter disaggregation. Disaggregation attempts decompose a measurement signal from a high level reading to the individual loads being measured. This domain is well-researched from a supervised model perspective but recent attempts at unsupervised, pattern-based disaggregation were developed to facilitate implementation on unlabeled smart meter data. Shao et al. use Dirichlet Process Gaussian Mixture Models to find and disaggregate patterns in sub-hourly meter data (Shao *et al.* 2013). Reinhardt and Koessler use a version of symbolic aggregate approximation (SAX) to extract and identify disaggregated patterns for the purpose of prediction (Reinhardt & Koessler 2014). These studies are also unique in that few of the disaggregation studies focus on commercial buildings as opposed to residential buildings.

2.4 Portfolio Analytics

Portfolio analysis is a domain in which a large group of buildings, often located in the same geographical area or owned or managed by the same entity, are analyzed for the purpose of managing or optimizing the group as a whole. Each subsection covers the publications reviewed in this domain that fall into three categories: characterization, classification, and targeting.

2.4.1 Characterization

Publications that address the characterization of a portfolio of buildings include unsupervised techniques meant to evaluate and visualize the range of behaviors and performance of the group. A majority of the techniques utilized are either clustering or visual analytics that provide a model of exploratory analysis that enable further steps. Seem produced an influential study that extracts days of the week with similar consumption profiles (Seem 2005). Further clustering work was completed by An et al. to estimate thermal parameters of a portfolio of buildings (An *et al.* 2012). Lam et al. used Principal Component Analysis to extract information about a group of office buildings (Lam *et al.* 2008). Approaches focused on visual analytics and dashboards were completed by Agarwal et al., Lehrer, and Lehrer and Vasudev (Agarwal *et al.* 2009; Lehrer 2009; Lehrer & Vasudev 2011). Granderson et al. completed a case study-based evaluation of energy information systems, in which some methods combine some unsupervised approaches with visualization (Granderson *et al.* 2010). Diong et al. completed a case study as well focused on a specific energy information system implementation (Diong *et al.* 2015). Morán et al. and Georgescu and Mezic developed hybrid methods that employed visual continuous maps and Koopman Operator methods respectively to visualize portfolio consumption(Morán *et al.* 2013; Georgescu & Mezic 2014). Miller et al. completed two studies focused on the use of screening techniques to automatically extract diurnal patterns from performance data and use those patterns to characterize the consumption of a portfolio of buildings (Miller & Schlueter 2015; Miller *et al.* 2015). Yarbrogh et al. used visual analytics techniques to analyze peak demand on a university campus (Yarbrough *et al.* 2015).

2.4.2 Classification

The concept of classifying buildings within a portfolio supplements the characterization techniques by assigning individual buildings to subgroups of relative performance for the purpose of benchmarking or decision-making. Santamouris et al. produced a report using clustering and classification to assign schools in Greece to subgroups of similar performance (Santamouris *et al.* 2007). Nikolaou et al. and Pieri et al. further extended this type of work to office buildings and hotels(Nikolaou *et al.* 2012; Pieri *et al.* 2015). Heidarinejad et al. released an analysis of clustered simulation data to classify LEED-certified office buildings (Heidarinejad *et al.* 2014). Ploennigs et al. created a platform for monitoring, diagnosing and classifying buildings and operational behavior within a portfolio to quickly visualizing the outputs (Ploennigs *et al.* 2014).

2.4.3 Targeting

Targeting is a concept that builds upon characterization and classification to identify specific buildings or measures to be implemented in a portfolio to improve performance. These publications are differentiated in that specific measures are identified in the analysis. Sedano et al. use Cooperative Maximum-Likelihood amongst other techniques to evaluate the thermal insulation performance of buildings (Sedano *et al.* 2009). Gaitani et al. used PCA and clustering to target heating efficiency in school buildings (Gaitani *et al.* 2010). Bellala et al. used various methods to find lighting energy savings on a campus of a large organization (Bellala *et al.* 2011). Petcharat et al. also found lighting energy savings in a group of buildings (Petcharat *et al.* 2012). Cabrera and Zareipour used data association rules to complete a similar study to find wasteful patterns (Cabrera & Zareipour 2013). Geyer et al. and Schlueter et al. test various clustering strategies to group different buildings within a Swiss alpine village according to their applicability for retrofit interventions (Geyer *et al.* 2016) and thermal micro-grid feasibility (Schlueter *et al.* 2016).

2.5 Operations, Optimization, and Controls

Unsupervised techniques focused on individual buildings themselves are placed in the category for building operations, optimization, and control. This class contains the largest number of publications, and it incorporates a wider range of applications. It is differentiated from Section 2.6 in that the applications are not as focused on detecting and fixing the anomalous behavior. This section evaluates publications within the sub-categories of occupancy detection, retrofit analysis, controls, and energy management.

2.5.1 Occupancy Detection

Occupancy detection using unsupervised techniques infers human presence in a non-residential building without a labeled ground truth dataset or as part of a semi-supervised approach using a subset of labeled data. This occupancy detection is then used for analysis or as inputs for control of systems. Augello et al. used multiple techniques to infer occupant presence on a campus in Italy (Augello *et al.* 2011). Dong and Lam used Hidden Markov Models to detect occupancy patterns that were then used in a simulation (Dong & Lam 2011). Thanayankizil et al. developed a concept called Context Profiling in which occupancy was detected temporally and spatially (Thanayankizil *et al.* 2012). Mansur et al. used clustering to detect occupancy patterns from sensor data (Mansur *et al.* 2015). The newest studies by Adamopoulou et al. and D'Oca and Hong use a range of techniques to extract rules related to occupancy (Adamopoulou *et al.* 2015; D'Oca & Hong 2015). A recent study using wavelets illustrates the correlation of occupancy with actual energy consumption (Ahn & Park 2016).

2.5.2 Controls

Controls optimization is an enduring field of study aimed at creating a state of the best operation and energy performance for a building system such as heating, cooling, ventilation or lighting. Kusiak and Song created a means of optimally controlling a heating plant with clustering as a key step (Kusiak & Song 2008). Patnaik et al. completed studies focused on using motif detection to find modes of chilled water plant operation that proved most optimal (Patnaik *et al.* 2010, 2009). Hao et al. built upon these concepts to create a visual analytics tool to investigate these motifs (Hao *et al.* 2011). May-Ostendorp et al. used rule extraction as a means of enhancing a model-predictive control process of mixed-mode systems (May-Ostendorp *et al.* 2011, 2013). Bogen et al. used clustering to detect usage patterns for building control system evaluation (Bogen *et al.* 2013). Fan et al. used clustering to enhance chiller power prediction with the ultimate goal of control optimization (Fan *et al.* 2013). Hong et al. used Empirical Mode Decomposition to spatially optimize the placement of sensors in a building (Hong *et al.* 2013). Domahidi et al. used support vector machines (SVM) to extract optimized rules for supervisory control (Domahidi *et al.* 2014). Habib and Zucker use SAX to identify common motifs of an absorption chiller for the purpose of characterization and control (Habib & Zucker 2015).

2.5.3 Energy Management

Energy management and analysis of an individual building using unsupervised techniques is becoming common due to the increasing amounts of raw building management (BMS) and energy management system (EMS) data. Users of these techniques are often facilities management professionals or consultants who undertake the process to understand how the building is consuming energy. Duarte et al. use visual analytics to process data from an EMS along with various pre-processing techniques (Duarte *et al.* 2011). Lange et al. created two overview studies focused spatiotemporal visualization of building performance data and its interpretation in various case studies (Lange *et al.* 2012, 2013). Gayeski et al. completed a recent survey of operations professionals on their use of graphical interfaces of BMS and EMS dashboards (Gayeski *et al.* 2015). Outside of the visual analytics realm, Fan et al., Xiao and Fan, and Yu et al. completed studies of an entire data mining using framework using data association rules to improve operational performance (Fan *et al.* 2015b; Xiao & Fan 2014; Yu *et al.* 2013).

2.6 Anomaly Detection

Anomaly detection for buildings focuses on the detection and diagnostics of problems occurring within a building, its subsystems, and components. This field is most often focuses on the use of novelty detection or clustering approaches to find anomalous behavior. The sub-categories for this section are divided according to the spatial hierarchy of systems within a building; the highest level is whole building consumption, down to the subsystems such as heating, cooling or lighting and then to the individual components of those systems.

2.6.1 Whole Building

Whole building anomaly detection uses the electricity or heating and cooling energy supply in coming to a building to determine sub-sequences of poor performance. This category is complimentary to many of the Smart Meter solutions as they both focus on the use of a single data stream for a building. Seem had an early work again in this category with his work in using novelty detection to find abnormal days of consumption in buildings (Seem 2006). Liu et al. used classification and regression trees (CART) (Liu *et al.* 2010) and Wrinch et al. use frequency domain analysis for the same purpose (Wrinch *et al.* 2012). Jacob et al. utilized hierarchical clustering to use as variables in regression models for whole building monitoring (Jacob *et al.* 2010). Fontugne et al. created a process known as the *Strip, Bind, and Search* method to automatically uncover misbehavior from the whole building level and subsequently detects the source of the anomaly (Fontugne *et al.* 2013b). Janetzko et al. developed a visual analytics platform to highlight anomalous behavior in power meter data (Janetzko *et al.* 2013). Chou and Telaga created a hybrid whole building anomaly detection process using K-means (Chou & Telaga 2014). Ploennigs et al. and Chen et al. created similar systems that use generalized additive models (GAM) (Ploennigs *et al.* 2013; Chen *et al.* 2014). In the most recent work, Capozzoli et al. and Fan et al. use various techniques as part of a framework to detect and diagnose performance problems (Capozzoli *et al.* 2015; Fan *et al.* 2015a).

2.6.2 Subsystems

Subsystem anomaly detection focuses on the use of a broader data set to detect and diagnose faults from a lower level. Yoshida et al. provided a semi-supervised approach that seeks to determine which variables within a building are most influential in contributing to overall building performance (Yoshida *et al.* 2008). Wang et al. use PCA to diagnose sensor failures (Wang *et al.* 2010). Forlines and Wittenberg visualized multi-dimensional data using what they call the Wakame diagram (Forlines & Wittenburg 2010). Linda et al. and Wijayasekara et al. use various techniques to diagnose system faults and visualize them spatially (Linda *et al.* 2012; Wijayasekara *et al.* 2014). Le Cam et al. use PCA to create inverse models to detect problems in HVAC systems (Le Cam *et al.* 2014). Li and Wen created a similar process using PCA in conjunction with wavelet transform (Li & Wen 2014). Sun et al. used data association rules to create fault detection thresholds for finding anomalies (Sun *et al.* 2015).

2.6.3 Components

Component level anomaly detection is a bottom-up fault detection approach that focuses on determining faults in individual equipment. Wang and Cui use PCA to detect component faults in chilled water plants (Wang & Cui 2005). Yu et al. and Fontugne et al. both compliment their work at the whole building level to find associated component performance anomalies automatically (Yu *et al.* 2012; Fontugne *et al.* 2013a). Zhu et al. use wavelets to diagnose issues in air handling units (AHU) (Zhu *et al.* 2012).

2.7 Discussion

Several challenges facing the use of unsupervised machine learning in building performance were uncovered through this process of review. The first relates to the effect of several traditional research sectors exploring techniques targeted on the improvement of building performance. It was found that different sets of terminology are used to describe similar concepts. For example, in the building energy analysis field, the term *fault* (such as (Zhu *et al.* 2012)) is used to describe a situation that is similar to what is labeled an *anomaly* in the data mining domain (such as (Fontugne *et al.* 2013a)). Thus, discussions between these fields are restrained and completing a review of knowledge is difficult.

A critical issue related to differences in domains is the inconsistency of success objectives. Often individual papers would discuss the accuracy or efficiency of the algorithm or technical process itself (such as (Iglesias & Kastner 2013)), while others focused exclusively on the end results of the evaluation such as how much energy was saved (such as (Seem 2006)). Several examples publications successfully address both types of issues. For example, Ploennings et al. published studies which both addressed the applicability of generalized additive models and discussed their implementation in a platform that is applied to real buildings (Ploennigs *et al.* 2013, 2014). Researchers should strive to optimize in both the theoretical and practical domains to have the most impact on real buildings.

Another observation relates to the lack of easy reproducibility amongst studies. Reproducibility provides the ability for a third-party researcher to easily recreate the results of a study through a release of the data or code developed. Recent prominent articles have outlined the importance of reproducibility in science ($_{j}o$ 2014) and the sharing of data and code to enhance this pursuit ($_{c}o$ 2014). The biomedical sciences research community is leading the way in this effort; editors from over 30 major journals, funding agency representatives, and scientific leaders from that field created guidelines for the enhancement of reproducibility ($_{j}o$ 2014). Research from the building performance analysis community should follow this lead, specifically on machine learning and other types of empirical analysis.

Another challenge discovered is the lack of clarity regarding which is the optimal technique for each application. For example, a number of studies were completed to test the ability of clustering techniques to group similar daily load profiles (Chicco & Ilie 2009; De Silva *et al.* 2011; Green *et al.* 2014; Gullo *et al.* 2009; Lavin & Klabjan 2014; Ramos *et al.* 2012; Shahzadeh *et al.* 2015). A researcher or analyst who is searching for the best technique can see a survey of implementations through these publications; however, it's hard for them to be compared against each other as each utilizes a different data set and incorporates different methodologies. An explanation of the amount of effort needed to implement a technique is missing in most studies as well. For example, to implement a certain algorithm on a potential use-case or data set, an analyst is interested in which parameters need to be tuned, what labeled ground truth data should be gathered, and what expertise is necessary for understanding and implementation. This lack of comparison stifles the ability to make conclusions about the efficiency, interpretability, and appropriateness of use of each algorithm.

This dissertation seeks to address each of these challenges through the development of a framework that bridges the gap between the building energy performance, computer science, and electrical engineering. This goal is accomplished through incorporation of many of the approaches and techniques found in this literature review on a large collected temporal data set from buildings. A library of techniques, both mainstream and newly developed, are implemented on these data. This library is implemented on a collected and open data set. These techniques and data are to be shared with a wider audience through various means of reproducible research to be outlined in the methodology and conclusion sections.

3 Methodology

As discussed in Section 1, a two-step process is presented as a means of extracting knowledge from whole building electrical meters. Figure 3.1 illustrates the intermediate steps in each of the phases.

The first step is to extract temporal features that produce quantitative data to describe various phenomenon occurring in the raw temporal data. This action is intended to transform the data into a more human-interpretable format and visualize the general patterns in the data. In this step, the data are extracted, cleaned, and processed with a library of temporal feature extraction techniques to differentiate various types of behavior. This library is outlined in Sections 4-6. These features are visualized using an aggregate heat map format that can be used evaluated according to expert intuition, comparison with design intent metrics, or with outlier detection. Section 3.1 gives a more detailed definition of temporal features and how they're utilized in this study.

The second step is focused on the characterization of buildings using the temporal features according to several objectives. This step allows an analyst to understand the impact each feature has upon the discrimination of each objective. Five test objectives are implemented in this study: principal building use, performance class, operations strategy, general industry class, and energy savings measure success. One of the key outputs of this supervised learning process is the detection and discussion of what input features are *most important* in predicting the various classes. This approach gives exploratory insight into what features are important in determining various characteristics of a particular building amongst a large set of its peers. These metadata are building blocks for many other techniques such as benchmarking, diagnostics and targeting. The motivation for choosing these particular objectives centers around the consistently available meta-data from the collected case study data and their relation to various other techniques in the building performance analysis domain. These topics are covered through qualitative discussion with several of the operations teams on the campuses where the data were collected and is discussed more thoroughly in Section 7.

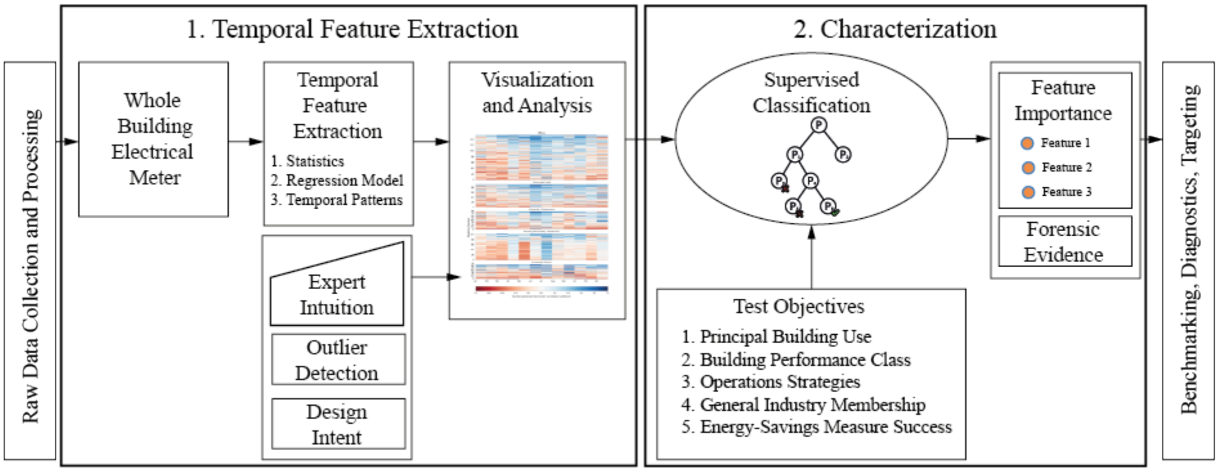

Figure 3.1: Overview of Data Screening Framework

3.1 Temporal Feature Extraction

Feature extraction is an essential process of machine learning and is the means by which objects are described quantitatively in a way that algorithms can differentiate between different types or classes. Figure 3.2 illustrates a hierarchical node diagram of the features, or metadata, about a building that is often necessary to accumulate to perform conventional analysis from the literature. Much of these data are needed when creating an energy simulation model, when setting thresholds for automated fault detection and diagnostics, or benchmarking a building. When performing analysis on a single building, these meta-data might be easy to accumulate. However, when such a process is scaled across hundreds or potentially thousands of buildings, a collection of these data is not a trivial procedure.

Modern, whole building electrical meters measure and report raw, sub-hourly, time-stamped data. Significant amounts of essential information can be extracted from temporal data to characterize a commercial building. The harvest of this information can assist in the implementation of conventional analysis techniques, as inputs to classify or benchmark a building, or to predict whether a building is a good candidate for individual energy savings measures. To extract information solely from these sensors, new features can be created from these raw data. These features are designated as temporal as they summarize behavior occurring in time-series data. To illustrate the concept of temporal features qualitatively, Figure 3.3 shows four example hourly electrical meters from different buildings. Even to the untrained eye, these data streams show obvious differences in the way each building operates. Building A seems to be an extremely consistent consumer of en-

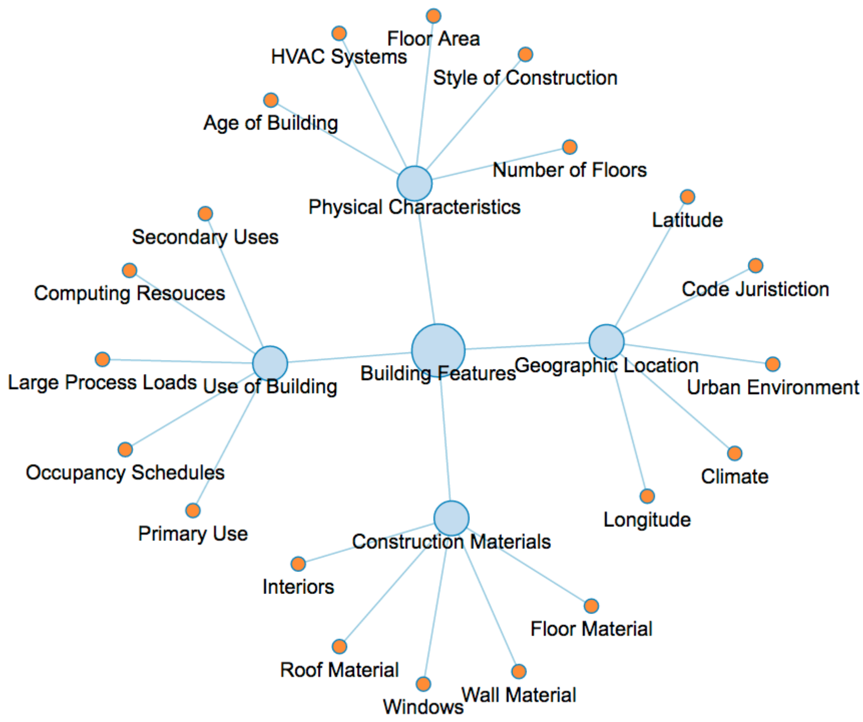

Figure 3.2: Conventional features, or metadata, about a building

ergy across the entire year. There are no steady-state shifts in operation and seemingly no influence from outside factors. Qualitatively, this data stream can be thought of as *consistent* or *predictable*. Building B is similar in operation but has an obvious influence from an external factor in the summer months. It is safely assumed that the consumption of this building is weather-dependent, and it has some kind of cooling system. Building C illustrates behavior that has *shifts* in consumption over the course of the year. This observation implies that this building has different schedules over the course of a year. Building D seems to have combinations of all of these attributes, with no obviously dominating phenomena.

Figure 3.4 illustrates the same four buildings with the time range constrained to two weeks of data. Short-term temporal effects at the weekly and daily level are now observed. Building A still appears very consistent with a predictable daily cyclical pattern and a few variations around August 4 and 5. Building B exhibits similar behavior, but with noticeable weekend differences on Saturdays and Sundays. Building C has less observable daily patterns but has a trend upwards in the last five days of the time range. Building D, again, has a combination of these attributes.

The goal of temporal feature extraction and analysis is to use various techniques to con-

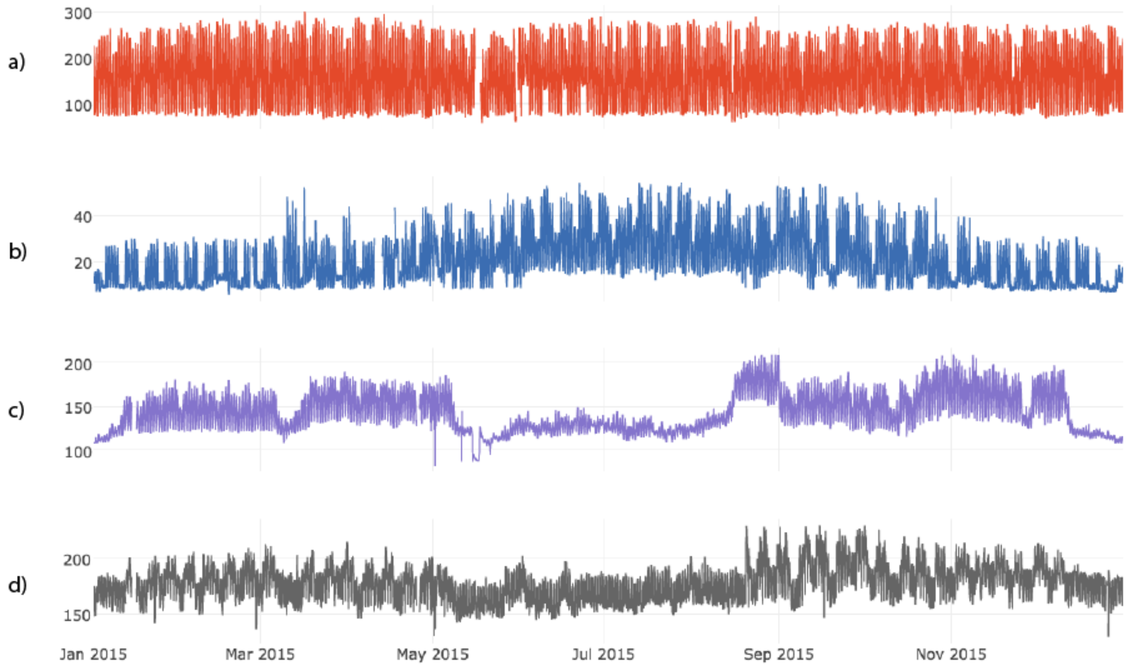

Figure 3.3: One year of example whole building electrical meter data that qualitatively exemplifying various temporal features

vert all these *qualitative* terms into a *quantitative* domain. For example, the descriptor *weather-dependency* can be quantified through the use of the Spearman rank order correlation coefficient with outdoor air temperature. Consistency or volatility of daily, weekly, or annual behavior can be quantified using various pattern recognition techniques. The primary focus of this study is to create and apply some temporal feature extraction techniques on commercial buildings for the purpose of characterization. Figure 3.5 illustrates the categories of temporal features created in this effort.

Temporal features are aggregations of the behavior exhibited in time-series data. They are characteristics that summarize sensor data in a way to inform an analyst through visualization or to use as training data in a predictive classification or regression model. Feature extraction is a step in the process of machine learning and is a form of dimensionality reduction of data. This process seeks to quantify various qualitative behaviors. This section provides and overview of the categories of temporal features extracted from the case study building data, the methods used to implement them, and visualized examples of a selected subset of features manifest themselves over a time range. Table 3.1 gives an overview the temporal features outlined in this section. A detailed list of the temporal

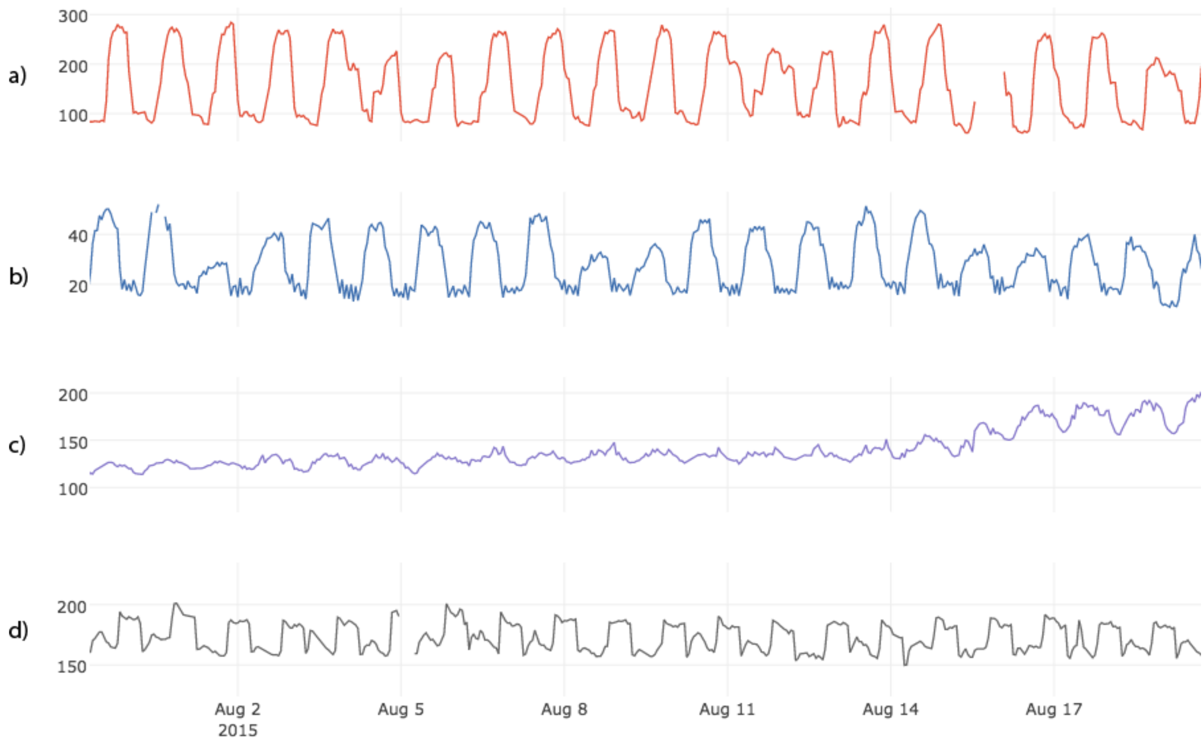

Figure 3.4: Two weeks of example whole building electrical meter data that qualitatively exemplifying various temporal features

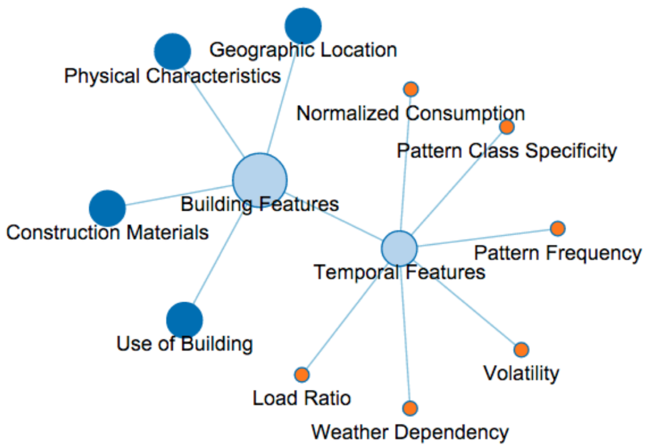

Figure 3.5: Temporal features extracted solely from raw sensor data

609 features created in this Section can be found in Appendix A.

Feature Category	**General Description**
Statistics-based	Aggregations of time series data using mean, median, max, min, standard deviation
Regression model-based	Development of a predictive model using training data and using model parameters and outputs to describe the data
Pattern-based	Extraction of frequent and useful daily, weekly, monthly, or long-term patterns

Table 3.1: Overview of feature categories

3.2 Characterization and Variable Importance

The primary goal of this dissertation is to get a better sense of what behavior in time-series sensor data is most characteristic of various *types* of buildings. As mentioned in the introduction, if this meta-data can be discriminated, the process of characterizing a building can be automated. In this section, the process of using random forest classification models and the input variable importance feature.

For each objective, several steps are taken to predict each objective and then to investigate the influence of the input features on class differentiation:

1. A random forest classification model is built using subsets of the generated features to predict the objectives class
2. The classification model provides an indication of the ability of the temporal features in describing the class based on its accuracy
3. Input feature importance is calculated by the classification model for insight on what the most informative features are in predicting class
4. An in-depth analysis comparison of two of the classes within each objective is completed to explore further the attributes that characterize a building

An overview of this process is found in Figure 3.6. After the technical analysis of the ability for the features sets to characterize building use type, a discussion is presented for each subsection on the practical insight gained from this process from discussions with the case study participants outlined in Section 3.3.1.

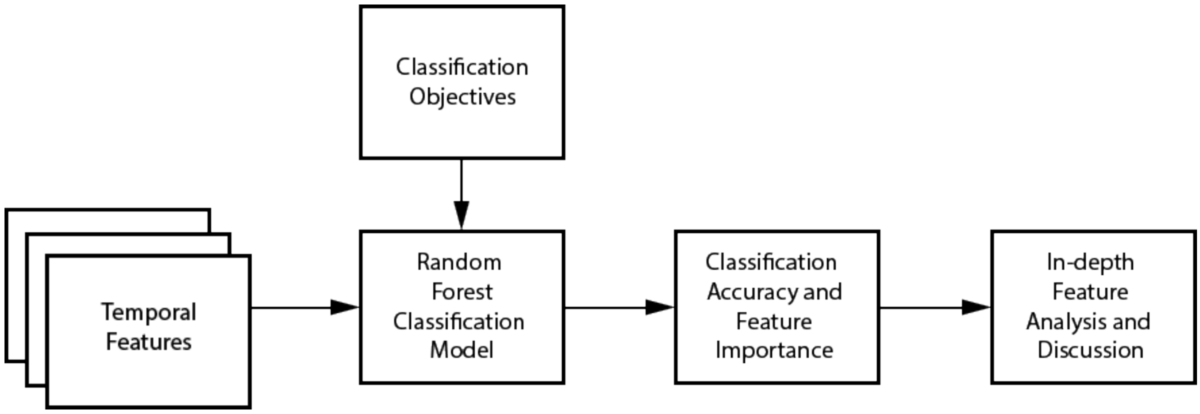

Figure 3.6: Characterization process to investigate the ability for various features to describe the classification objectives

Random forest classification models were chosen based on their ability to model diverse and large data sets in a robust way Breiman *et al.* (n.d.). These models use an ensemble of decision trees to predict various characteristic labels about each building based on its features. The literature describes decision trees as the "closest to meeting the requirements for serving as an off-the-shelf procedure for data mining" Hastie *et al.* (2009). Figure 3.7 illustrates an example of a decision tree using features to determine whether a patient is sick or healthy using two features Geurts *et al.* (2009).

Decision trees often over-fit data due to high variance. Random forest models work by creating a set of decision trees and averaging all of their predictions to overcome this variance. Figure 3.8 illustrates a set of four decision trees that is more accurately able to distinguish between the two classes than a single tree model.

Random forests use a form of cross-validation by training and testing each tree using a different bootstrapped sample from the data. This process produces an *out-of-bag error (OOB)* that acts as a generalized error for understanding how well each class can be predicted. This accuracy is used to determine how well the generated temporal features can delineate the class objectives. Random forests can also calculate the importance of the input features and how well they lend themselves to predicting the objectives. This attribute is useful in that it allows us to understand exactly which temporal features are most characteristic of various objectives. Variable importance is calculated using Equation 3.2.1. The importance of input feature X_m for predicting Y by adding up the weighted impurity decreases $p(t)\Delta i(s_t, t)$ for all nodes t where X_m is used, averaged over all N_T

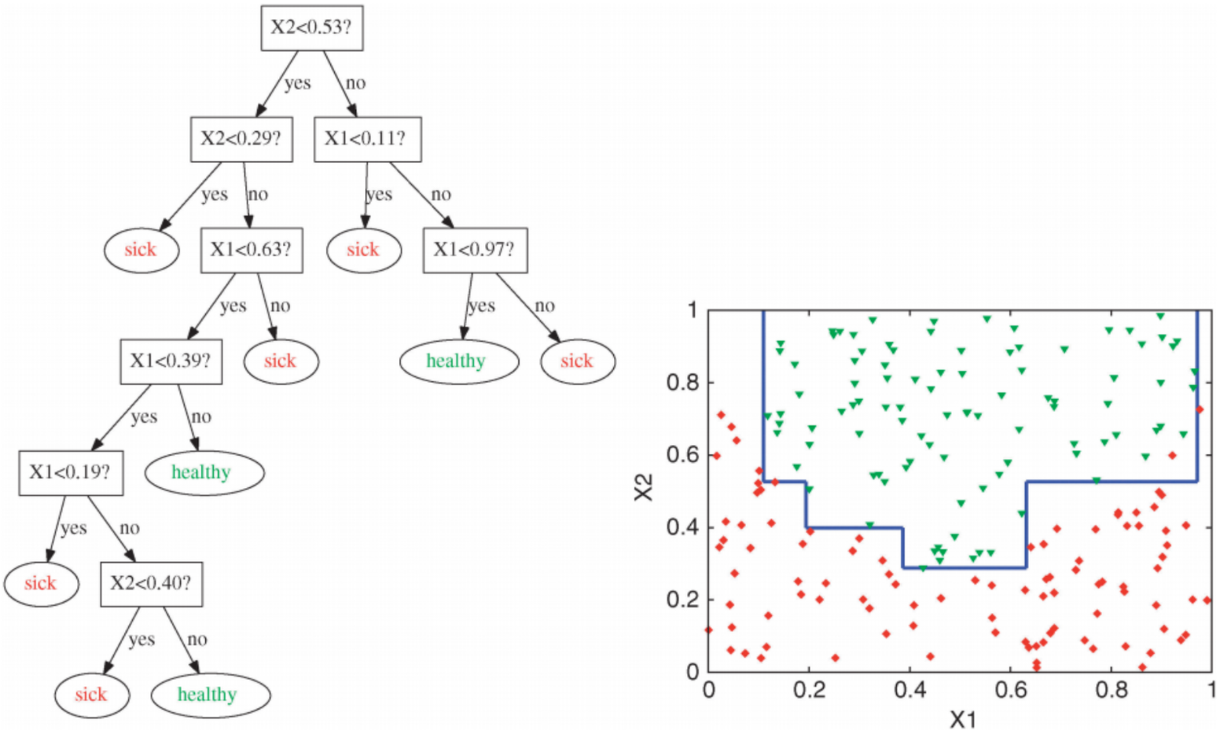

Figure 3.7: An example of a decision tree (left) with the decision boundary for two features, $X1$ and $X2$ (right). Adaption with permission from Geurts *et al.* (2009).

trees in the forest Louppe *et al.* (2013).

$$Imp(X_m) = \frac{1}{N_T} \sum_{T} \sum_{t \in T : v(s_t) = X_m} p(t) \Delta i(s_t, t) \tag{3.2.1}$$

3.3 Case Study, Empirical Data Collection, and Qualitative Research

One of the main goals of this research is the testing and implementation of the temporal feature extraction techniques on empirical sensor data collected from real buildings. Various raw data sets were obtained from case study buildings and campuses around the world to test the developed methods. The target of these interactions was to collect at least one year of hourly data from whole building electrical meters, resulting in at least 8760 measurements per building. Several of these data sets were collected through a series of site visits and interviews. These interactions are detailed in Section 3.3.1 by giving an

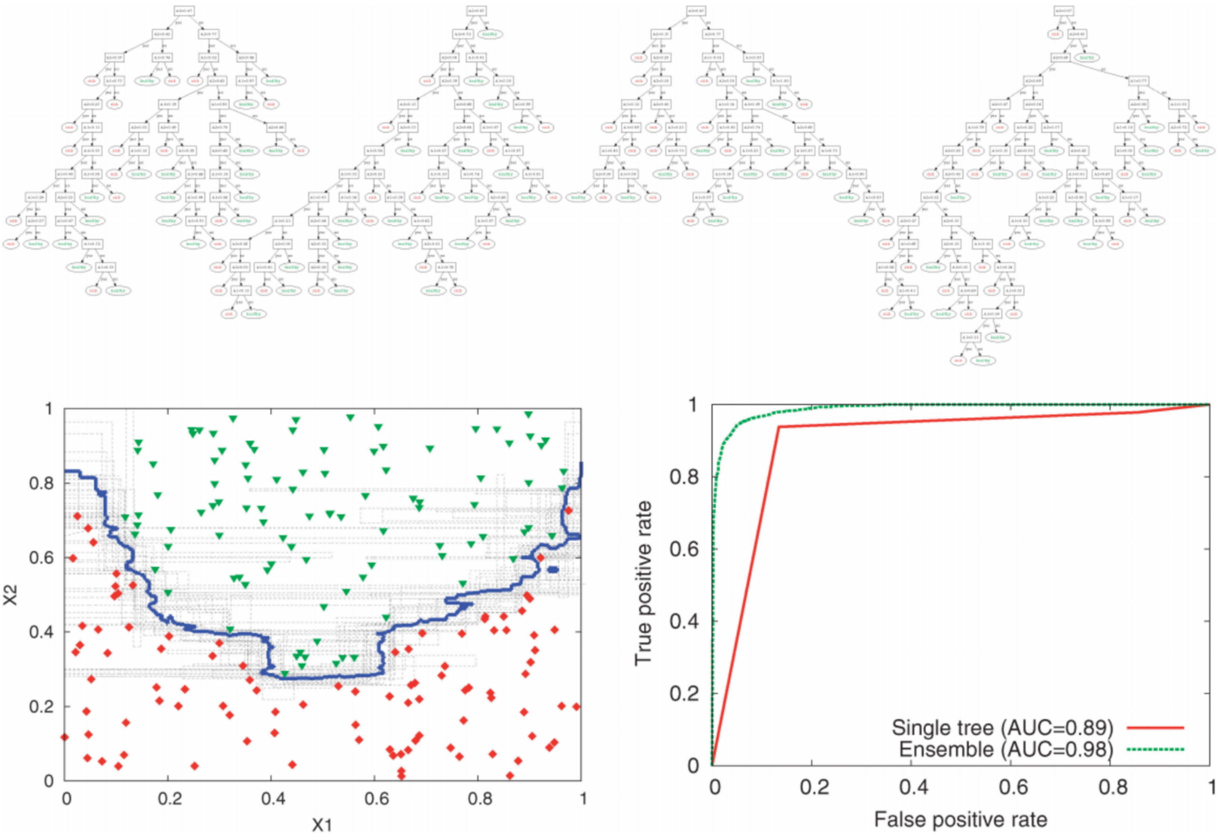

Figure 3.8: Ensemble of decision trees (top) that produces a more accurate decision boundary (lower left) and comparison with a single tree model (lower right). Adapted with permission from Geurts *et al.* (2009).

in-depth overview of these case studies by discussing the current performance data acquisition systems and the standard methods of utilizing those data for every day tracking activities. A key goal of the collection of these data was that they would be a basis for an open, shareable data repository for building performance research. This goal was discussed with the case study participants. Several other raw data sets were collected from open data sources on the Internet and were included in this study, albeit often with less metadata available. These case studies are described in Section 3.3.2.

In addition to the quantitative data collected from each of these case studies, qualitative feedback was gathered to get a better sense of *how useful* the implementation and interpretation of the framework would be in the day-to-day operations of various types of stakeholders. The results of these qualitative interviews are included in this Section 7.

3.3.1 Site Visits for Case Studies

Throughout the course of two years, from February 2014 to April 2016, several site visits were conducted to interview operations staff at seven campuses. The purpose of this effort was two-fold: first, to collect as much raw, temporal data from each site as possible and, second, to discuss the status quo of building energy analysis as performed on their campus. This section discusses these site visits, the types of data that were collected, and a few of the lessons-learned from the process. A consistent theme in the site visits was that each campus has been investing in electrical metering and data acquisition systems over the past decade. In every one of the case study interviews, the operations staff discussed the underutilization of the data being collected. A common phrase was, "We have more meter data than any time before, and we don't know what to do with it." Another common situation was that a campus had a large electrical metering infrastructure but did not know how to extract raw data for this research project. This scenario occurred on three of the seven campuses after the first interview, and data was still not available even after a follow-up visit on two of those campuses. Therefore, only four of the seven case studies had data available and will be discussed in the following subsections.

Case Study 1

The first case study is a campus in a continental climate in the Midwest region of the United States. It is a university with 226 buildings spread across two main campuses. Altogether, these buildings have a total floor area over 2.3 million square meters (25 million square feet). An initial interview was conducted with the lead statistician of the facilities management in March 2015. Information was gathered on the building and energy management systems of the campus and a discussion regarding the typical utilization of the data was conducted. It was found that there are over 480 electrical meters on the campus and that these data were primarily used for billing of the individual academic departments. They have a custom metering data management platform with some capabilities for data export. A second site visit was conducted in June 2015 to facilitate the collection of a sample one year data set. In this site visit, a facilities management professional with experience in SQL databases was able to directly query the underlying back-end of the energy management system to extract one year of raw data from all of the metering infrastructure on the campus. An accompanying meta-data spreadsheet was discovered that included information on floor area, primary space usage, EnergyStar score, and address. These data were then used for the analysis and feature extraction, and some of the results were compiled and presented to the entire facilities management department of this university in March 2016. This presentation gave an overview of the feature creation techniques and an understanding of how the buildings on their campus compare to other universities. More discussion on the feedback from this presentation are discussed in Section 7.

Case Study 2

The second case study is a campus in the Northeast region of the United States. It is also a University and it has 180 buildings on a single main campus. An initial meeting was organized in April 2015 with the facilities management team. This campus has well-organized building and energy management systems with a strong emphasis on data acquisition and management. The campus has an analytics and automated fault detection software platform that is connected to the underlying controls systems. A follow-up campus visit was conducted in August 2015 to facilitate the download of a raw, example data set from the buildings on campus. At this point, a log-in to a new data management platform was given for the purposes data extraction. Several issues arose from the use of this platform and ultimately, a database query by the software developers of the system was used to extract the one year of electrical meter data from the campus buildings. Once again, a spreadsheet of meta-data was shared that included information on floor area and primary building use type. A final site visit was conducted in April 2016 to discuss some of the results of the data acquisition and upcoming plans for upgrades. A formal presentation of the results was not able to be given; thus only limited feedback of the implementation progress was collected.

Case Study 3

The third case study is a campus in the Midwest region of the United States. Once again, it is a university campus with 25 buildings encompassing 204,000 square meters (2.2 million square feet) of floor space. An initial site survey and discussion of the campus was conducted in March 2015 with the campus lead mechanical and energy engineers. This campus has its electrical meters connected to a campus energy management platform that includes various visualizations and analytics techniques. This platform also can easily provide raw data download for analysis in this study. This platform resulted in this campus being by far the most user-friendly on data collection out of the case study set, including the open, on-line data sources. Raw data in flat files was easily downloaded for all data points at once. The meta-data for this campus was also extracted from this energy management platform, albeit in a more manual method from the user interface. A follow-up visit to this campus was conducted in March 2016 with initial results of characterizing the data according to a subset of the tested features. A significant amount of feedback for this case study was given by the facilities management department regarding the ability for these insights to assist in their decision-making processes.

Case Study 4

The fourth case study is an international school campus in tropical Southeast Asia. This campus includes five buildings with approximately 58,000 square meters (625,000 square feet). It was built and opened in 2010 and includes some sustainable design features such as an optimized chilled water plant, solar thermal cooling system, and an innovative, fresh air delivery system. The building management and data acquisition system have been a primary focus of the operations director of the campus for many years. Discussions and interviews with the operations staff have occurred numerous times over the course of the last five years. The key focus for this campus has been maintaining an optimized chilled water system. The operations team of this organization has been an active contributor to the development of the methodology.

Case Study 5

The final case study to be outlined in this section is a university campus located in Switzerland. This campus includes 22 building encompassing more than 150,000 square meters (1.6 million square feet). This campus has an energy management system with the ability to extract raw data, albeit only one point at a time. Data from this campus was utilized in a previous research project focused on campus and building-scale co-simulation and modeling. Only email correspondence with the campus facilities managers of this campus was conducted. A significant amount of meta-data was available from the facilities department through a spreadsheet that provided the breakdown of primary uses of the spaces in each building.

3.3.2 Online Open Case Studies

Several large data sets were found through a search of openly accessible data on-line. This section gives an overview of these data sources and the methods in which the data was

Source Name	**Description**	**Website**
Cornell University	EMCS Portal	http://portal.emcs.cornell.edu/
University of California - Berkeley	Berkeley Campus Energy Portal	http://berkeley.openbms.org/
Arizone State University	Campus Metabolism	https://cm.asu.edu
Carbon Culture	Community Open Data Platform	https://platform.carbonculture.net
EnerNOC	EnerNOC GreenButton Data	https://open-enernoc-data.s3.amazonaws.com/anon/index.html
University of Southampton	Open Data Service	http://data.southampton.ac.uk/

Table 3.2: Open, online data sources

extracted and pre-processed for analysis. Table 3.3.2 illustrates these sources, a short description of the platform in which the data was downloaded, and the URL of the platform. As in the site visit case studies, one year of hourly whole building electrical meter data was collected from each of these sources for as many buildings as possible.

3.4 Overview of Data Collected

Through data collection from the on-site case study interviews and on-line data sources, whole building electrical meter data from 1238 buildings was collected. Figure 3.9 illustrates the locations of these building around the world. A majority of the buildings are located in the United States, with the highest concentrations in the northeast region. A wide range of building types are included in the data set, from Education and Government to Agriculture and Heavy Industry.

From these groups of primary use types, the buildings are distributed across various time zone regions as seen in Figure 3.13. The east coast of the United States is the largest group due to the number of campuses and buildings from the EnerNOC data source. All of the buildings from the Carbon Culture data source are located in the United Kingdom.

Figure 3.11 and 3.12 illustrate the industries and sub-industries that the case study buildings are collected from. The number of university campuses is strongly evident in both

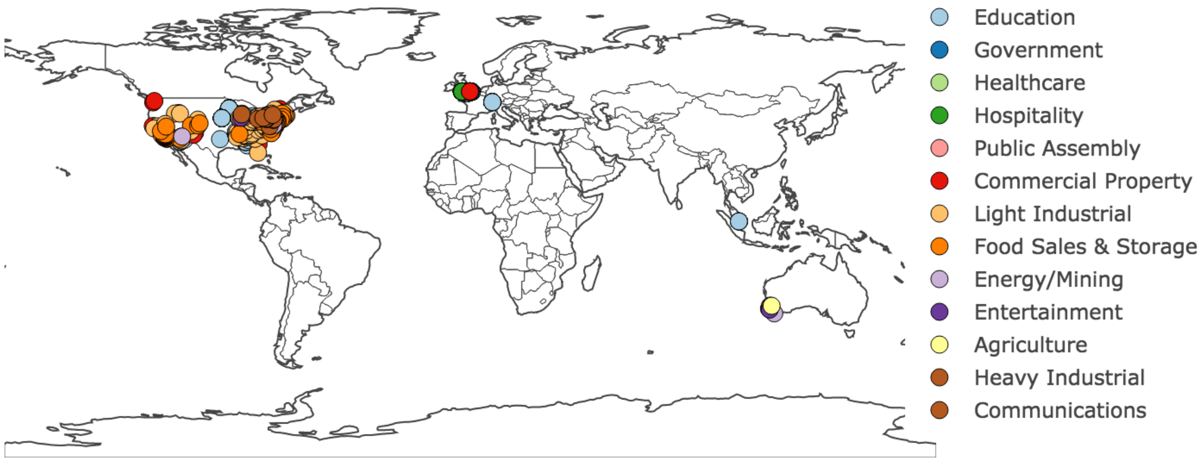

Figure 3.9: Locations of 1238 case study buildings collected from across the world

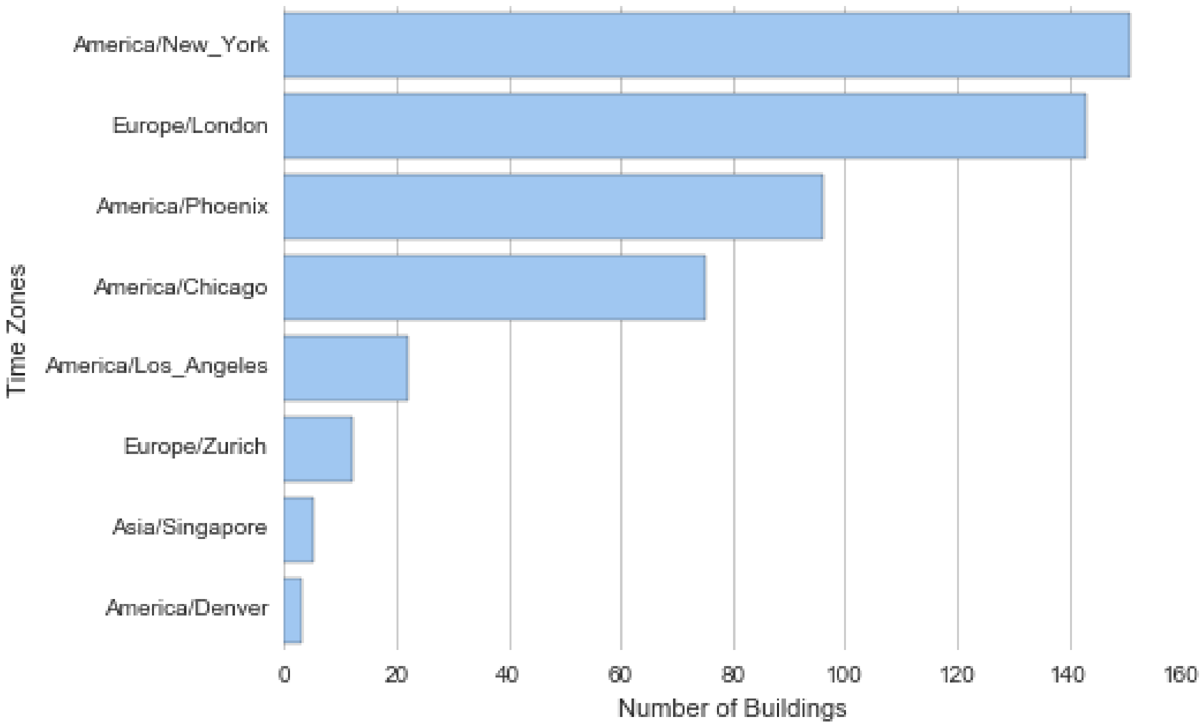

Figure 3.10: Distribution of case study buildings amongst time zones

784 charts.

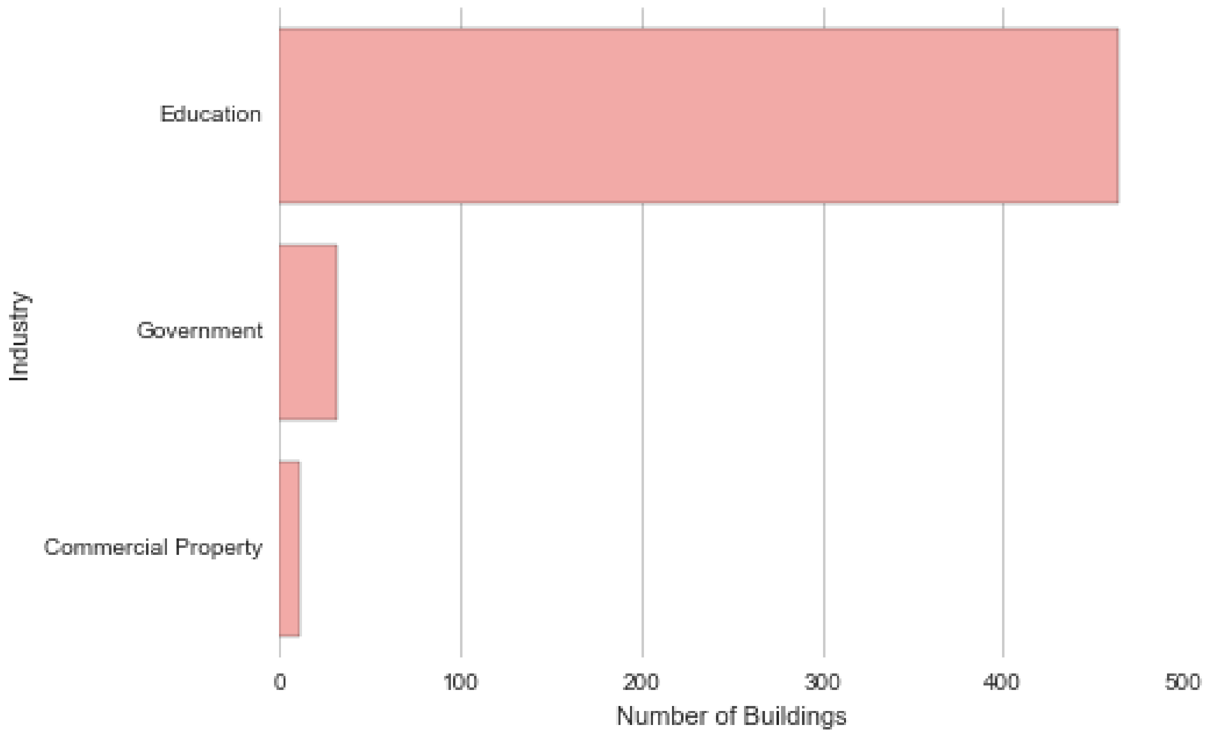

Figure 3.11: Distribution of case study buildings amongst general industries

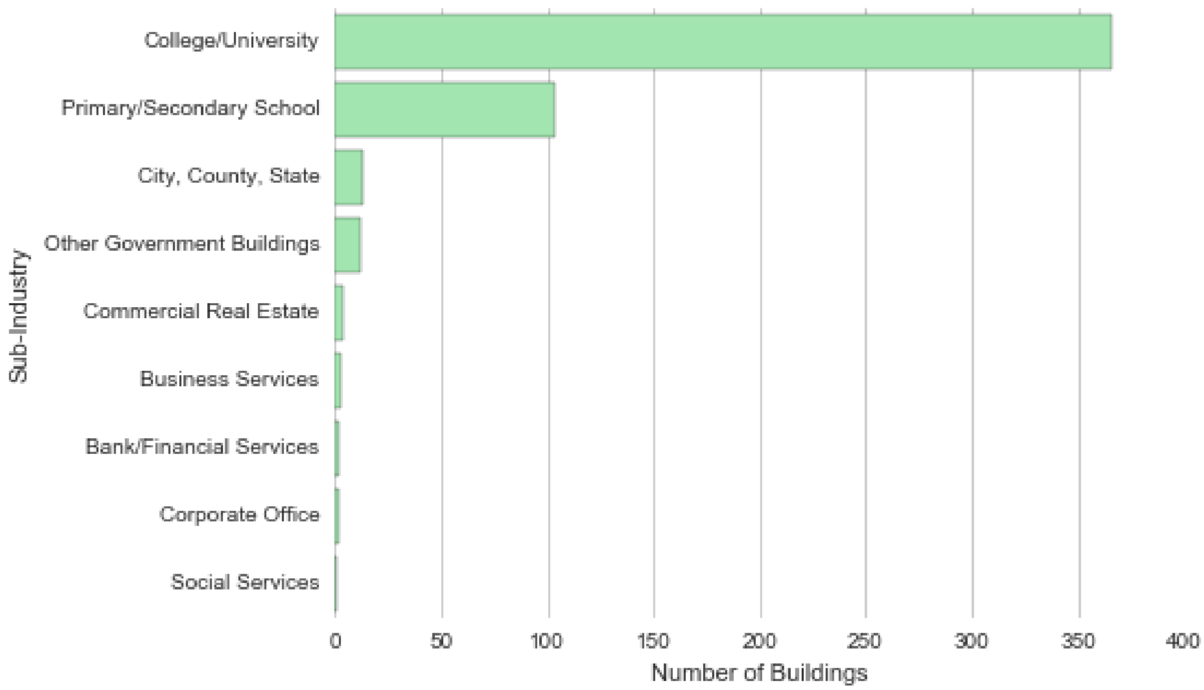

Figure 3.12: Distribution of case study buildings amongst sub-industries

3.4.1 Selection of Case Study Subset for Feature Implementation

A subset of buildings was chosen based on limiting criteria for inclusion in the implementation sections of this thesis. The primary consideration for inclusion is that the building is a member of one of the top primary use types: Offices, Primary/Secondary Schools, University Laboratories, University Classrooms, or Dormitories. These categories and the number of buildings in one are shown in Figure 3.13.

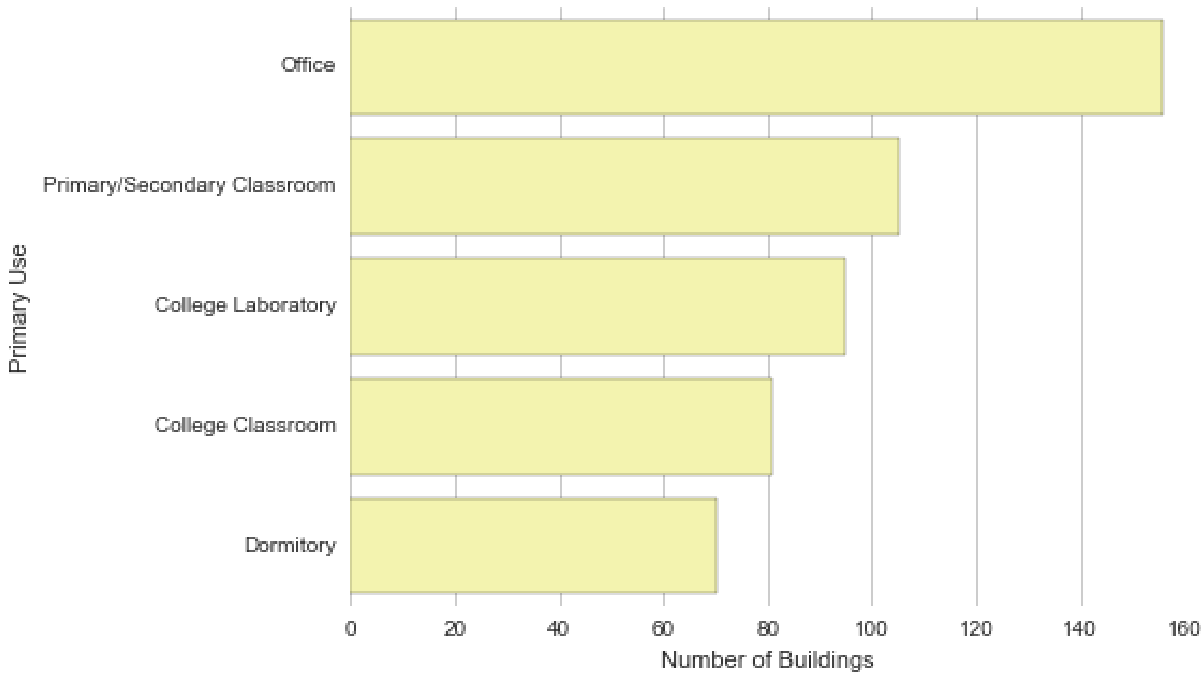

Figure 3.13: Distribution of case study buildings amongst primary space uses

3.5 Advanced Metering Infrastructure Case Study

A larger data set of almost 10,000 non-residential buildings is gathered in this thesis from an organization tasked with using the data to target buildings for performance improvement measures. These data are from an Advanced Metering Infrastructure (AMI) implementation. Different types of meta-data are available for these buildings, including industry and energy savings measure implementation. The primary goal of this data set is to provide a context of scalability on a larger data set. These data are strictly private and detailed data cannot be included in the methodology or development of the framework.

4 Statistics-based Features

Statistics-based temporal features are the first and most simplified category of temporal features developed. The main classes of features are basic temporal statistics, ratio-based, and the Spearman rank order correlation coefficient.

4.1 Theoretical Basis

4.1.1 Basic Temporal Statistics

The first set of temporal features to be extracted are basic statistics-based metrics that utilize the time-series data vector for various time ranges to obtain information using mean, median, maximum, minimum, range, variance, and standard deviation. Many of these features are developed through the implementation of the VISDOM package in the R programming language (Borgeson & Kwac 2015).As a simple example, if a time-series vector is described as X, with N values of $X = x_1, x_2, ..., x_n$, the most common statistical metric, mean (or μ), can be calculated using Equation 4.1.3 (Mitsa 2010).

$$\mu = \frac{\sum\limits_{i=1}^{N} x_i}{N} \tag{4.1.1}$$

The mean is taken not just for the entire time series, but also from the summer and winter seasons. The variance of the values are taken for the whole year, the summer and winter seasons as well. The variance of daily mean, minimum, and maximum values are determined to understand the breadth of values across the time range. Variance is calculated according to Equation 4.1.2.

$$\sigma^2 = \frac{\sum\limits_{i=1}^{n}(x_i - \mu)^2}{n} \tag{4.1.2}$$

The maximum and minimum electrical demand are calculated. Additionally, the hour and date at which the maximum demand occurs are determined to understand when peak consumption occurs. Additionally, the temperature at the maximum and minimum is account for weather influence. The 97th and 3rd percentiles are calculated to exclude any extreme outliers, a value that's often more useful than the maximum and minimum.

A series of hour-of-day (HOD) metrics are calculated that relate to aggregating the behavior occurring at each of hour the day-four metrics. The first of these calculates the most current hour of the top demand of the top 10% hottest days and the most common hour of the top 10% temperatures to inform roughly about cooling energy consumption. These metrics are repeated from the bottom 10% coldest days and temperatures. Another set of twenty-four metrics is calculated to account simply for the mean demand of each hour of the day.

A set of metrics is calculated individually for January and August to account for potential heating and cooling seasons. The daily maximum, minimum, mean, range and load duration are calculated for these seasonal periods. The complete list of these features can be found in Appendix A.

4.1.2 Ratio-based Statistical Features

The second major category of statistical features is ratio-based features. Simply, these are metrics in which two or more of the previously calculated statistical metrics are combined as a ratio. These features often have a *normalizing effect* in which buildings can be more appropriately compared to each other. The first extracted metric of this type is one of the most commonly calculated for building performance analysis: the consumption magnitude of electricity normalized by the floor area of the building. This metric seeks to provide a basis for comparison between buildings and is used as a key metric within numerous benchmarking and performance analysis techniques. Figure 4.1 illustrates a single building example of this metric per hour across a time range of two weeks at the end of the year. The top line chart of this figure shows the magnitude of hourly electrical consumption for one of the case study buildings. The middle portion of the figure repeats this information in the form of a color-based, one-dimensional heatmap. In this example, the daily weekday profiles manifest themselves as light-colored bands and weekend and unoccupied periods as darker bands. The color bar at the bottom of the figure is key in interpreting the color values. This figure is an example of a single building demonstration of this particular feature and is a type of graphic that is used throughout this entire section.

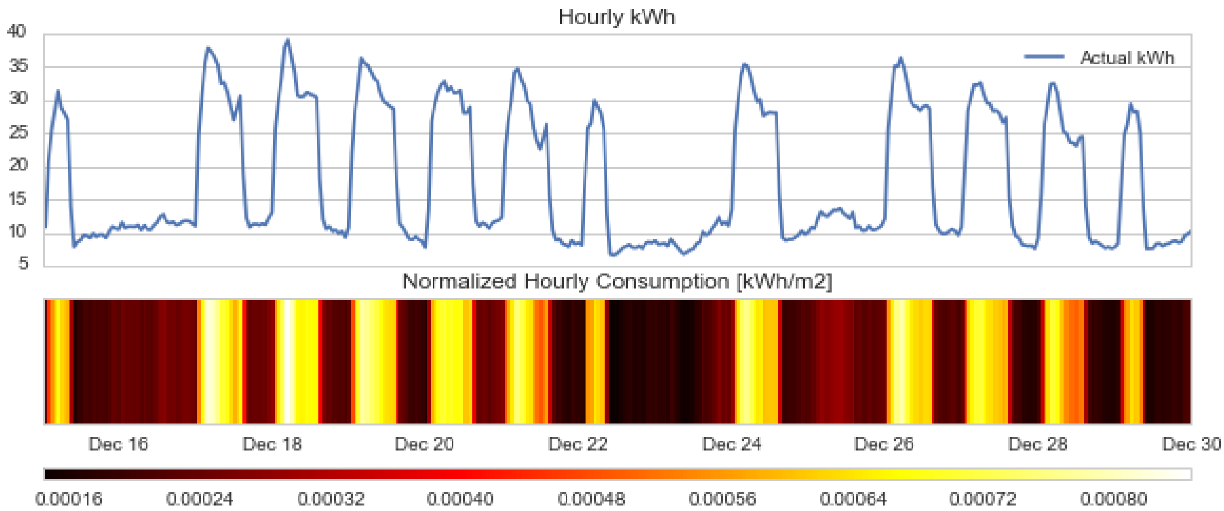

Figure 4.1: Single building example of area normalized magnitude

After normalized consumption, the first set of temporal features to be extracted are primary statistics-based metrics that utilize the time-series data vector for various time ranges to retrieve information using mean, median, maximum, minimum, range, variance, and standard deviation. If a time-series vector is described as X, with N values of $X = x_1, x_2, ..., x_n$, the most common statistical metric, mean (or μ), can be calculated using Equation 4.1.3 (Mitsa 2010).

$$\mu = \frac{\sum_{i=1}^{N} x_i}{N} \tag{4.1.3}$$

The median value of a vector is simply the middle value in an ordered set if the number of values is odd. If the length of the vector is even, then the median is the mean of the two middle values. The minimum and maximum values are the first and last in an ordered set. Vectors of values can also be described according to percentiles. Percentiles are cutpoints dividing the range of a probability distribution based on the percentage of values below a given threshold. For example, the value at the 95% percentile is found 95% of the way along an ordered set, with only 5% of the values remaining before reaching the maximum. In this section, aggregation ratios of many of these collection techniques are applied to the 24 hours from a single day to characterize various types of typical behavior quickly. The first example of these ratios is the minimum versus maximum ratio or load ratio. This rate is calculated by taking the daily minimum and dividing it by the daily maximum. Figure 4.2 illustrates a single building example of this ratio on one month of data from a case study building. These load ratios indicated whether a daily profile is more diverse, resulting in a lower load ratio, or more flat, resulting in a higher load ratio. In this example, weekends and holidays are a darker shade of blue as compared to generally-occupied weekdays. Load ratio can be used an indicator also of the relative magnitude of the unoccupied baseline. Buildings that have a lower average load ratio often have higher than average baselines, such as in laboratories or hospitals.

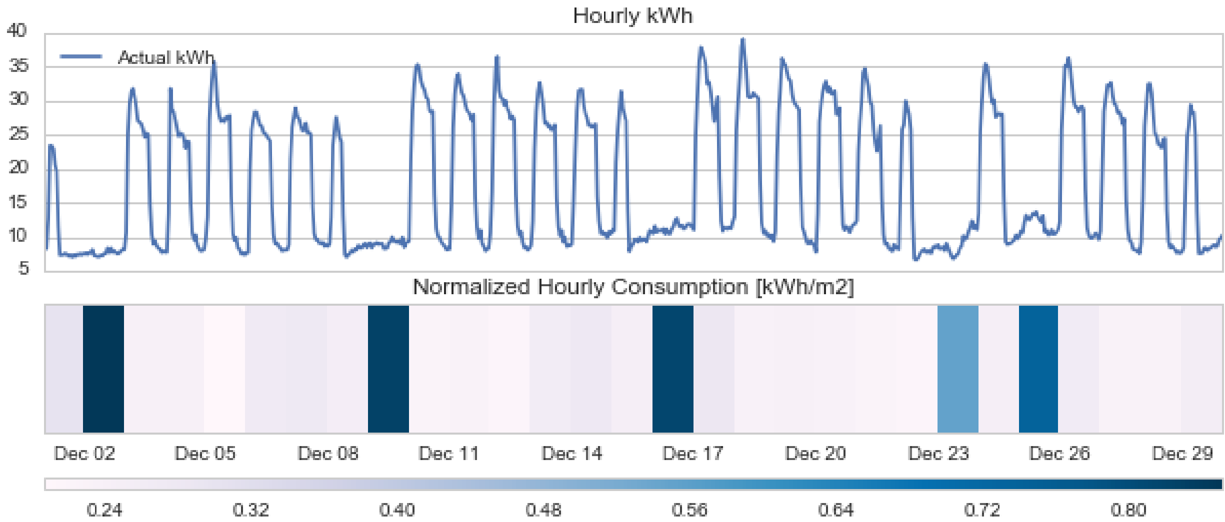

Figure 4.2: Single building example of the daily load ratio statistic

A library of similar load ratio daily metrics is designed and implemented on the case study buildings. These other ratios are daily mean versus maximum, minimum versus 95% percentile, and mean versus 95% percentile. The use of the 95% metric is mean to mitigate against outliers skewing the load ratios. These ratios are calculated on all days in the set, as well as just for weekend and weekdays. A full list of the features generated is found in Appendix A.

4.1.3 Spearman Rank Order Correlation Coefficient

Data stream influence characterization is the process of roughly classifying the dataset into streams and subsequences based on weather conditions sensitivity. A feature is developed in a study of evaluation of campus data for simulation feedback, and the following is a summarization of this technique (Miller & Schlueter 2015). This evaluation is important in understanding what measured performance is due to heating, cooling, and ventilation systems (HVAC) responses to outdoor conditions and what is due to schedule, occupancy, lighting, and different loading conditions which are weather independent. Performance data that is influenced by weather can be used to understand the HVAC system operation better or be weather-normalized to understand occupant diversity schedules.

The Spearman Rank Order Correlation (ROC) is used to evaluate the positive or negative correlation between each performance measurement stream and the outdoor air dry bulb temperature. This technique has been previously used for weather sensitivity analysis (Coughlin *et al.* 2009). The ROC coefficient, ρ, is calculated according to a comparison of two data streams, X, and Y, in which the values at each time step, X_i, and Y_i, are converted to a relative rank of magnitude, x_i and y_i, according to its respective dataset. These rankings are then used to calculate ρ that varies between +1 and -1 with each extreme corresponding to a perfect positive and negative correlation respectively. A value of 0 signifies no relationship between the datasets. This ρ value for a time-series is calculated according to Equation 4.1.4.

$$\rho = 1 - \frac{6\sum d_i^2}{n(n^2-1)} \tag{4.1.4}$$

The difference between the data stream rankings, x_i and y_i, is signified by a difference value, d_i, and the number of samples compared to each dataset is signified by n. Figure 4.3 illustrates the calculation of the ROC coefficient, ρ for three examples. The cooling sensitive data set shows a strong positive correlation between outside air temperature and energy consumption with a ρ value of 0.934. As the outside air temperature increases, the power consumption measured by this meter increases. The heating sensitive dataset shown has a strong negative correlation with a ρ of -0.68. A weather-insensitive dataset is shown in the middle which has a ρ of 0.0, signifying no weather relationship, which is evident due to the four levels of consumption which are independent of outdoor air conditions.

The correlation coefficient can be visualized for a single case as seen in Figure 4.4. The coefficient, in this case, is calculated individually for each month. This process results in twelve calculations of the metric using between 29-31 samples. In this case, consumption in January to May is noticeably more heating sensitive, a fact that can be observed clearly from the line chart, as well as the one dimension heat map. May to November is more cooling sensitive. It is interesting that September appears to be the most cooling sensitive month, a fact perhaps related to use schedules during that month. This coefficient is not

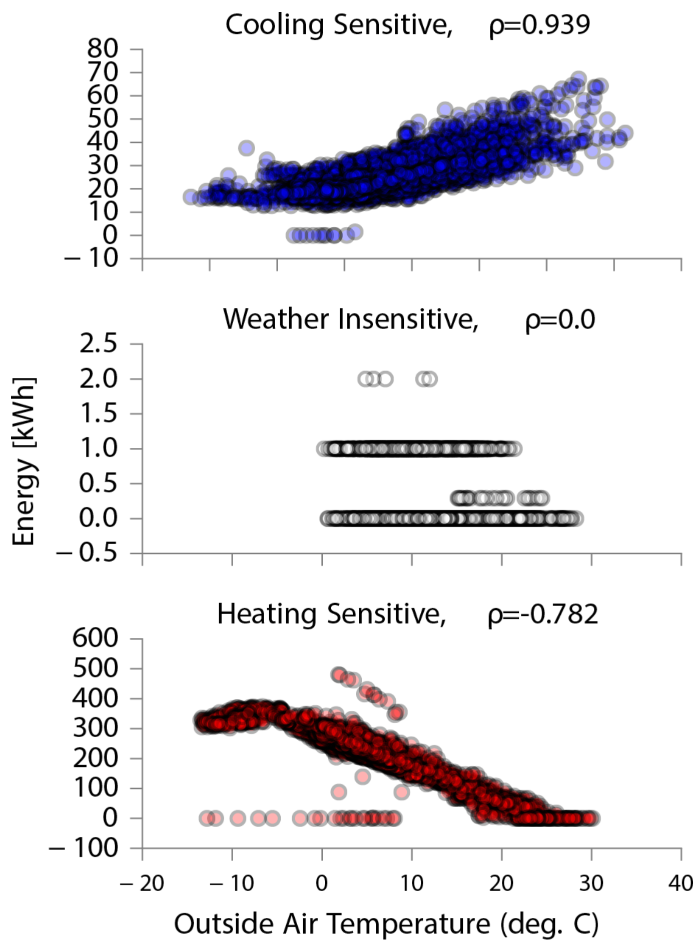

Figure 4.3: Weather sensitivity examples as energy vs. outdoor air temperature (from (Miller & Schlueter 2015))

a perfect indicator of HVAC consumption; it just detects a correlation. However, it is fast and easy to calculate and is the first phase of detecting weather dependency. More detailed and informative weather influence extraction features are investigated in Section 5.

4.2 Implementation and Discussion

Figure 4.5 illustrates the same normalized consumption metric as applied to all of the case study buildings. There are five segments of buildings based on the primary use types within the set: offices, university laboratories, university classrooms, primary/secondary schools, and university dormitories. These metrics are visualized in this way to understand

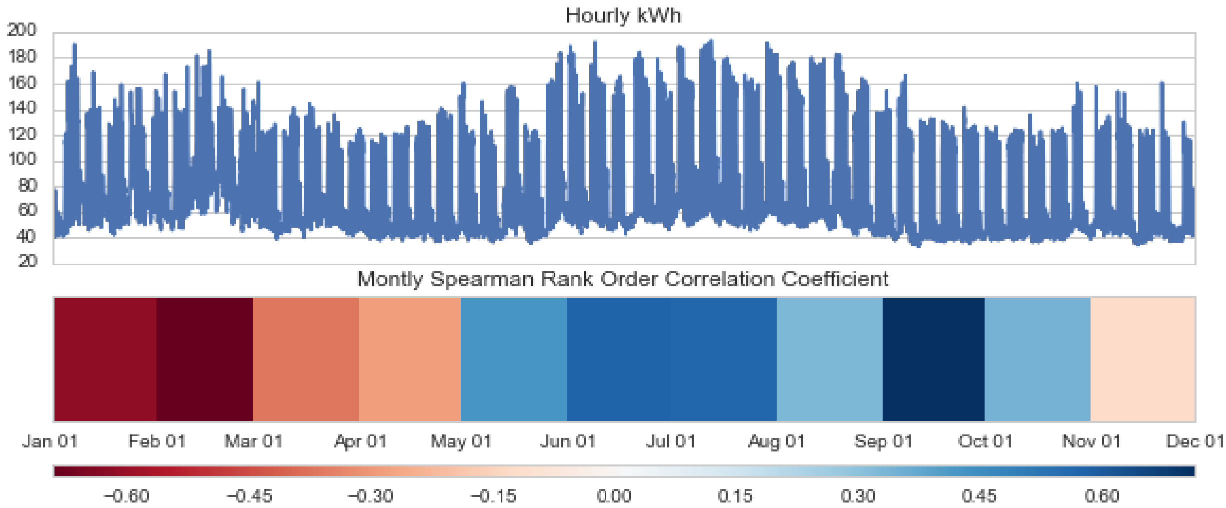

Figure 4.4: Single building example of the spearman rank order correlation coefficent with weather

the difference between each of these use types for each of the presented metrics. Each row of the heatmap for each segment is the values of the feature for a single building, while the x-axis is the time range for all buildings. Not all of the case study buildings have a January to December time range. For these cases, the data was rearranged so that a continuous set of January to December data is available to be visualized in the heat map. The aggregation metrics themselves are not calculated with this rearranged vector; it is only for visualization purposes. Like Figure 4.1, this type of graphic is used to visualize many of the temporal features in this section. From this metric in particular, one will notice that university labs have a systematically higher consumption over time as compared to the other use types. One will also see the dark vertical lines across the time range indicating weekend use as compared to the weekday. This particular pattern is absent from university dormitories due to their more continuous energy consumption.

Figure 4.6 illustrates this metric as applied to all case study buildings. As in the normalized magnitude, various patterns are more apparent including the weekday versus weekend phenomenon.

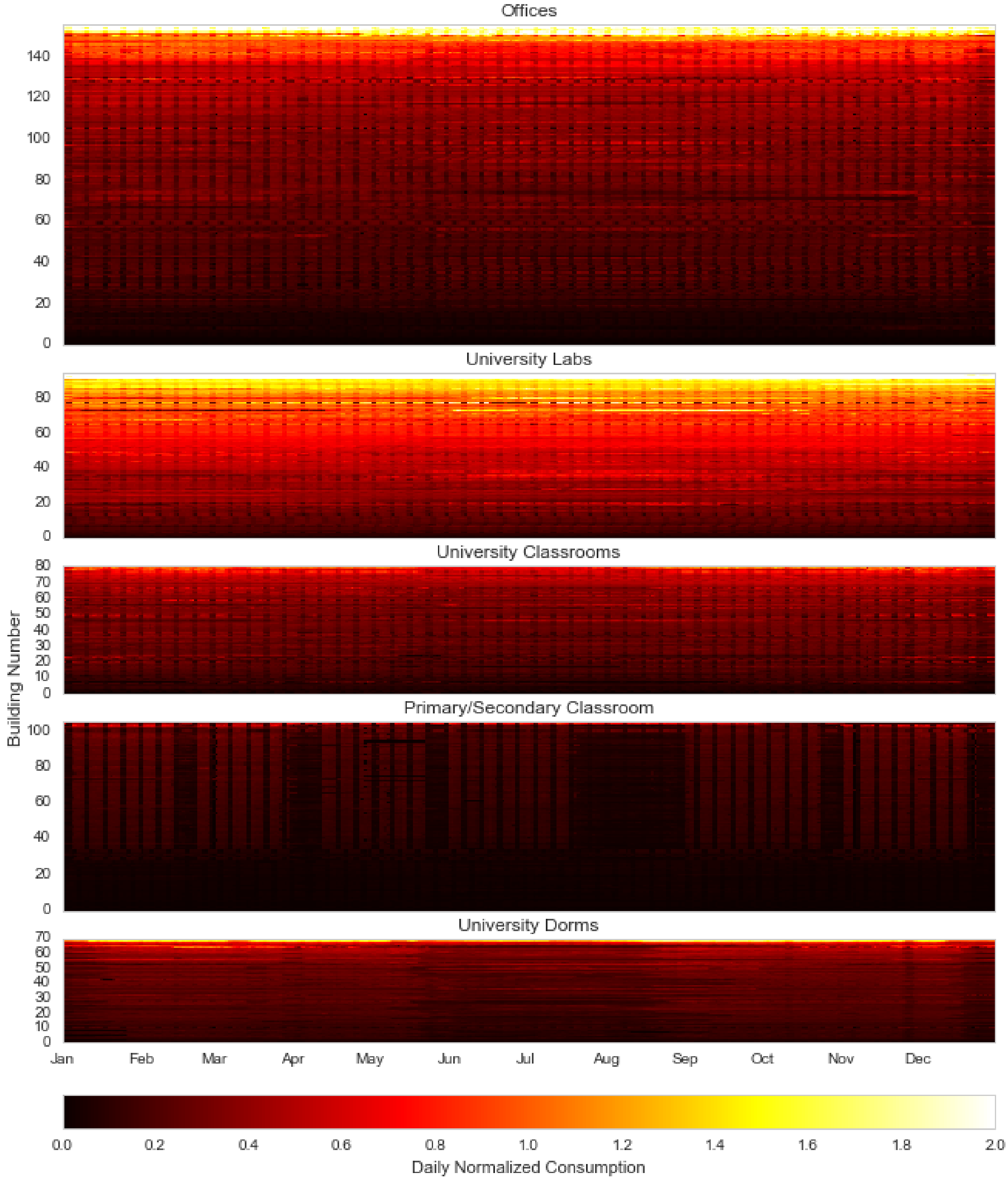

Figure 4.5: Heat map representation of normalized magnitude

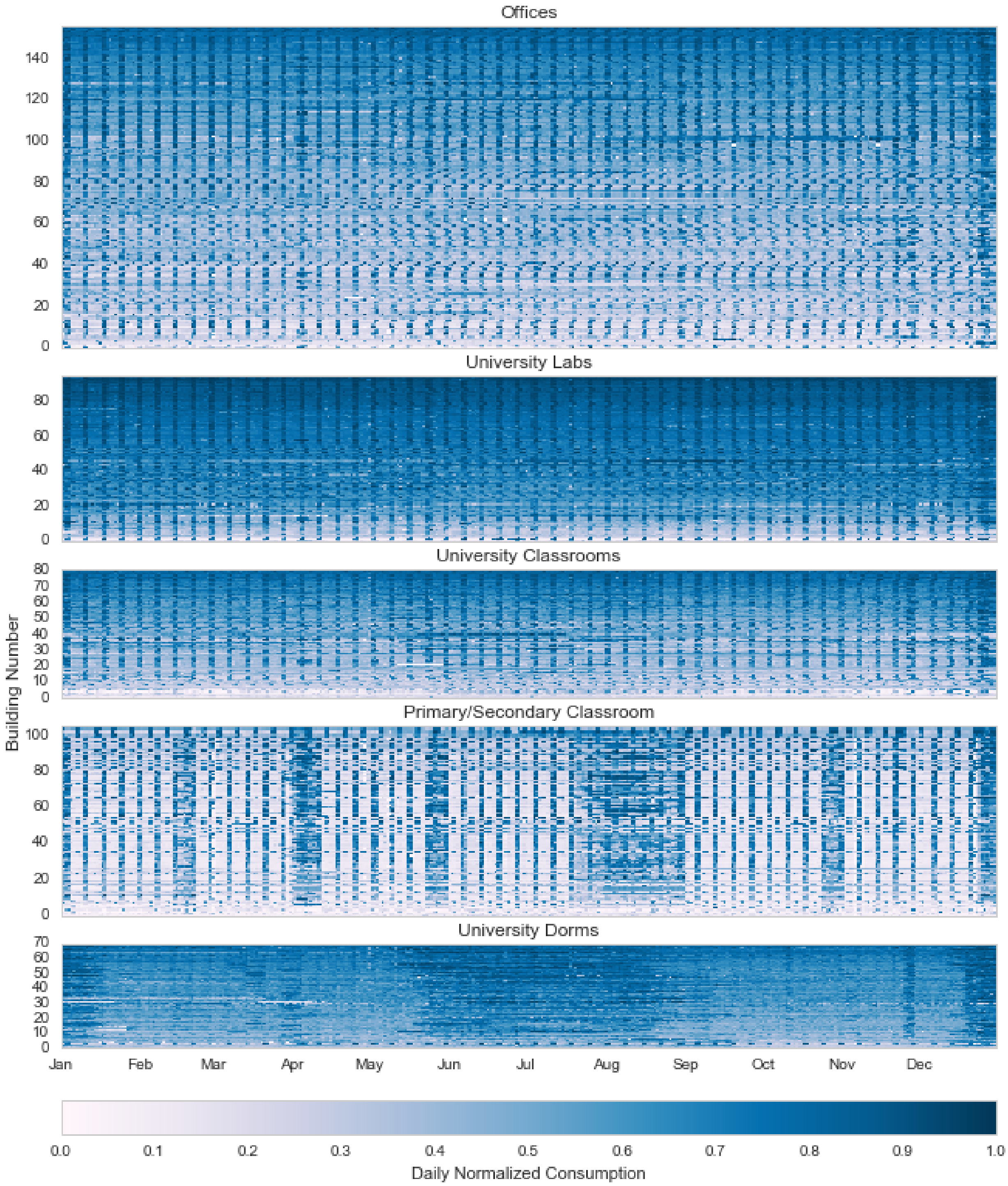

Figure 4.6: Heatmap of daily load ratio statistic for all case study buildings

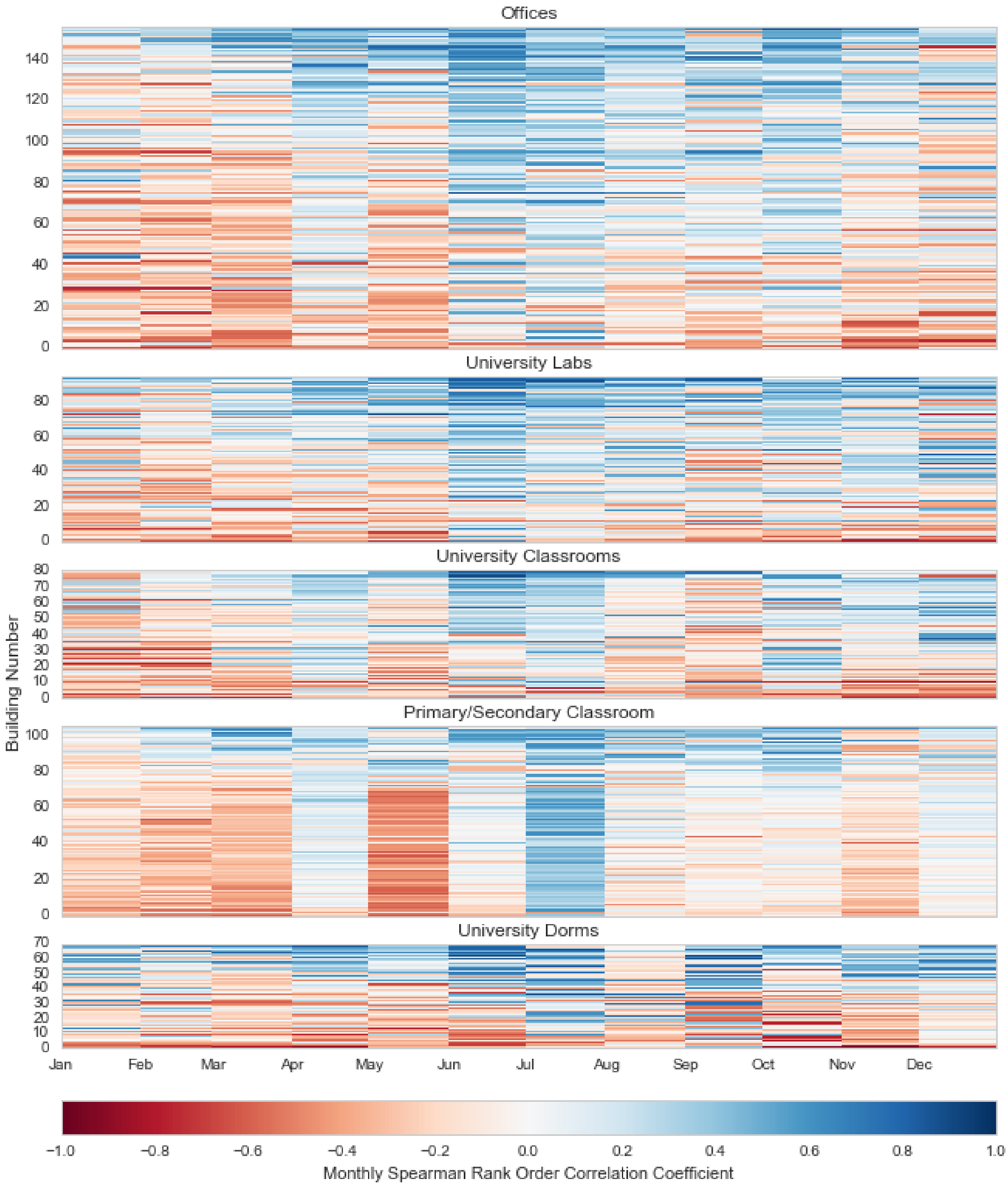

Figure 4.7: Heatmap of spearman rank order correlation coefficient for all case study buildings

5 Regression Model-based Features

Semi-physical behavior about a building can be extracted by using performance prediction models and using output parameters and goodness-of-fit metrics for characterization. This section covers the use of several common electrical consumption prediction models to create sets of temporal features useful for characterization of buildings. Section 5.1 covers the theory underlying each technique and Section 5.2 discusses the implementation of the case study data with a focus on underlying trends as related to building use type.

5.1 Theoretical Basis

5.1.1 Load shape regression-based Features

Prediction of electrical loads based on their shape and trends over time is a mature field developed to forecast consumption to detect anomalies and analyze the impact of demand response and efficiency measures. The most common technique in this category is the use of heating and cooling degree days to normalize monthly consumption (Fels 1986). Over the years, various other techniques have been developed using techniques such as neural networks, ARIMA models, and more complex regression (Taylor *et al.* 2006). However, simplified methods have retained their usefulness over time due to ease of implementation and accuracy. In the context of temporal feature creation, a regression model provides various metrics that describe how well a meter conforms to conventional assumptions. For example, if actual measurements and predicted consumption match well, the underlying behavior of energy-consuming systems in the building has been captured adequately. If not, there is the uncharacterized phenomenon that will need to be obtained with a different type of model or feature.

A contemporary, simplified load prediction technique is selected to create temporal features that capture whether the electrical measurement is simply a function of time-of-week scheduling. This model was developed by Matthieu et al. and Price and implemented mostly in the context of electrical demand response evaluation (Price 2010; Mathieu *et al.*

2011). The premise of the model is based on two features: a time-of-week indicator and an outdoor air temperature dependence. This model is also known as the *Time-of-week and Temperature or (TOWT)* model or *LBNL regression model* and is implemented in the *eetd-loadshape* library developed by Lawrence Berkeley National Laboratory[1].

According to the literature, the model operates as follows (Price 2010). The time of week indicator is created by dividing each week into a set of intervals corresponding to each hour of the week. For example, the first interval is Sunday at 01:00, the second is Sunday at 02:00, and so on. The last, or 168th, interval is Saturday at 23:00. A different regression coefficient, α_i, is calculated for each interval in addition to temperature dependence. The model uses outdoor air temperature dependence to divide the intervals into two categories: one for occupied hours and one for unoccupied. These modes are not necessarily indicators of exactly when people are inhabiting the building, but simply an empirical indication of when occupancy-related systems are detected to be operating. Separate piecewise-continuous temperature dependencies are then calculated for each type of mode. The outdoor air temperature is divided into six equally sized temperature intervals. A temperature parameter, β_j, with $j = 1...6$, is assigned to each interval. Within the model, the outdoor air temperature at time, t, occurring at time-of-week, i, (designated as $T(t_i)$) is divided into six component temperatures, $T_{c,j}(t_i)$. Each of these temperatures is multiplied by β_j and then summed to determine the temperature-dependent load. For occupied periods the building load, L_o, is calculated by Equation 5.1.1.

$$L_0(t_i, T(t_i) = \alpha_i + \sum_{j=1}^{6} \beta_j T_{c,j}(t_i) \tag{5.1.1}$$

Prediction of unoccupied mode occurs using a single temperature parameter, β_u. Unnoccupied load, L_u, is calculated with Equation 5.1.2.

$$L_0(t_i, T(t_i) = \alpha_i + \beta_u T_{c,j}(t_i) \tag{5.1.2}$$

The primary means of temporal feature creation from this process is through the analysis of model fit. The first metric calculated is a normalized, hourly residual, R, that can be used to visualize deviations from the model. It is calculated from the actual load, L_a, and

[1]https://bitbucket.org/berkeleylab/eetd-loadshape

the predicted load, L_p. The residual at a specific hour, t, is calculated using Equation 5.1.3.

$$R_t = \frac{L_{t,a} - L_{t,p}}{max_{L_a}} \tag{5.1.3}$$

An example of the TOWT model implemented on one of the case study buildings is seen in Figure 5.1. Two primary characteristics are captured from a model residual analysis. The first is the building's primary deviation from a set time-of-week schedule and behavior causing the model to highly over-predict. These deviations are most often attributed to public holidays, breaks in normal operation, or changes in normal operating modes. In the single building study, one of the most obvious daily deviations, Christmas Day, is observed. This day is significantly over-predicted due to the model not being informed of the Christmas Day holiday. The automated capture of this phenomenon can inform whether the building is of a certain use-type or in a certain jurisdiction. The second characteristic captured are periods of under prediction when the building is consuming more electricity than expected. These data inform whether a building is being consistently utilized, or whether there is volatility in its normal operating schedule from week-to-week.

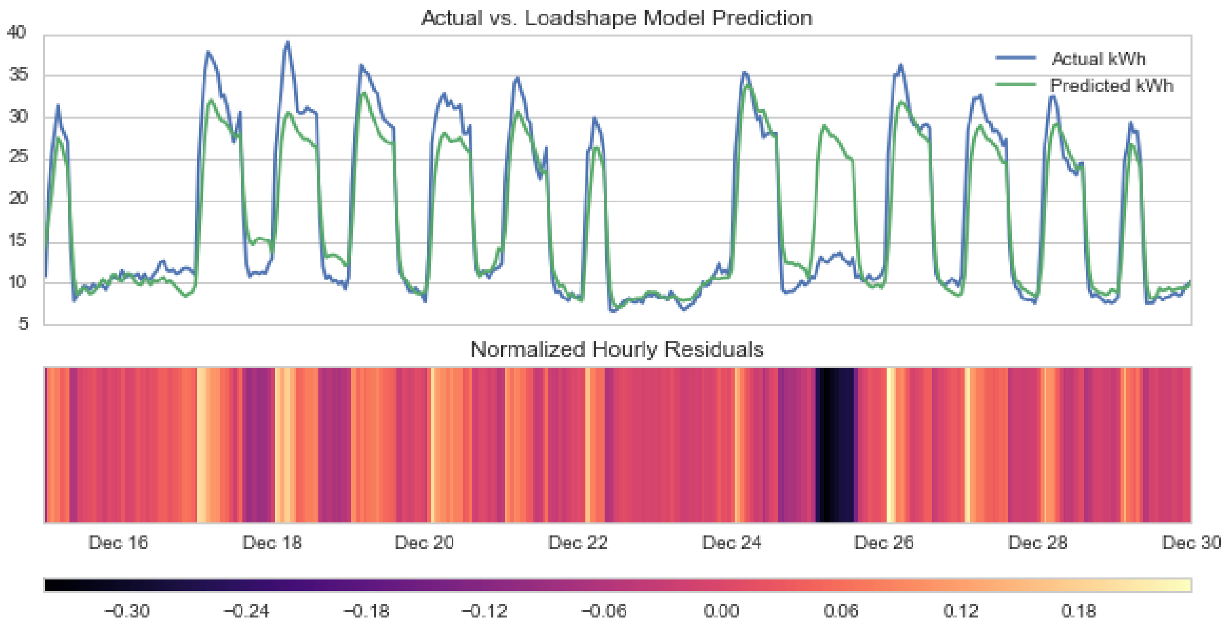

Figure 5.1: Single building example of TWOT model with hourly normalized residuals

5.1.2 Change Point Model Regression

Another means of performance modeling that takes weather characterization into consideration is the use of linear change point models. The outputs of these models can be interpretable in approximating the amount of energy being used for heating, ventilation, and air-conditioning (HvAC). This type of model has its basis in the previously-mentioned PRISM method and has been continuously utilized, recently by Kissock and Eger (Kissock & Eger 2008). This multivariate, piece-wise regression model is developed using daily consumption and outdoor air dry-bulb temperature information. A linear regression model is fitted to data detected to be correlated with outdoor dry-bulb air temperature, either positively for cooling energy consumption or negatively for heating energy consumption. For example, as the outdoor air temperature climbs above a certain point, the relationship between electrical consumption and every degree increase in temperature should be a linear line with a certain slope if the building has an electrically-driven cooling system. The point at which this change occurs is considered the cooling balance point of the building and the slope of the line is the rate of cooling energy increase due to outdoor air conditions. This example can be seen in Figure 5.2a in which the base load of the building is designated as β_1, the slope of the cooling energy line is β_2, and the change point temperature is β_3. Heating energy, as seen in Figure 5.2b, is similar except that the slope of the line will be negative; as temperature decreases, the heating energy increases. An optimization algorithm is used to detect each of these parameters from either hourly or daily raw data.

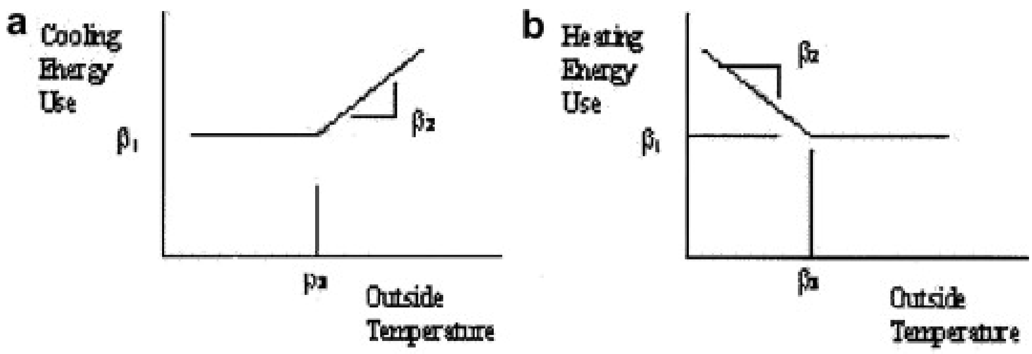

Figure 5.2: Example of an (a) 3 point cooling and (b) 3 point heating change point models (Used with permission from (Kelly Kissock & Eger 2008))

Equations 5.1.4 and 5.1.5 are used to predict energy energy consumption based on an outdoor air temperature, T. This equation can also predict the heating $(\beta_2(T - \beta_3))$

or cooling ($\beta_2(\beta_3 - T)$) components of the electrical consumption to a certain level of accuracy.

$$E_c = \beta_1 + \beta_2(T - \beta_3) \tag{5.1.4}$$

$$E_h = \beta_1 + \beta_2(\beta_3 - T) \tag{5.1.5}$$

Figure 5.3 illustrates a change point model fit on an office building in a continental climate that includes both heating and cooling seasons. It should be noted that the model is not perfectly characterizing the data due to two modes of daily operation; this situation is due to there being an offset between occupied and unoccupied operation. This model is used to generate features of approximate heating and cooling energy and in general, the slopes of these two modes can safely be assumed to be similar in most cases. The Open Meter Python library is used to regress these models for each building in this study [2].

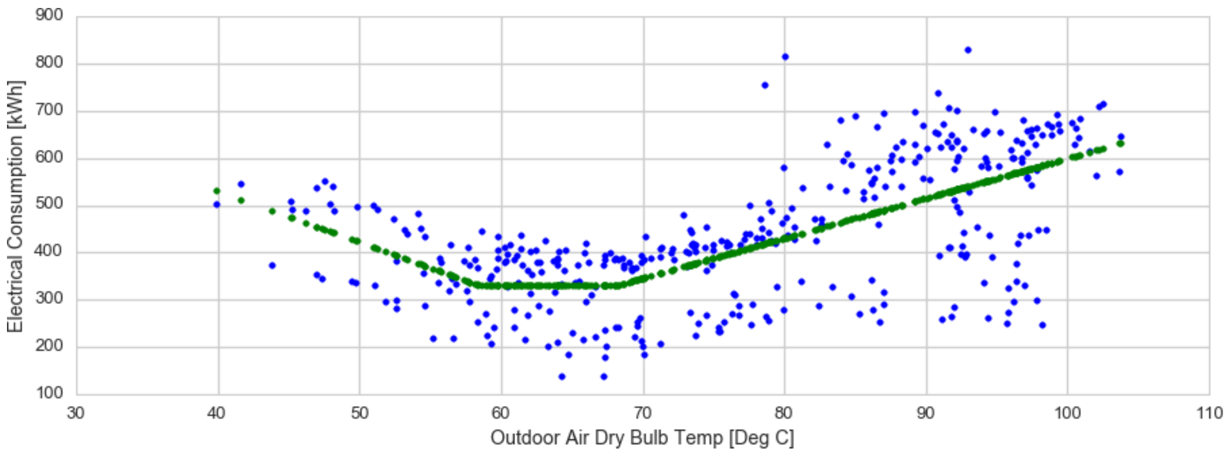

Figure 5.3: Single building example of change point model of a building

Figures 5.4 and 5.5 illustrate single building examples of using the regression model to extract the approximate heating and cooling electrical consumption from the overall power meter. The cooling consumption example illustrates cooling consumption primarily in the summer-time season, as expected. An interesting aspect of this example is that there are a couple days of predicted cooling consumption in November and December. These days are due to outdoor air temperature crossing the balance point in anomalous ways during that

[2] http://www.openeemeter.org/

season. The heating consumption example also resembles an intuitive understanding how the heating season from December to mid-April. In each example, one notices a correlation between the cooling and heating consumption in the heat-map and slight increases in the line charts indicating seasonality.

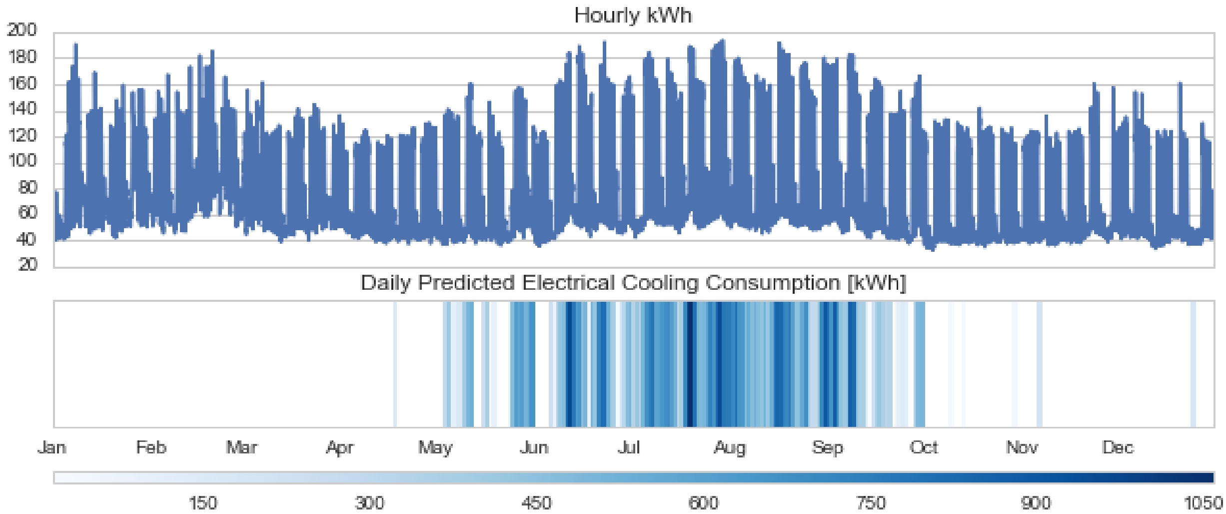

Figure 5.4: Single building example of predicted electrical cooling energy using change point model

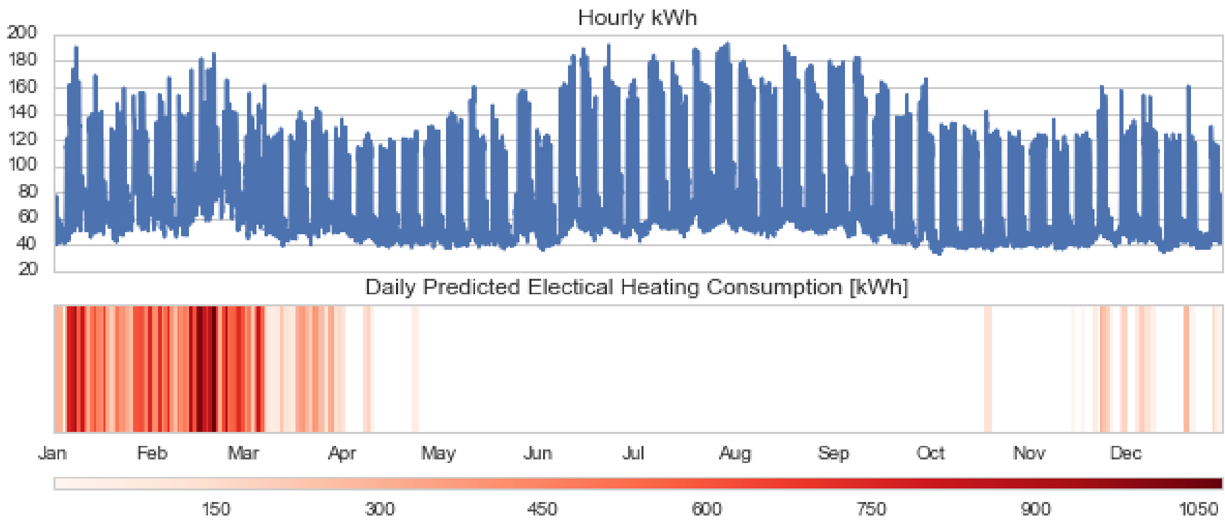

Figure 5.5: Single building example of predicted electrical heating energy using change point model

5.1.3 Seasonality and Trend Decomposition

Temporal, or time series data, from different sources, often exhibit similar types of behavior that are studied within the field of forecasting and temporal data mining. Electrical building meter data fits within this category, and the same feature extraction techniques can be applied as what is commonly done for financial or social science analysis. These techniques often seek to decompose time-series data into several components that represent the underlying nature of the data (Mitsa 2010). For example, the electrical meter data collected from buildings is often cyclical in its weekly schedule. People are utilizing buildings each day of the week in a relatively predictable pattern. A very common example of this behavior is found in office buildings where occupants are typical white collar professionals who come into work on weekdays at a particular time and leave to go home at a certain time. Weekends are unoccupied periods in which there is little to no activity. This behavior is an example of what's known as seasonality within time series analysis. Seasonality is a fixed and known period of consistent modulation and is a feature that is often extracted before creating predictive models.

Trends are another feature commonly found in temporal data. A trend is a long-term increase or decrease in the data that often doesn't follow a particular pattern. Trends are commonly due to factors that are less systematic than seasonality and are often due to external influences. For building energy consumption, trends manifest themselves as gradual shifts in consumption over the course of week or months. Often these shifts are due to weather-related factors having an influence on the HVAC equipment. Other causes of trends are changes in occupancy of degradation of system efficiency.

To capture these features to understand their impact on characterizing buildings, the seasonal-trend decomposition procedure based on loess is used to extract each of these features from the case study buildings (Cleveland *et al.* 1990). This process is used to remove the weekly *seasonal* patterns from each building, the long-term trend over time, and the residual remainders from the model developed by those two components. The input data is aggregated to daily summations and weather normalized by subtracting the calculated heating and cooling elements from the change point model described in Section 5.1.2. This step is done to reduce the influence weather plays in the trend decomposition. The *STL* package in R is used for this process to extract the seasonal, trend, and irregular components [3].

The details of the inner algorithms of the *STL* procedure are described by Cleveland et al. (Cleveland *et al.* 1990). The process uses an inner loop of algorithms to detrend and

[3]https://stat.ethz.ch/R-manual/R-devel/library/stats/html/stl.html

deseasonalize the data by creating a trend component, T_v, and a seasonal component, S_v. The remainder component, R_v, is a subtraction of the input values, Y_v as seen in Equation 5.1.6.

$$R_v = Y_v - T_v - S_v \tag{5.1.6}$$

An output of the process of the *STL* package is seen in Figure 5.6. The *data* component is the weather-normalized electrical meter data, the *seasonal* component is decomposed weekly pattern, the *trend* is the smoothed trend component, and the *remainder* is the residual after the other components have been subtracted out.

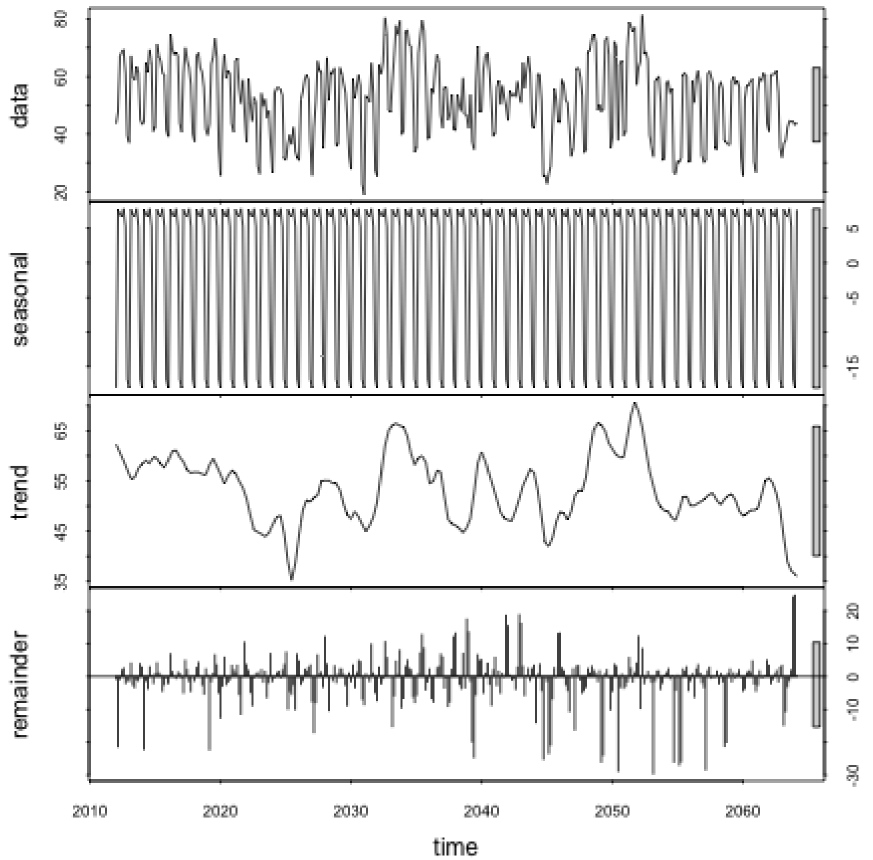

Figure 5.6: Output of seasonal decomposition process using loess for a single building.

The seasonal component of this decomposition process can then be extracted to get an understanding of the typical weekly pattern of a building's electrical consumption. Figure 5.7 illustrates this situation for a single building that has a typical office-style utilization

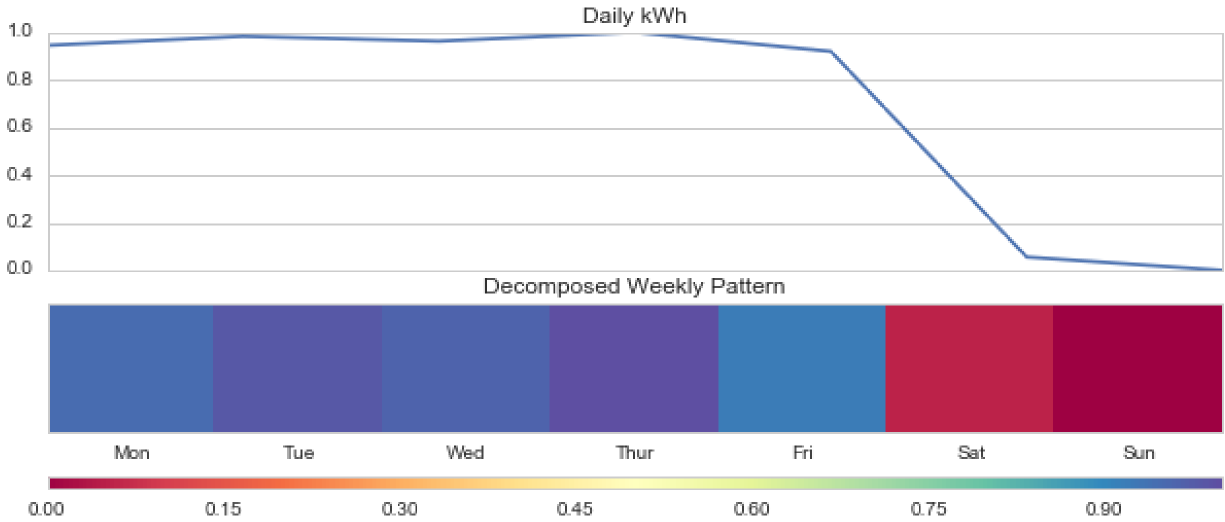

Figure 5.7: Single building example of decomposed weekly patterns using the *STL* process

schedule with a Monday to Friday working schedule with Saturday and Sunday off. This metric has been normalized to make it comparable to other buildings.

The general trend over the course of the year of data is another example of quantifying the seasonal patterns in utilization of a building. Weather influence has been reduced or removed using the change point models. Therefore, a trend could be the result of changes in building occupancy due to breaks, changes in equipment or space functions that would significantly increase or decrease the consumption, or gradual faults in systems of equipment. Figure 5.8 illustrates a single building example of a decomposed trend for a building. January to May is in the middle range of consumption trend with a noticeable dip in April. From June to Oct, there is a trend upwards of higher than normal consumption, perhaps due to higher utilization of the space. October to the end of the year is back to average with a slight dip during the last few weeks of the year.

The remainder values of the *STL* decomposition process are indicators of days that fall outside of the *STL* model's prediction. This situation is similar to the residuals of the *loadshape* models in Section 5.1.1. Figure 5.9 illustrates an example of the residual days. Once again, this metric is normalized, however not on a 0 to 1 range. Instead, negative values indicate a lower than expected consumption for the day, while positive values are higher than average. In this example, the residuals aren't exceptionally systematic. However, a few identifiable days can be seen including Thanksgiving in November.

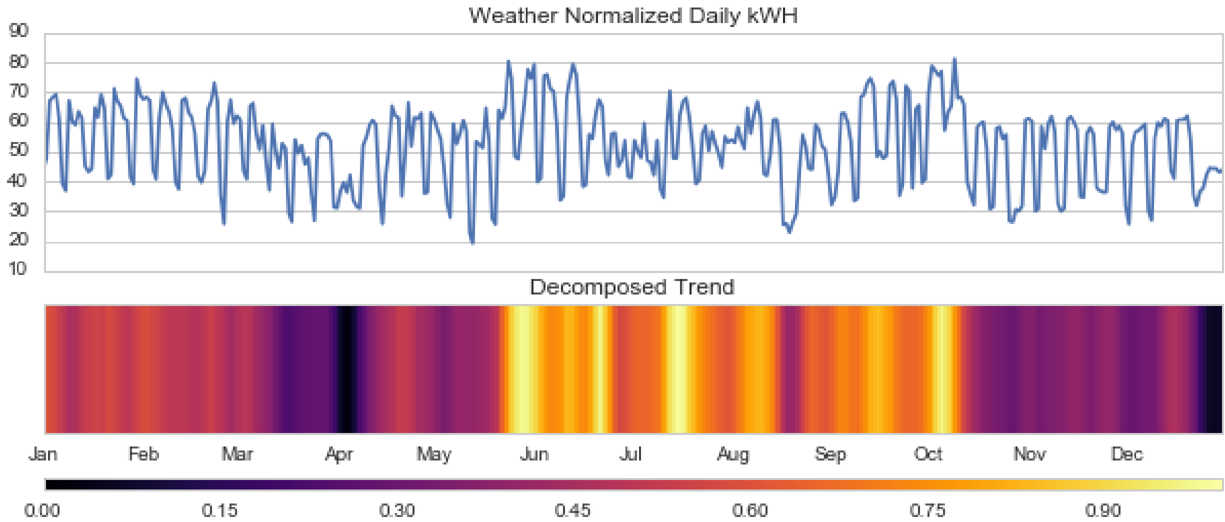

Figure 5.8: Single building example of decomposed trend using the *STL* process

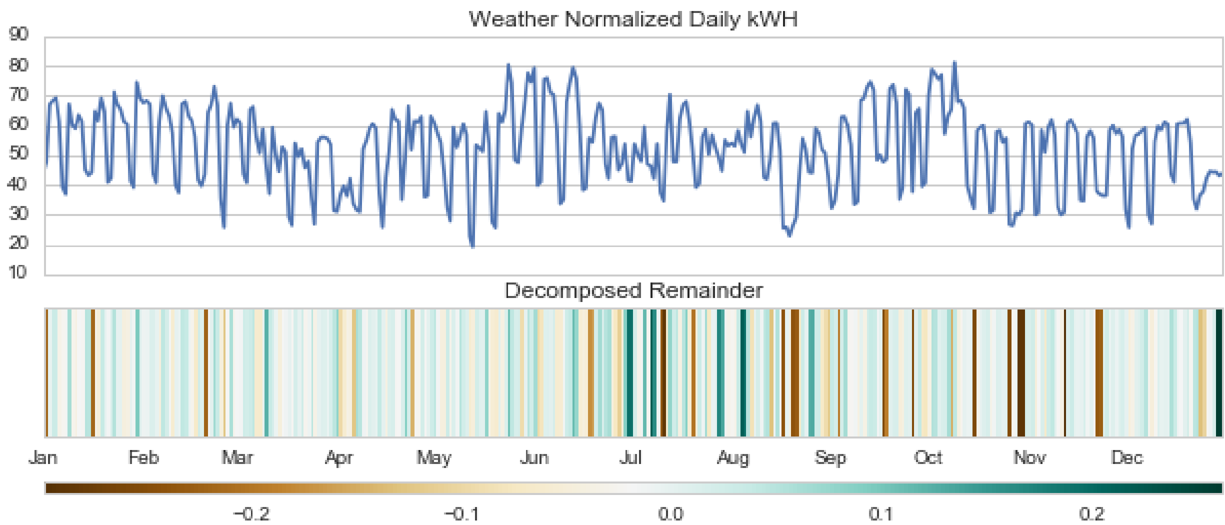

Figure 5.9: Single building example of decomposed remainder component using the *STL* process

5.2 Implementation and Discussion

Based on the theoretical basis of model-based approaches, the techniques are then applied to the 507 targeted case study buildings. This process enables the analysis of various patterns and phenomenon occurring in the data as a result of the building use type. Figure 5.10 illustrates an overview of an implementation of the loadshape model on all the buildings across the various building use types in the study. The differences between each use type can be noticed from a high level due to the nature of residuals. The darker areas of the visualization indicate when the model is highly over-predicting consumption and lighter areas indicate when the model is under-predicting. Common holiday periods such as spring, summer and winter breaks and holidays such as the American Labor Day and Thanksgiving are seen as darker areas. Offices, labs and classrooms seem to have similar residual patterns, likely due to their scheduling being similar. Slight key differences are seen such as the fact that classrooms have more general areas of over-prediction, likely due to less consistent occupancy. Primary/Secondary schools and dormitories are clearly less predictable on an annual basis due to their strong seasonal patterns of use; this fact is intuitive and model residuals of this type are accurate in automatically characterizing this behavior.

Figures 5.11 and 5.12 illustrate heating and cooling energy regression for all case study buildings. These figures have been normalized according to floor area. Each building's response to outdoor air temperature is indicative of the type of systems installed in addition to the efficiency of energy conversion of those systems. Approximately 15-20% of offices, labs, and classrooms have a certain amount of cooling electrical consumption, while the rest have little to none. Many of those buildings are on district heating and cooling systems, therefore, weather dependent electrical consumption is likely due to air distribution systems or auxiliary pumps. Several of the labs have year-round cooling consumption, likely due to climate and the high internal loads that accompany laboratory environments.

Figure 5.13 illustrates the weekly pattern decomposition for all of the case study buildings. For offices, most of the other cases also exhibit a typical Monday to Friday schedule, with a few exceptions that have various weekday differences and several that have higher values on Saturday. Tuesday seems to be the most consistent across the range of buildings on the peak day of consumption. University labs and classrooms appear to have the same amount of diversity and a similar schedule to offices, perhaps with slightly less use of Fridays. Primary/Secondary school classrooms appear to be the most consistent in their weekly Monday to Friday schedule and have an entirely consistent lack of Saturday and Sunday utilization. University dormitories are the most diverse in their weekly patterns with approximately half of the buildings having dominant weekday schedules and half having dominant weekend schedules.

Figure 5.14 illustrates the trend decomposition as applied to the entire case study set of buildings. Offices appear to have quite a bit of diversity over time, with a few observable systematic low spots in the spring and autumn periods at the bottom of the heat map. Laboratories reflect that behavior, while university visually has an opposite effect with

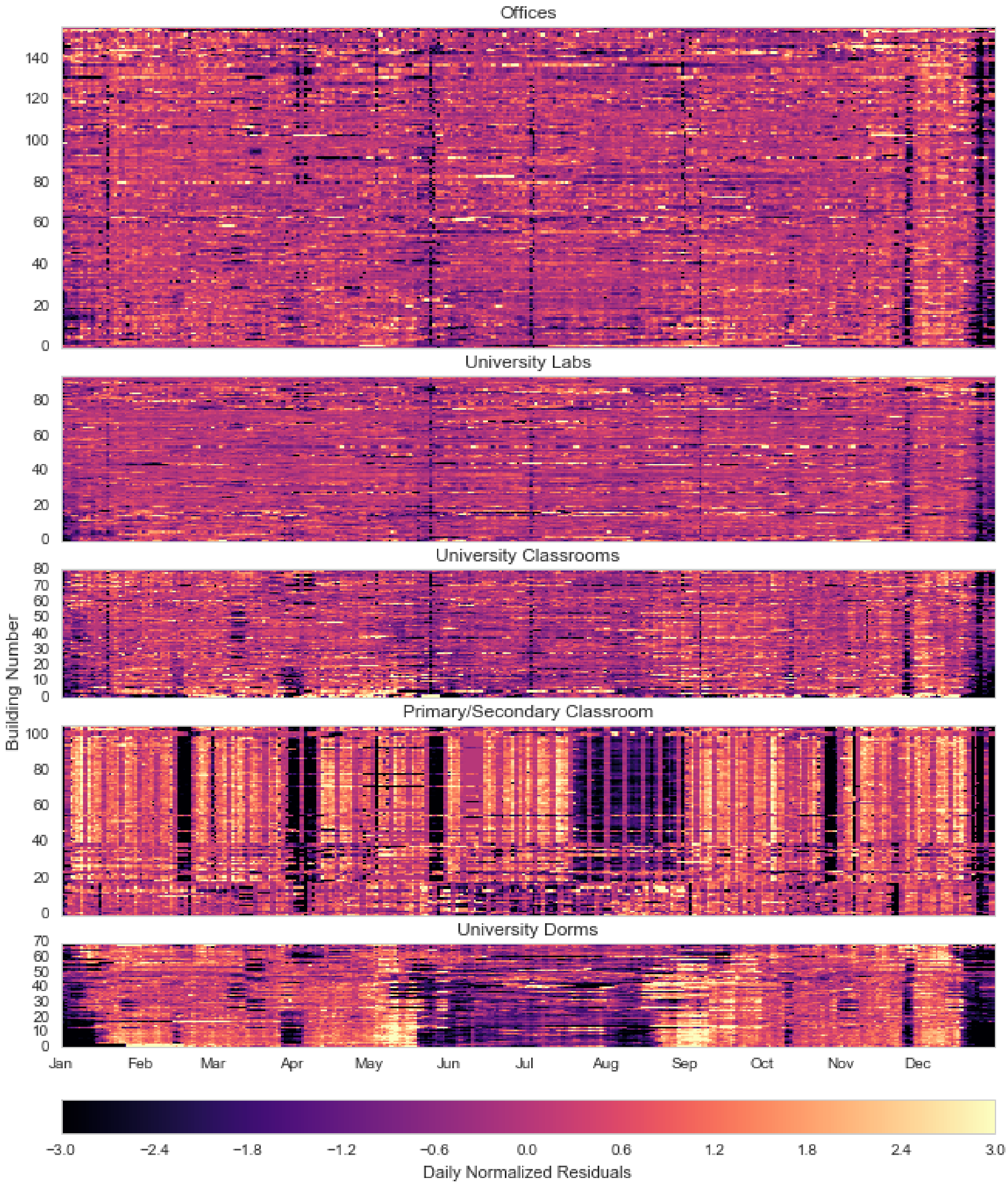

Figure 5.10: Heatmap of normalized daily residuals for all case study building

lower than the average trend in the summer months. Primary/Secondary school classrooms have a very distinct delineation between when school is in session and out of session

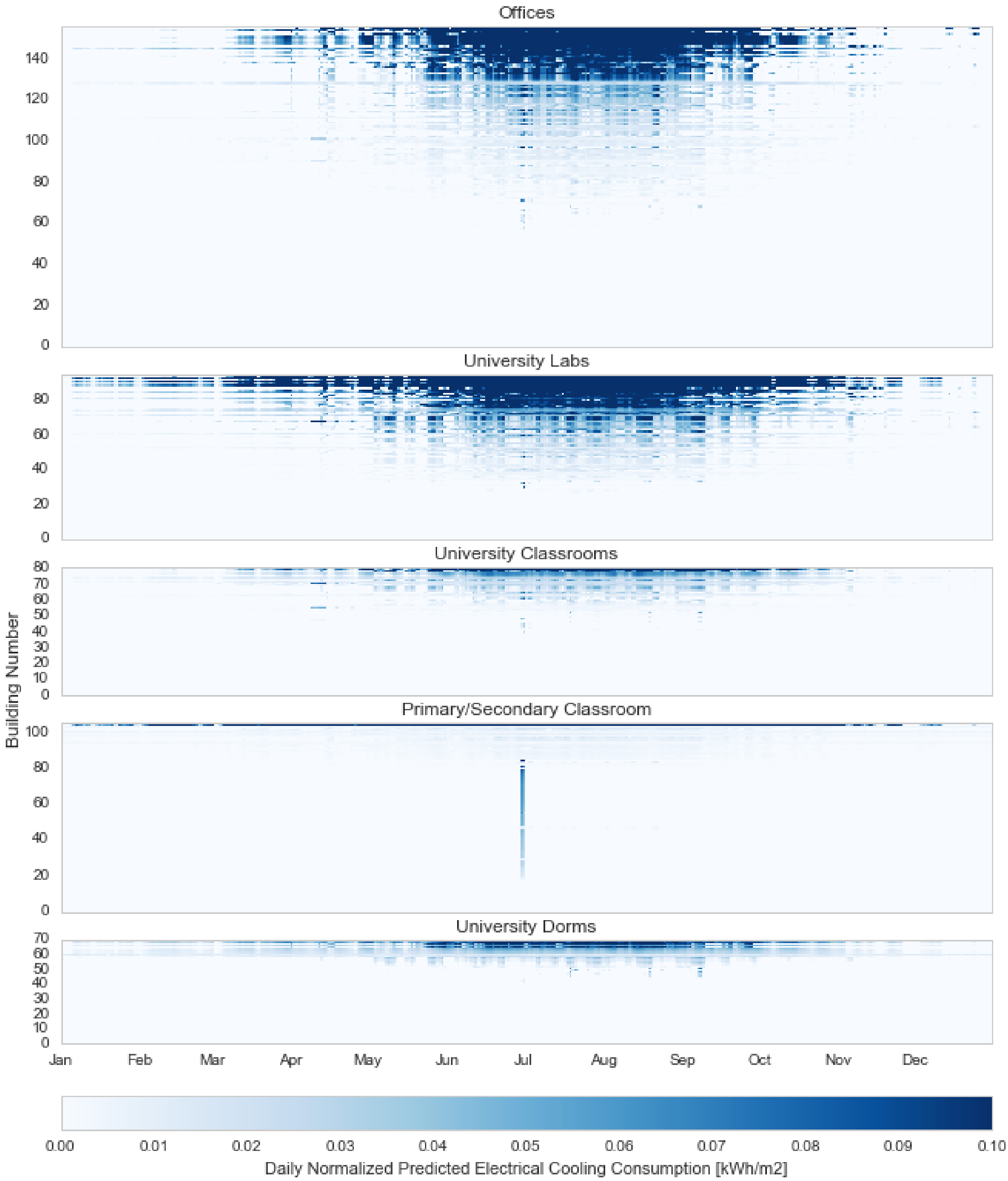

Figure 5.11: Heatmap of normalized predicted electrical cooling energy for all case study buildings

1158 during the summer and various breaks. As many of these schools are in the UK, their
1159 out-of-session periods appear to line up naturally. University dormitories also have clear

Figure 5.12: Heatmap of normalized predicted electrical heating energy for all case study buildings

delineations between occupied and unoccupied periods and they seem also to match up quite well, despite the diversity of data sources of these buildings.

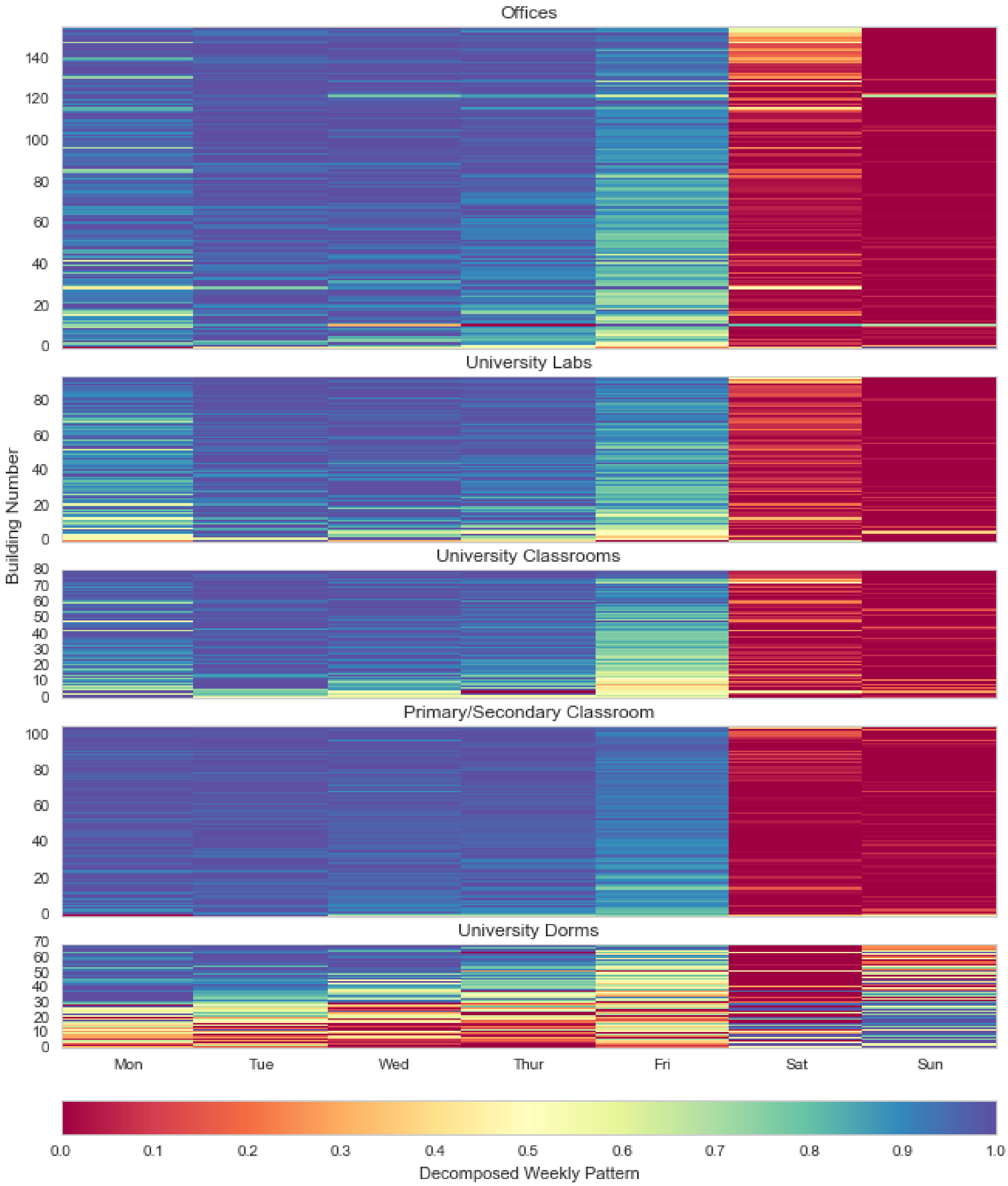

Figure 5.13: Heatmap of decomposed weekly patterns for all case study buildings

Figure 5.15 illustrates the residuals applied across all of the case study buildings. Some similarity between all of the university offices, labs, and classrooms are apparent regarding

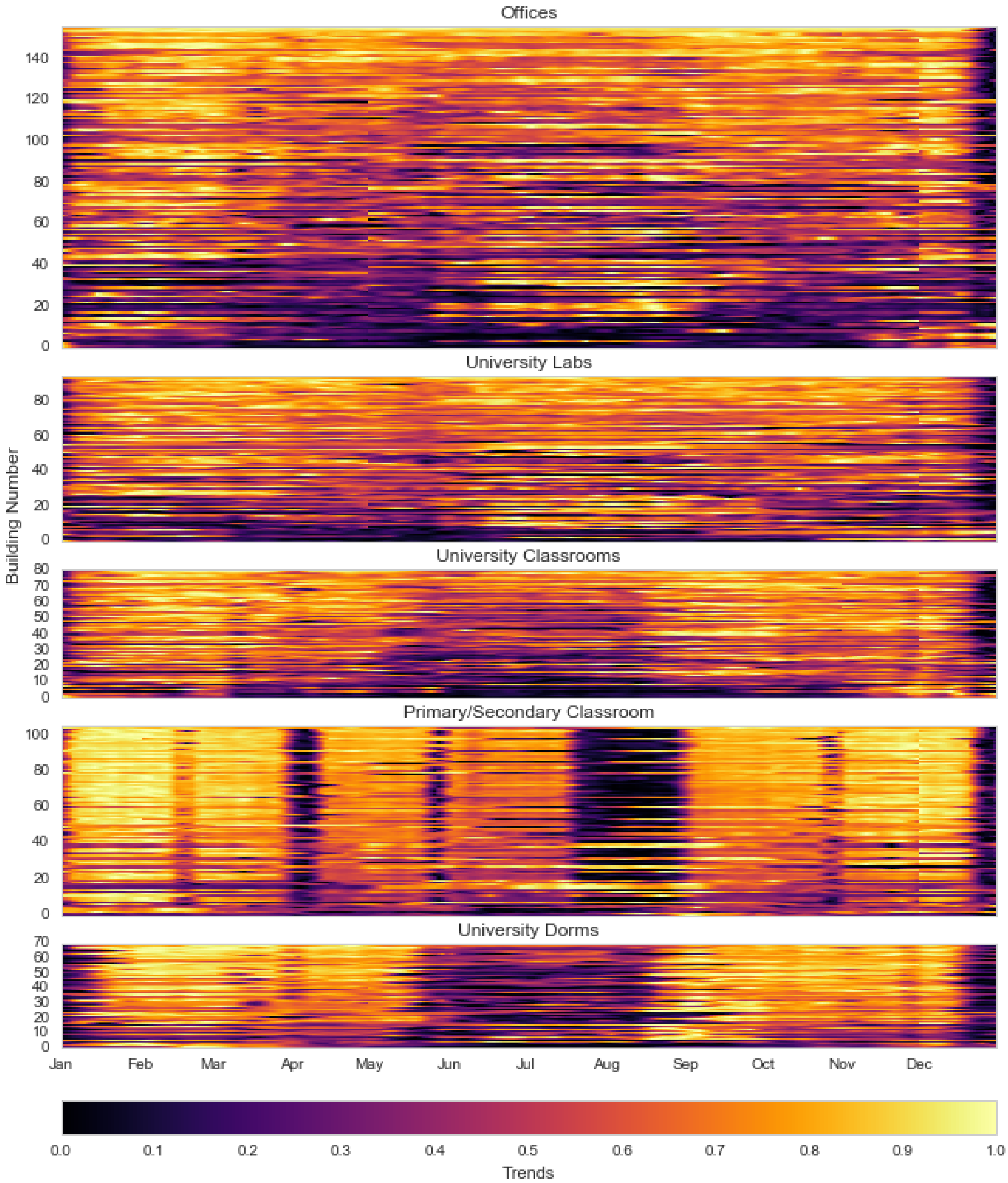

Figure 5.14: Heatmap of decomposed trend over time for all case study buildings

the holidays detected. The most consistent ones include the American memorial day in May, American Independence Day in July, Thanksgiving in November and Christmas Day in December. However, University Labs have a slightly less dramatic range of values.

Primary/Secondary schools have appeared to have many more dramatic differences from the *STL* model.

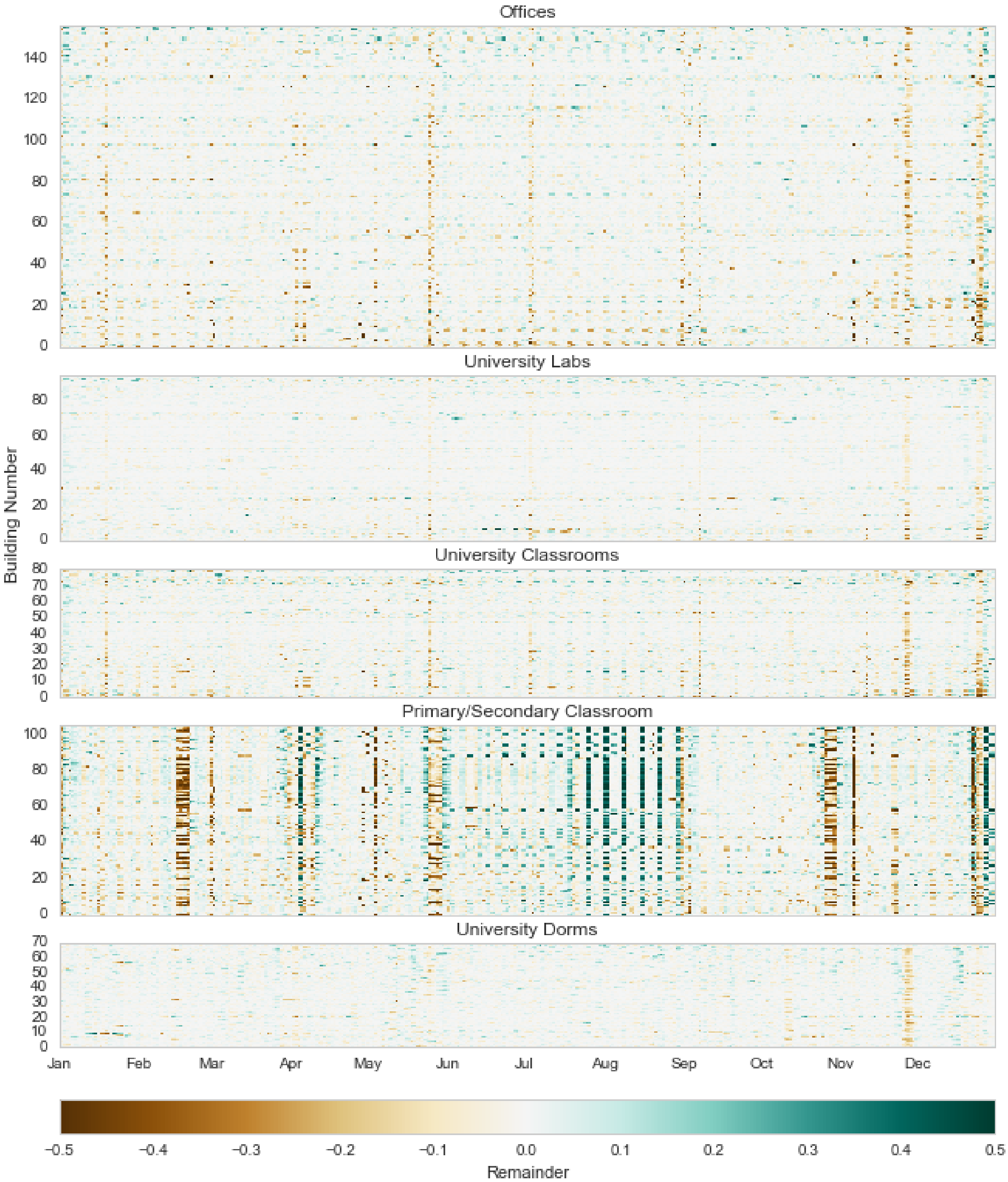

Figure 5.15: Heatmap of decomposed remainder residuals for all case study buildings

Overall, model-based temporal features are good at highlighting several different phenomena occurring in a building's behavior. The first, and most important, is essentially how *predictable* a building is across an annual time range and what systematically anomalous days are occurring, such as holidays and break periods. Weather-related models are helpful in understanding what consumption is likely due to heating and cooling systems. This feature is different than the spearman coefficient from the statistic-based section in that it provides more information related to *when* a building goes into climate control modes regarding outside air temperature.

6 Pattern-based Features

Temporal data mining for performance monitoring focuses on the extraction of patterns and model building of time series data. These techniques are, in some ways, similar to many existing building performance analysis approaches; however, different concepts and terminology are used. Two key concepts to understand when applying data mining to buildings are that of *motifs* and *discords*. A motif is a common subsequence pattern that has the highest number of non-trivial matches (Patel *et al.* 2002), thus, a pattern that is frequently found in the dataset. A discord, on the other hand, is defined as a subsequence of a time series that has the largest distance to its nearest non-self match (Keogh *et al.* 2005). It is a subsequence of a univariate data stream that is least like all other non-overlapping subsequences and is, therefore, an unusual pattern that diverges from the rest of the dataset. These definitions are more general than that of a *fault* and therefore more appropriate for the goal of higher level information extraction with less parameter setting. In short, the goal is to find *interesting or infrequent* behavior efficiently and not create a detailed list of specific problems that could be occurring in individual systems.

6.1 Theoretical Basis

To work with standard temporal mining approaches, Symbolic Aggregate approXimation (SAX) representation of time-series data (Lin *et al.* 2003) is used. SAX allows discretization of time series data which facilitates the use of various motif and discord detection algorithms. The process breaks time series data into subsequences which are converted into an alphabetic symbol. These symbols are combined to form strings to represent the original time series enabling various mining and visualization techniques. Regarding application, an example of a process using SAX-based techniques is the VizTree tool that uses augmented suffix tree visualizations designed for usability by an analyst (Lin *et al.* 2004). A particular application of VizTree is the analysis of collected sensor data from an impending spacecraft launch in which thousands of telemetry sensors are feeding data back to a command center where experts are required to interpret the data. Visualization and filtering tools are needed that allow a natural and intuitive transfer of mined knowledge to the monitoring task. Human perception of visualizations and the algorithms behind them must work in unison to achieve an understanding of significant amounts of original data streams.

6.1.1 Dirunal Pattern Extraction

Towards the development of diurnal motif and discord extraction, a new technique was developed as an application of temporal data mining to building performance data. It is a process called *DayFilter* and it includes five steps designed to filter structure incrementally from daily raw measured performance data. These steps, as seen in Figure 6.1, are intended to bridge the gap between contemporary top-down and bottom-up techniques. The arrows in the diagram denote the execution sequence of the steps. Note that steps 3, 4, and 5 produce results applicable to the implementation of bottom-up techniques. Much of the graphics and explanation for this section are contained in a publication explaining DayFilter and its uses (Miller *et al.* 2015).

The whole building and subsystem metrics are targeted for analysis to determine high-level insight. The process begins with a data preprocessing step which removes obvious point-based outliers and accommodates for gaps in a univariate data set of variable length. Next, the raw data is transformed into the SAX time-series representation for dimensionality reduction by creating groups of SAX words from daily windows. This step enables the quick detection of *discords*, or regular patterns of performance that fall outside what is considered normal in the dataset according to the frequency of patterns. The discords are filtered out for future investigation while the remaining set of SAX words is clustered to create performance *motifs* or the most common daily profiles. The additional clustering step beyond the SAX transformation and filtering adds the ability further to aggregate daily profiles beyond the SAX motif candidates. These clusters are useful in characterizing what can be considered *standard* performance. Finally, these data are presented using visualization techniques as an aid to interpreting the questionable discords and the common clusters. In the following simplified example, each of these steps is detailed. The input parameter selections in this section are based on suggestions from other studies using SAX aggregation and clustering approaches.

As in any data mining approach, data preprocessing is an important step to clean and standardize the data. In the proposed method, extreme point measurements are removed that fall outside of three standard deviations, 3σ, of the mean, μ, of the selected univari-

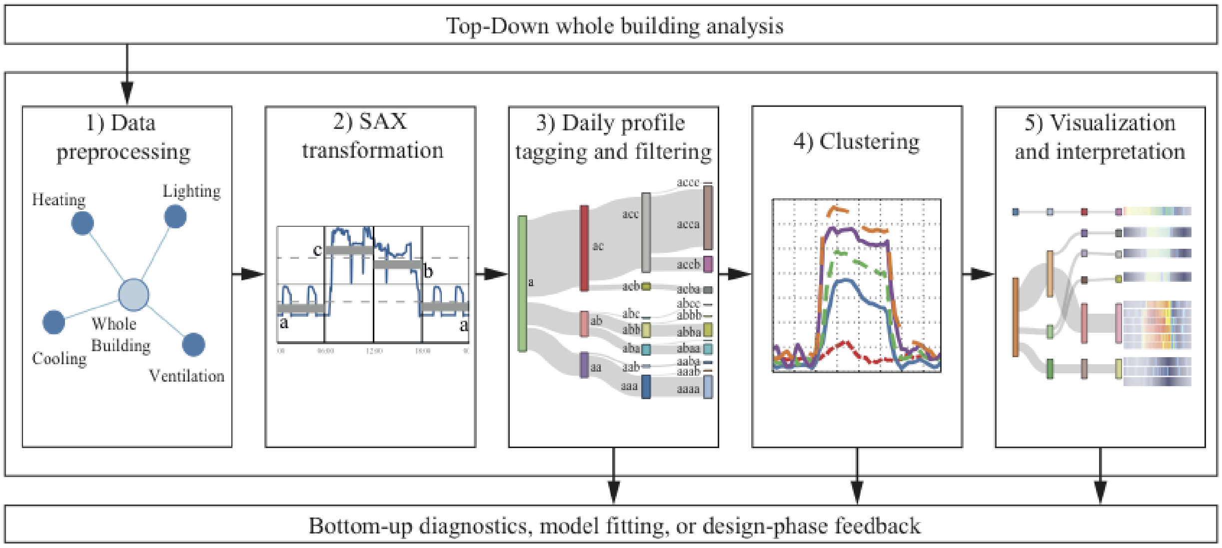

Figure 6.1: Diagram of the five steps in the *DayFilter* (from (Miller *et al.* 2015))

ate data stream $x(t)$. The data are then normalized to create a dataset, $Z(t)$ with an approximate 0 mean and a standard deviation of close to 1 (Goldin & Kanellakis 1995):

$$Z(t) = \frac{x(t) - \mu}{\sigma} \tag{6.1.1}$$

In the second step, $Z(t)$ is transformed into a symbolic representation using SAX. It is one of the many means of representing time-series data to enhance the speed and usability of various analysis techniques. SAX is a type of Piecewise Aggregate Approximation (PAA) representation developed by Keogh et. al and it has been used extensively in numerous applications (Lin *et al.* 2007).

In brief, the SAX transformation is as follows. The normalized time-series, $Z(t)$, is first broken down into N individual non-overlapping subsequences. This step is known as *chunking*, and the period length N is based on a context-logical specific period (Lin *et al.* 2005). In this situation, N is chosen as 24 hours due to the focus on daily performance characterization. Each chunk is then further divided into W equal sized segments. The mean of the data across each of these segments is calculated and an alphabetic character is assigned according to where the mean lies within a set of vertical breakpoints, $B = \beta_1, ..., \beta_{a-1}$. These breakpoints are calculated according to a chosen alphabet size, A, to create equiprobable regions based on a Gaussian distribution, as seen in Table 6.2.

β_i \ A	3	4	5
β_1	-0.43	-0.67	-0.84
β_2	0.43	0	-0.25
β_3		0.67	0.25
β_4			0.84

Figure 6.2: Example breakpoint lookup table from Keogh et. al (Keogh *et al.* 2005) for $A = 3, 4, 5$ calculated from a Gaussian distribution (Miller *et al.* 2015)

Based on a chosen value of W segments and alphabet size A, each N size window is transformed into a SAX *word*. An example of this process is seen in Figure 6.3. This example shows two daily profiles which are converted to the SAX words, *acba* and *abba*. The SAX word is useful from an interpretation point of view in that each letter corresponds consistently to a subsequence of data from the daily profile. For example, the first letter explains the relative performance for the hours of midnight to 6:00 AM. Therefore if the size of A is set to 3, a SAX word whose first letter is a would have low, b would indicate average, and c would correspond to high consumption. Larger sizes of A would create SAX words with a more diverse range of characters and would capture more resolution magnitude-wise.

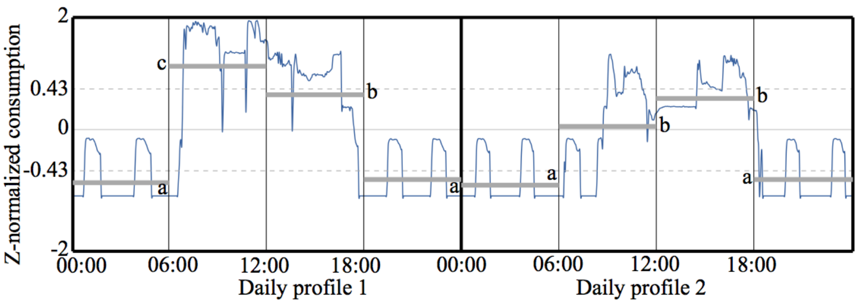

Figure 6.3: SAX word creation example (based on figure from Keogh et. al (Keogh *et al.* 2005)) of two days of 3 minute frequency data, parameters are N=480, W=4, and $A = 3$ and the generated representative word for daily profile 1 is *acba* and daily profile 2 is *abba* (from (Miller *et al.* 2015))

The individual subsequences, N, are not normalized independently. This particular decision is divergent from the generalized shape-based discord approaches and is because, at this level of analysis and the context of building performance data, there is interest in discovering interesting subsections based on both magnitude and shape.

The targeted benefits of using SAX in this scenario are that discretization uniformly reduces the dimensionality and creates sets of words from the daily data windows. This transformation allows the use of hashing, filtering, and clustering techniques that are commonly used to manipulate strings (Lin *et al.* 2007).

Once the SAX words are created, each pattern is visualized and tagged as either a motif or discord. The results of applying the SAX process to a two-week sample power dataset are shown in Figure 6.4. The diagram shows how each daily chunk of high-frequency data is transformed into a set of SAX characters. In this example, an alphabet size, A, of 3 and a subsequence period count, W, of 4 are used for each character aggregating the data from 6 hours of each profile. These parameters are the same as used in the more simplified two-day example from Figure 6.3

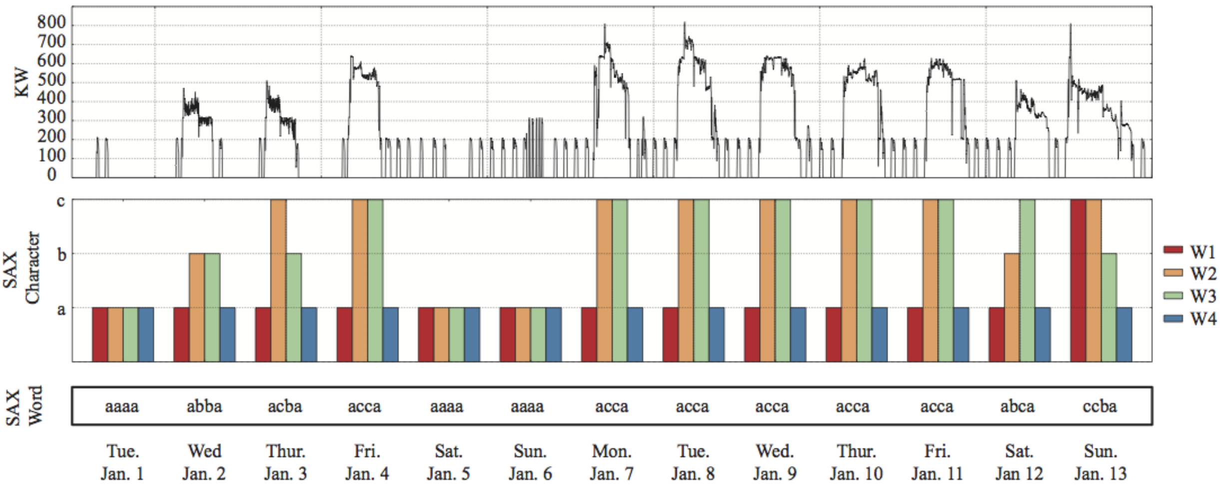

Figure 6.4: Creation of SAX words from daily non-overlapping windows: W1: 00:00-06:00, W2: 06:00-12:00, W3: 12:00-18:00, W4: 18:00-24:00. Time series data is transformed according to a SAX character creation and then as a string, or SAX word (Miller *et al.* 2015)

Figure 6.5 visualizes the frequency of the SAX strings and substrings in the form of an augmented suffix tree. Suffix trees have been an integral part of string manipulation and mining for decades (Weiner 1973). Augmented suffix trees enable a means of visualizing the substring patterns to show frequency at each level. This figure incorporates the use of a Sankey diagram to visualize the tree with each substring bar height representing the number of substring patterns existing through each window of the day-types. The more frequent patterns are categorized as *motifs* or patterns which best describe the average behavior of the system. One can see the patterns with the lower frequencies and their indication as *discords* or subsequences that are least common in the stream.

Heuristically, a decision threshold is set to distinguish between motifs and discords. This threshold can be based on the word frequency count for each pattern as a percentage of the number of all observations. This threshold can be tuned to result in a manageable number of discord candidates to be further analyzed. More details about setting this limit will be discussed the applied case studies.

In the two-week example, this process yields two patterns which have a frequency greater than one and thus are the motif candidates. A manual review of the data confirms that those patterns match with an expected profile for a typical weekday (*acca*) and weekend (*aaaa*). The less frequent patterns are tagged as discords and can be analyzed in more detail. In this case, it can be determined that the patterns *abba*, *abca*, and *acba*, despite being infrequent, are not abnormal due to the occupancy schedule for those particular days. Pattern *ccba*, however, is not explainable within the scheduling and is due to a fault causing excessive consumption in the early morning hours.

This step leads into the next phase of the process focused on further aggregating the motif candidates of the dataset. The size and number of potential motif filtered in this step will give an indication of the number of clusters that will likely pick up the exact structure from the dataset.

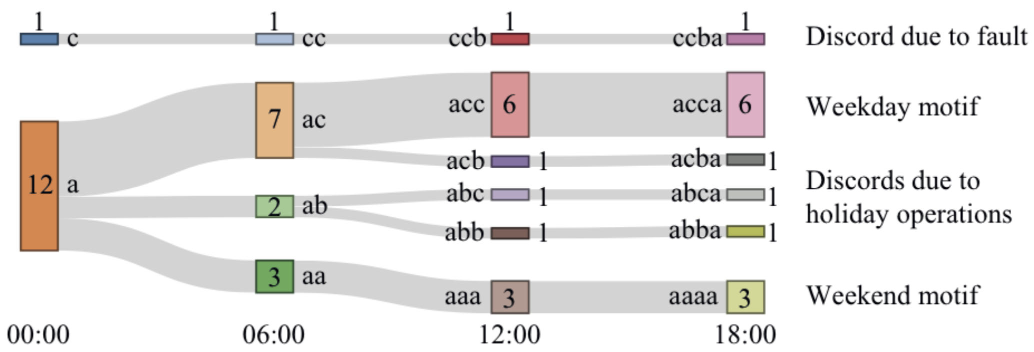

Figure 6.5: Augmented suffix tree of SAX words. Each level from left to right represents the $W1 - W4$, the substrings are noted adjacent to each bar, and the bar thickness is proportional to the number of days within each pattern type. The pattern frequency in number of days is noted in this graphic within or just adjacent to each bar. (from (Miller *et al.* 2015))

As the final step, interpretation and visualization are critical for *DayFilter* for a human analyst to visually extract knowledge from the results, and to make decisions regarding further analysis. The *Overview, zoom and filter, details-on-demand* approach (Shneiderman 1996) and the previously mentioned VizTree tool (Lin *et al.* 2004) are used for insight into this process. The hidden structures of building performance data are revealed through the SAX process, and visualization is used to communicate this structure to an analyst. The method uses a modified Sankey diagram to visualize the augmented suffix tree in a way which the count frequency of each SAX word can be distinguished. Figure 6.6 shows how this visualization is combined with a heat map of the daily profiles associated with each of the SAX words using the same two-week example data from Figures 6.4 and 6.5. The Sankey diagram is rearranged according to the frequency threshold set to distinguish between the motif and discord candidates.

In Figure 6.6, the discords are shown as the top four days, Jan. 2, 3, 12, and 13 and the remaining days are shown as more frequent potential motifs below. Each daily profile is shown adjacent to the right of the Sankey diagram and is expressed as a color-based heatmap. Each horizontal bar of the heat map is an individual day, and they are grouped according to pattern with the associated legend informing the viewer the magnitude of energy consumption across the day. This visualization is designed to present quickly the patterns arranged according to a sort of hierarchy provided by the suffix tree. One can more easily distinguish seemingly *normal* versus *abnormal* behavior with this combination of visualizations.

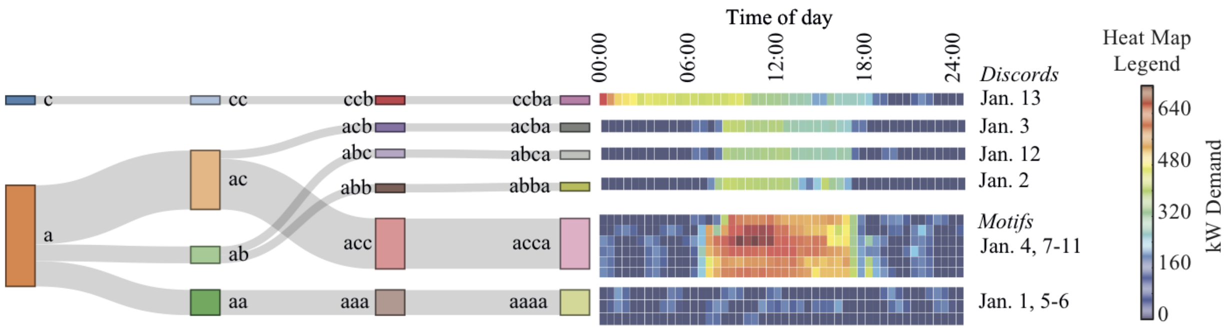

Figure 6.6: Example suffix tree with heatmap from the two week dataset. The sankey diagram illustrates the divisions according to pattern and the general categories of motif vs. discord candidates. Each horizontal line in the heatmap represents a single daily profile to illustrate consumption magnitude of each SAX word. (Miller *et al.* 2015)

DayFilter is applied on a large energy performance datasets to demonstrate the usability and results in real-life scenarios. The process is applied to a 70,000 square meter international school campus in the humid, tropical climate of Singapore. It was built in 2010 and includes a building management system (BMS) with over 4,000 measured data points taken at 5-minute intervals from the years of 2011-2013 - resulting in close to 800 million records of raw data. This collection includes 120 power meters and 100 water meters in the energy and water management system. The data from this study are a seed dataset in an open repository of detailed commercial building datasets (Miller *et al.* 2014).

The chilled water plant electricity consumption is targeted in this case due to its importance in this climate and the potential savings opportunities available through chilled water plant optimization. Measured kilowatt-hour (kWh) and kilowatt (kW) readings were taken from July 12, 2012, to October 29, 2013, with 474 total daily profiles analyzed. Figure 6.7 illustrates a Sankey diagram with a heat map of the output of the *DayFilter* process with parameters set to A=3 and W=4. The discord and motif candidates are separated in this case according to a decision threshold which quantifies a discord as a day-type with a frequency count less than 2% of total days available. This distinction results in 39 days with patterns tagged as discord candidates, which is 8.2% of the total days in the dataset.

In general, there are six primary motif candidates with two candidates appearing to be typical weekday types, two holiday or half-capacity types, and two-weekend unoccupied types. Pattern *aaaa* and *abaa* are predominantly flat profiles common to non-occupied cooling consumption. Patterns *abba* and *acba* are representative of days in which school is out of session, but staff still occupies the office spaces. Pattern *acca* represents a regular full-occupied school day, and it is by far the most common with 202 days tagged out of 474. Pattern *accb* is similar to *acca* with slightly more use in the late afternoon and early evening. This phenomenon is due to extracurricular activities planned outside the normal operating schedule of the facility.

For characterization, a metric is developed from the *DayFilter* process that approximates the presence of motifs and discords. This metric is a daily frequency calculation of each day's pattern count versus the total number of days. An example of this metric is seen in Figure 6.8.

6.1.2 Pattern Specificity

Another way to leverage SAX to characterize the case study data is to use it to extract which patterns are most indicative of a particular building use type. This information is obtained using the SAX-VSM process pioneered by Senin and Malinchik that uses SAX and Vector Space Model technique from the text mining field (Senin & Malinchik 2013b). Conventionally this method is utilized as a classification model to predict which class a certain time-series belongs. A by-product of the process is that the subsequences of

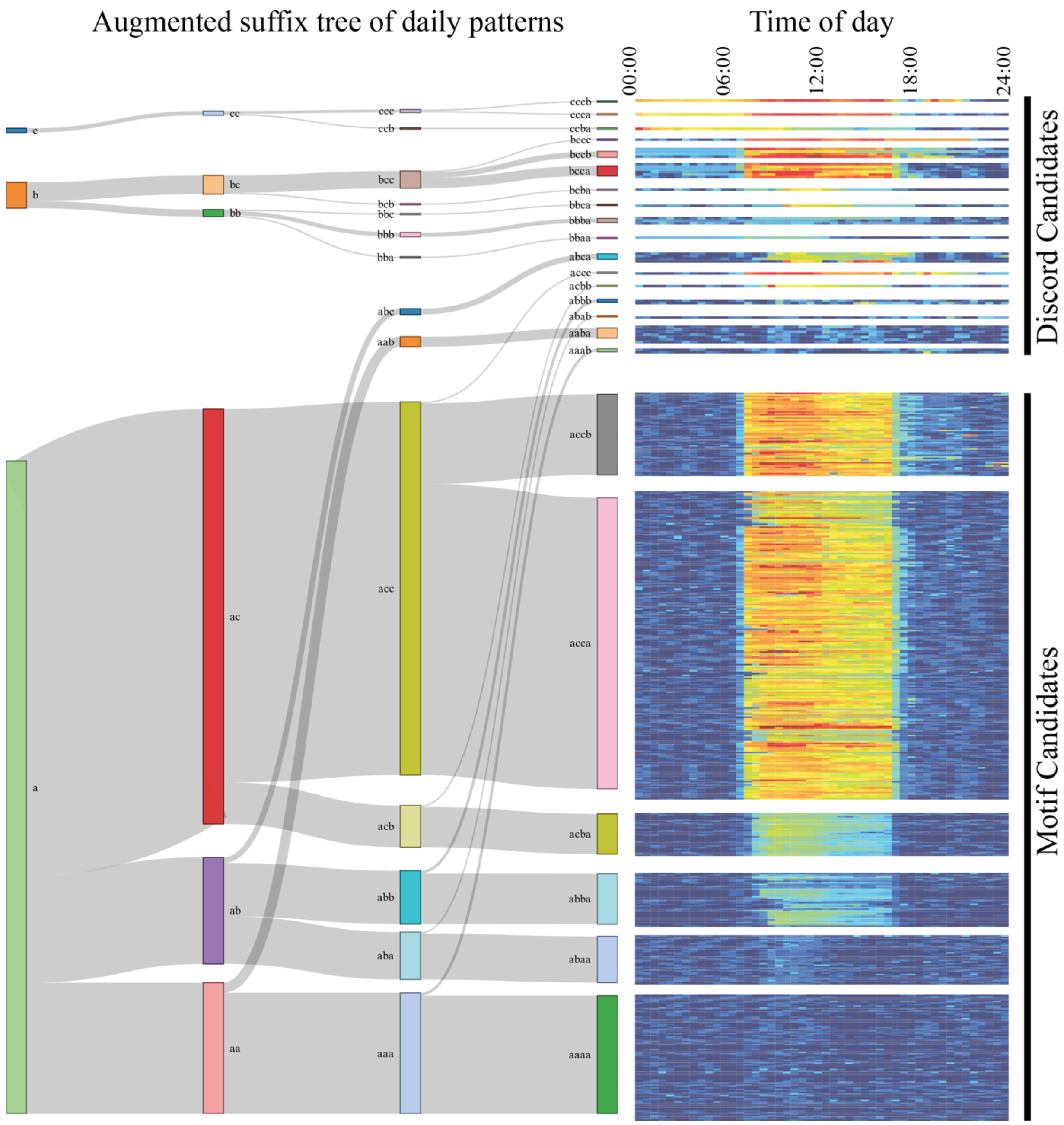

Figure 6.7: Cooling electricity consumption representation of the day-types from the Day-Filter process (Miller *et al.* 2015)

each data stream are assigned a metric indicating their specificity. Pattern specificity is a concept that quantifies how well a meter *fits within its class*. This technique is used to determine whether a building is operating similar to other supposed peer buildings of the same type.

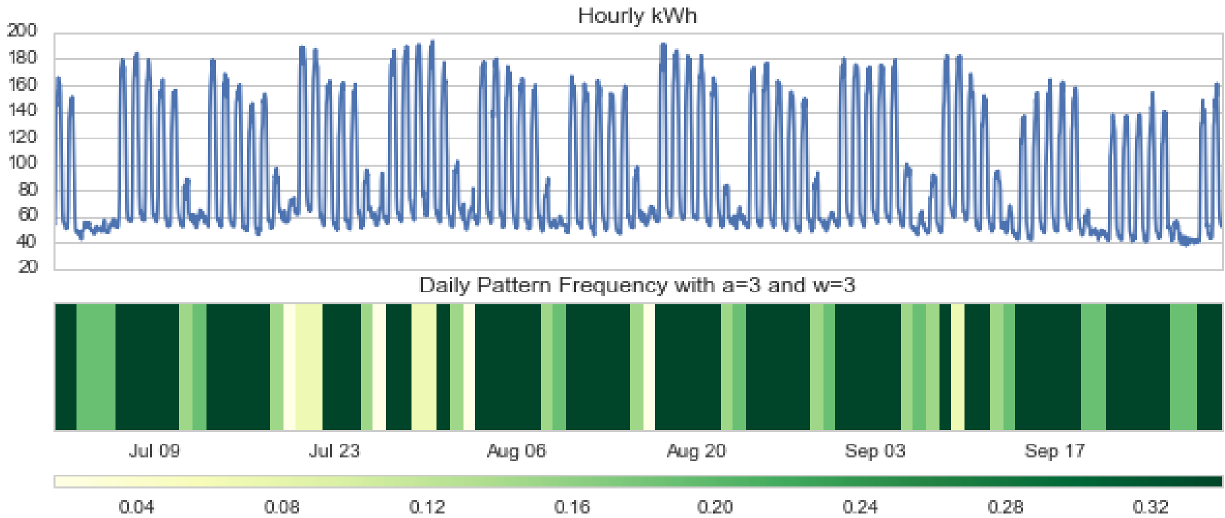

Figure 6.8: Single building example of daily pattern frequency using *DayFilter*, a=3 and w=3

The SAX-VSM process begins with the SAX word creation, similar to *DayFilter* as shown in Figure 6.4. However, the key difference is that the conventional SAX process extracts word patterns from overlapping windows as opposed to simply *chunking* each daily profile. Each data stream within a particular class of a training data set is converted to SAX words using the same input variables of alphabet size, A, and subsequence period count, W. In addition, a P variable is chosen to indicate the size of the sliding window. With SAX-VSM, all of the SAX words for a certain use type class, such as Offices, are then combined into a large Bag of Words (BOG) representation called a corpus, and then used to build a term frequency matrix. This model is then used to calculate a $tf * idf$ weight coefficient, which is the product of the term frequency (tf) and the inverse document frequency (idf). The term frequency is a logarithmically scaled metric based on the incidence of a pattern in the BOG. The inverse document frequency is computed as the log of the ratio of the number of classes to the number of bags where each pattern occurs (Manning *et al.* n.d.). Once this matrix of weight vectors is computed, the cosine similarity of an individual data stream can be calculated to determine how similar to each class it is.

In this study, the goal is not to use SAX-VSM to classify each data stream, but to extract instead temporal features that can be used to characterize them. Thus, the in-class cosine similarity is calculated for each building's data set as compared to the class it was assigned. This process is not conventional from the classification sense as it is considered over-fitting due to all samples being included in the training set. This situation is tolerated in this analysis as it is desired to quantify only how much the patterns of use for a building

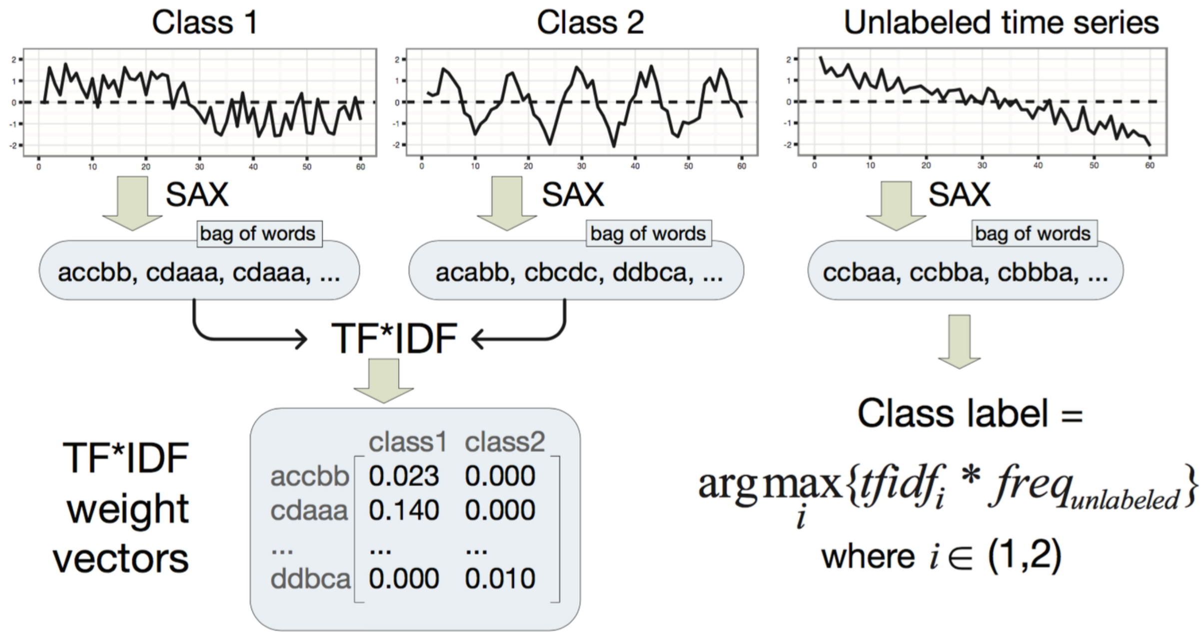

Figure 6.9: Overview of SAX-VSM algorithm: first, labeled time series are converted into bags of words using SAX; secondly, $tf * idf$ statistics is computed resulting in a single weight vector per training class. For classification, an unlabeled time series is converted into a term frequency vector and assigned a label of a weight vector which yields a maximal cosine similarity value (figure and caption used with permission from (Senin & Malinchik 2013a)).

compare to those of its labeled *peers*.

The specificity metric for each data stream is calculated for each sliding window by subtracting all other $tf * idf$ weights for each pattern from the in-class weighting. An example of this weighting

The specificity calculation process is implemented on each of the building test data sets. A single building example of this process is seen in Figure 6.11. This building is within the *Office* use-type classification; thus the color spectrum indicates how precise each subsequence is to this building's behavior as an office as compared to the entire training data set. This example is using the input metrics of $a = 8$, $p = 8$, and $w = 24$ to capture the specificity of daily patterns. These parameters settings include the use of a 24-hour sliding window that is divided into eight segments of three hours length, and the normalized magnitude assigns a symbol from a range of eight letters, a, b, c, d, e, f, g, h.

The specificity calculation process also implemented using input parameters designed to capture patterns of weekly behavior. In this situation, the input metrics of $a = 6$, $p = 14$,

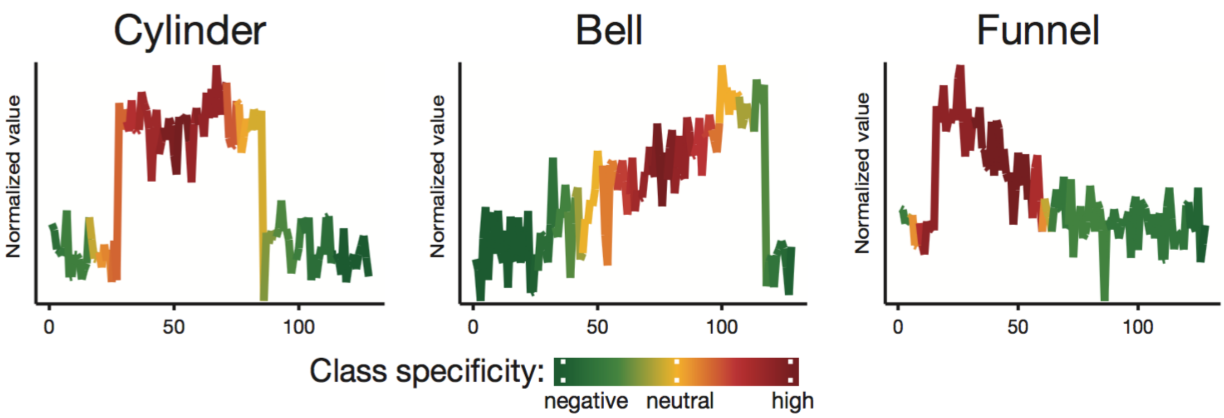

Figure 6.10: An example of the heat map-like visualization of subsequence *importance* to a class identification. Color value of each point was obtained by combining $tf * idf$ weights of all patterns which cover the point. The highlighted class specificity corresponds to a sudden rise, a plateau, and a sudden drop in Cylinder; to a gradual increase in Bell; and to a sudden rise followed by a gradual decline in Funnel (figure and caption used with permission from (Senin & Malinchik 2013a))

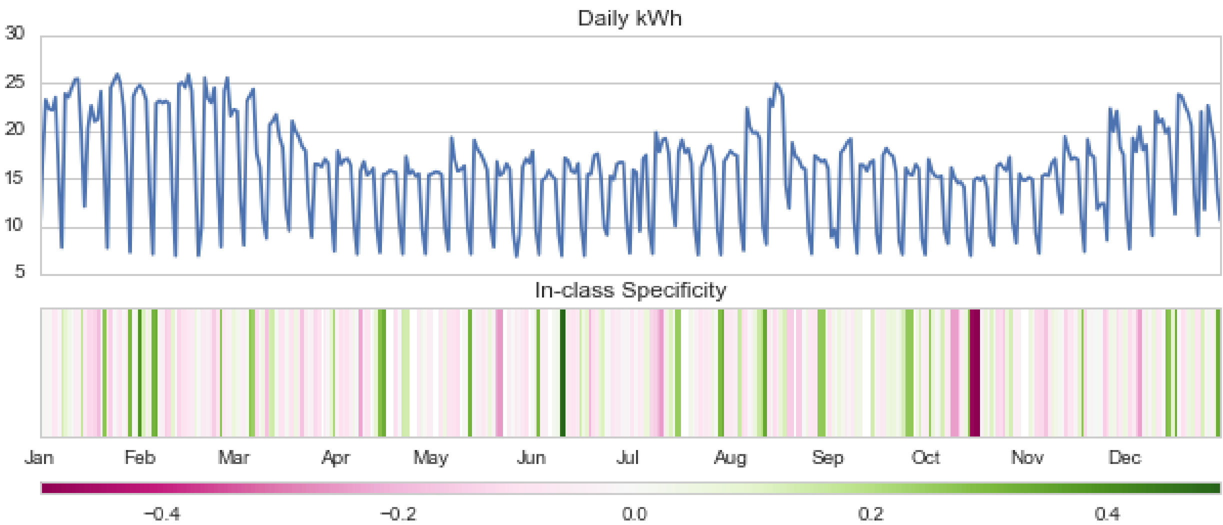

Figure 6.11: Single building example of daily in-class specificity, a=8, p=8, and w=24 for an office building. Positive specificity indicates behavior that is characteristic of a certain class, while negative values indicates behavior of a different class.

and $w = 168$ are chosen to capture this behavior. These parameters settings model a

168-hour sliding window (one week) that is divided into 14 segments of 12 hours length, and the normalized magnitude assigns a symbol from a range of six letters, a, b, c, d, e, f. A single building example is seen in Figure 6.12. This building is also within the *Office* use-type classification; thus the color spectrum indicates how precise each weekly subsequence is to this building's behavior as an office as compared to the entire training data set.

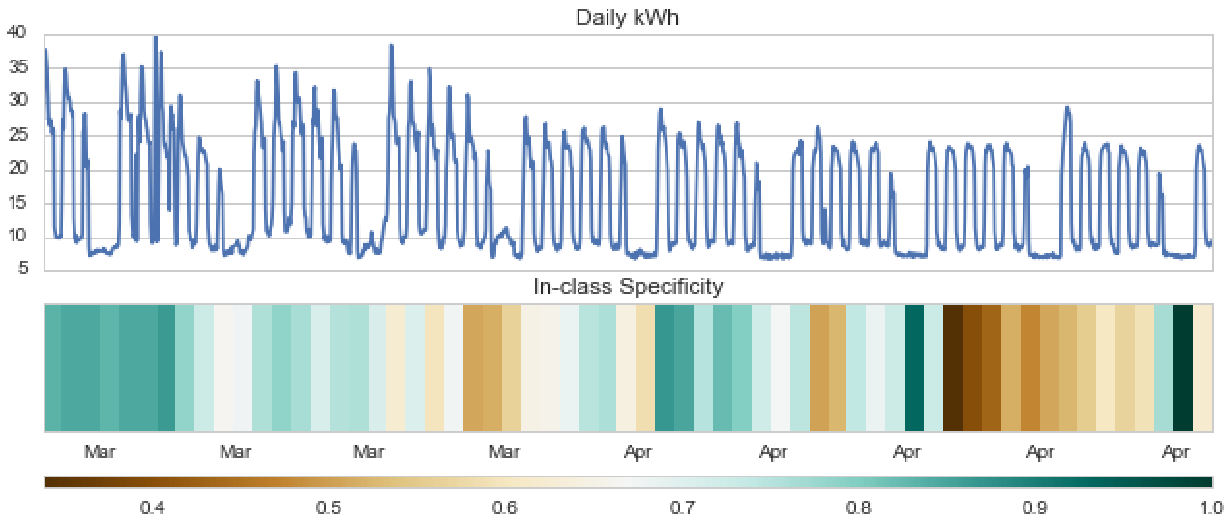

Figure 6.12: Single building example of weekly in-class specificity, a=x, w=X, and p=X

6.1.3 Long-term Pattern Consistency

Breakout detection screening is a process in which each data stream is analyzed according to the tendency to shift from one performance state to another with a transition period in between. This metric is used in this context to quantify long-term pattern consistency, and much of the explanation and graphics in this section are from a previous study (Miller & Schlueter 2015). Breakout detection is a type of change point detection that determines whether a change has taken place in a time series dataset. Change detection enables the segmentation of the data set to understand the nonstationarities caused by the underlying processes and is used in multiple disciplines involving time-series data such as quality control, navigation system monitoring, and linguistics (Basseville & Nikiforov 1993). Breakout detection is applied to temporal performance data to understand general, continuous areas of performance that are similar and the transition periods between them.

In this process, an R programming package, *BreakoutDetection*, is utilized, which is also developed by Twitter to process time-series data related to social media postings[1]. This package uses statistical techniques which calculate a divergence in mean and uses robust metrics to estimate the significance of a breakout through a permutation test. The specific technical details of the breakout detection implementation can be found in a study by James et al. (James *et al.* 2014). *BreakoutDetection* uses the E-Divisive with Medians (EDM) algorithm, which is robust amongst anomalies and can detect multiple breakouts per time series. It can identify the two types of breakouts, mean shift and ramp up. Mean shift is a sudden jump in the average of a data stream, and ramp up is a gradual change of the value of a metric from one steady state to another. The algorithm has parameter settings for the minimum number of samples between breakout events that allows the user to modulate the amount of temporal detail.

The goal in using breakout detection for building performance data is to find directly when macro changes occur in sensor data stream. This discovery is particularly exciting in weather-insensitive data to understand when modifications are made to the underlying system in which performance is being measured. Figure 6.13 data from a single building data stream. Each color represents a group of continuous, steady-state operation and each change in color is, thus, a breakout. These breakouts could be the result of schedule or control sequence modifications, systematic behavior changes, space use type changes, etc. Creation of diversity factor schedules should target data streams which have few breakouts and the data between breakouts is the most applicable for model input. One parameter setting for breakout detection is the minimum breakout size threshold. This parameter prevents breakouts from being detected to close together, thus capturing potentially noisy behavior for the particular data set.

6.2 Implementation and Discussion

Figure 6.14 shows this pattern frequency metric as applied to all the case study buildings.

Figure 6.15 illustrates this process applied to all 507 case studies as divided amongst the use types. Clear differences in patterns across the time ranges are visible for each of the building use types. Offices, university laboratories, and university classrooms all seem to have similar phases of specificity at similar times of the year, while dorms and primary/secondary schools are often differentiated by their breaks.

Figure 6.16 illustrates weekly specificity as applied to all the buildings. The transition between specific and non-specific patterns is smoother in this case due to the weekly time

[1]https://github.com/twitter/BreakoutDetection

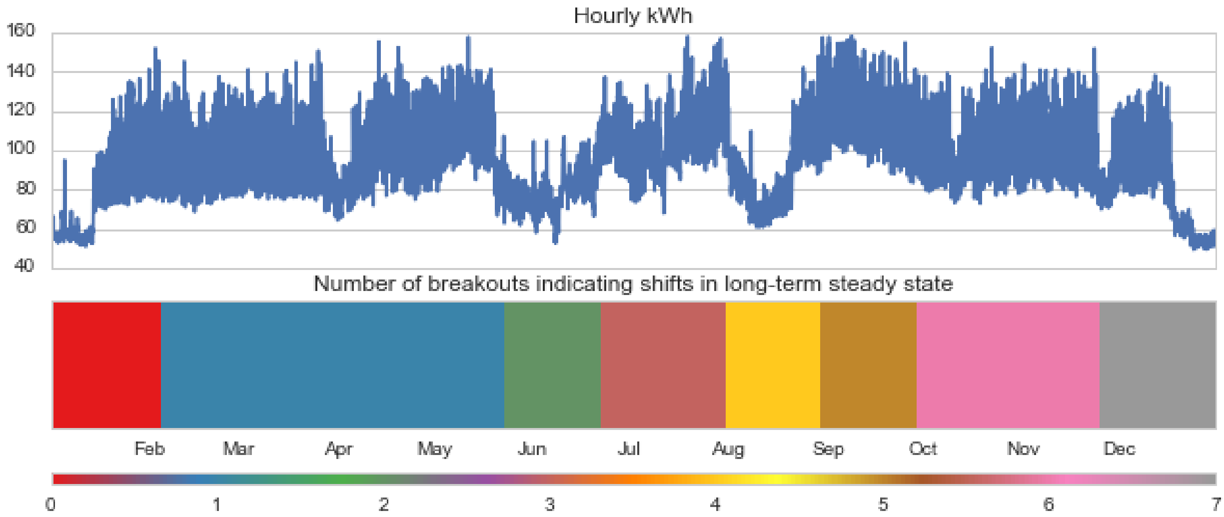

Figure 6.13: Single building example of breakout detection to test for long-term volatility in an university dormitory building. A minimum threshold of 30 days is chosen in this case, which explains the lack of threshold shift in April, a break that may be attributed to spring break for this building

range. It is also apparent that the most distinct behavior patterns for each building use type are correlated to when that particular building has behavior related to lower occupancy such as summer breaks or holiday periods. These phenomena need to be somewhat consistent across all the buildings within a classification for it to indicate specificity.

Figure 6.17 illustrates breakout detection across the building use types in this study. This implementation uses the same input parameter of a 30 day minimum between breakouts. One notices somewhat of consistency amongst offices, labs, and classrooms regarding the distribution of breakout numbers, while university dormitories and primary/secondary classrooms have a noticeably higher number of breakouts across the range of behavior.

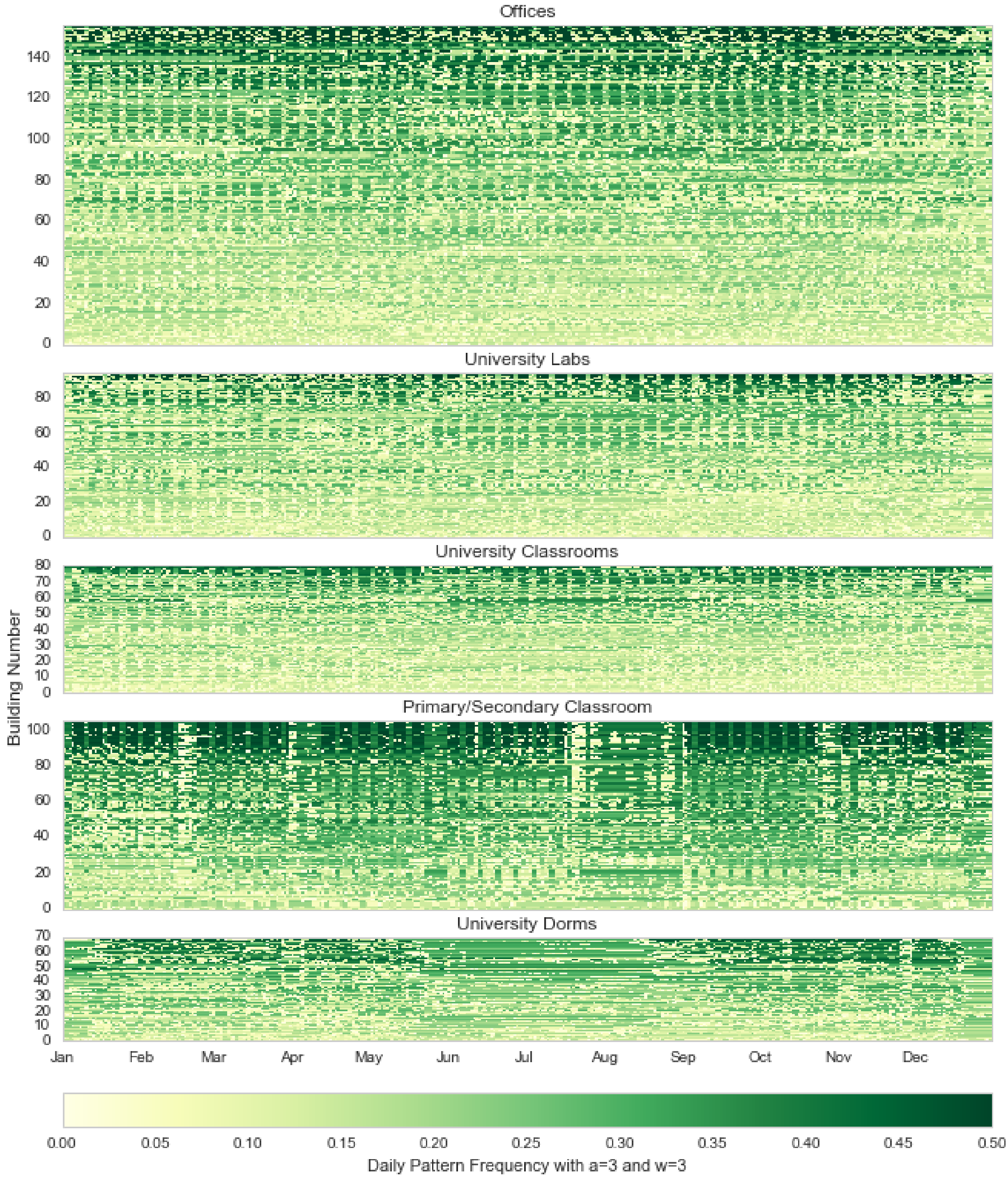

Figure 6.14: Heatmap of daily pattern frequencies using *DayFilter* with a=3 and w=3

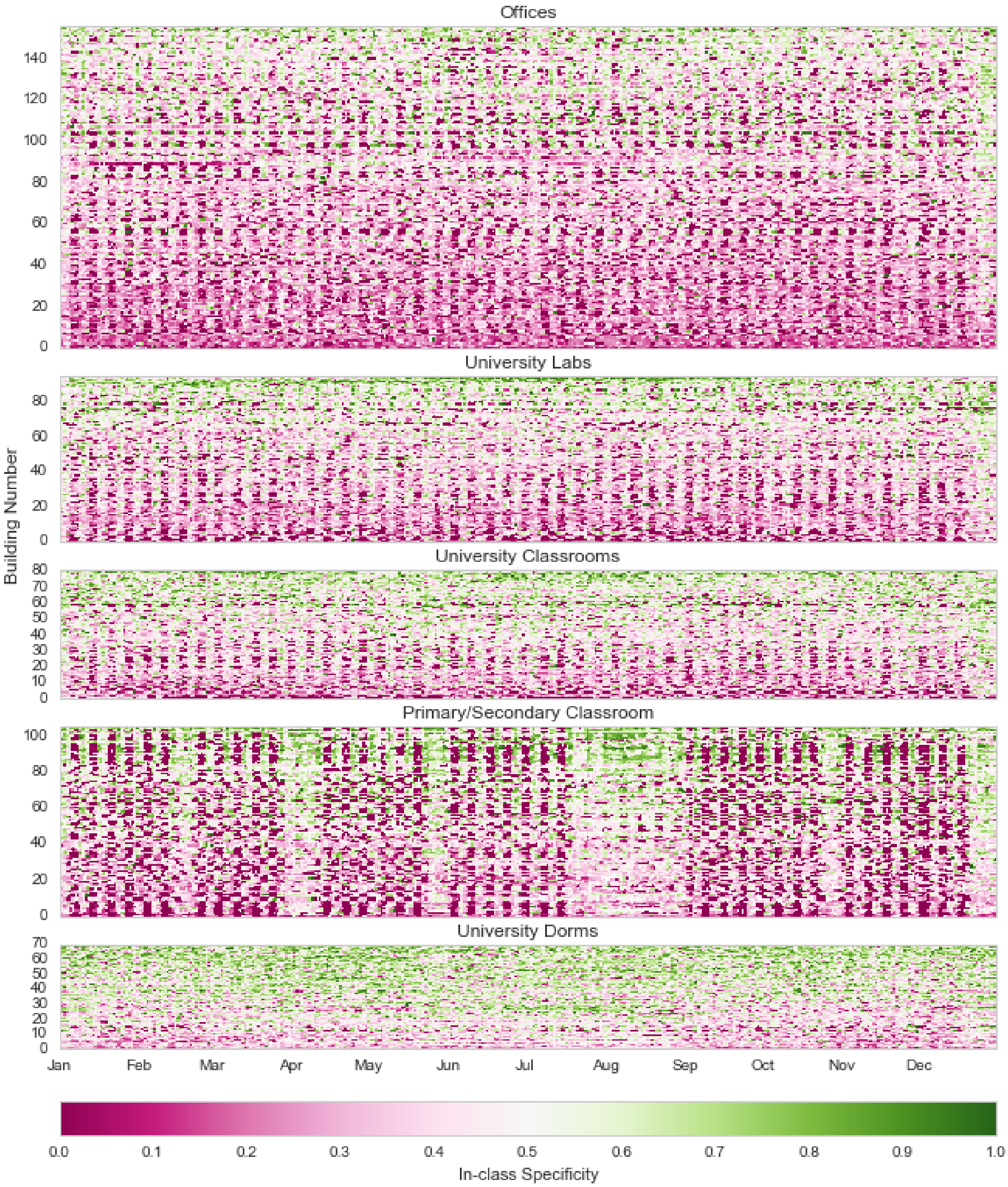

Figure 6.15: Heatmap of in-class specificity with p=24, a=8, w=8

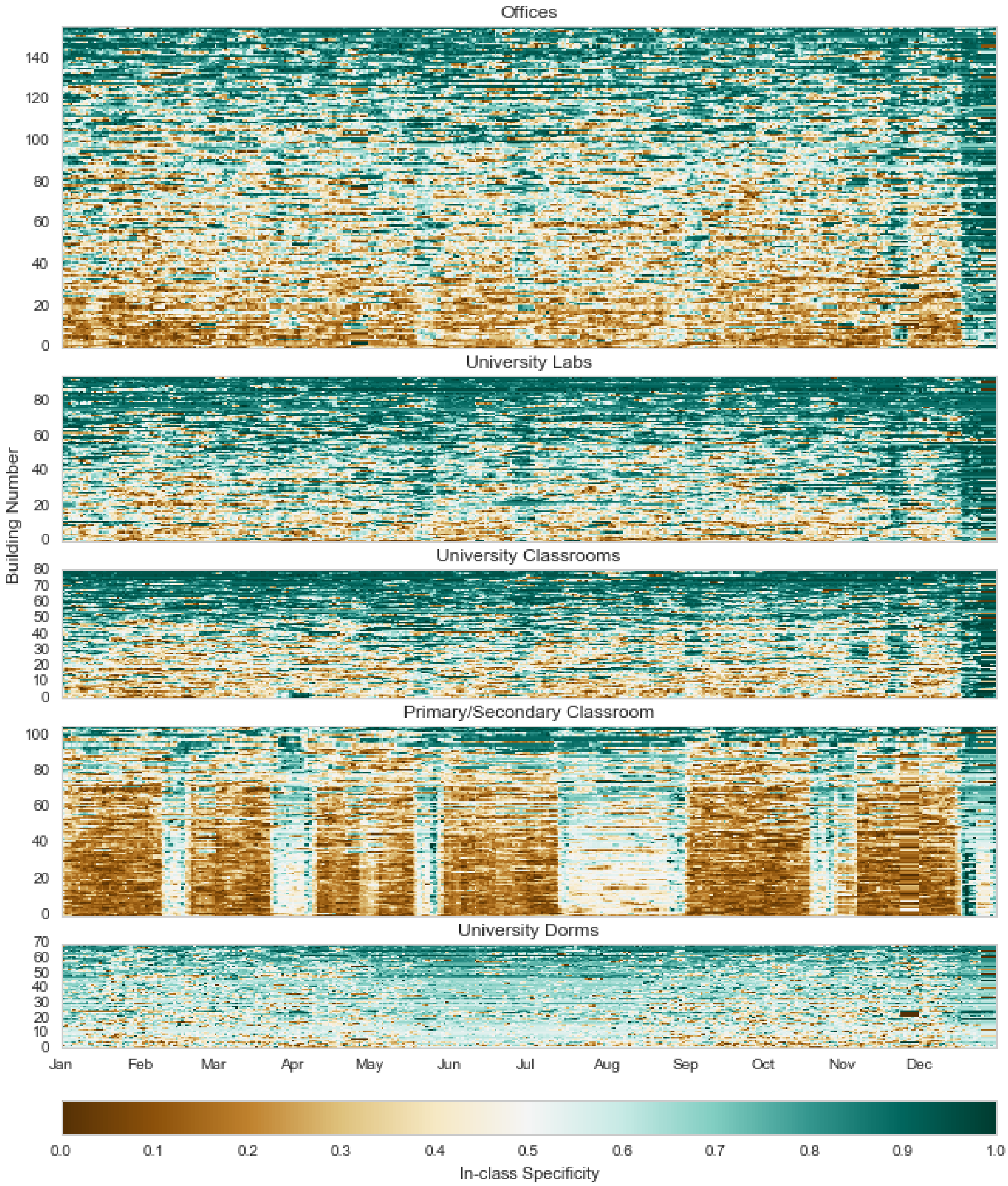

Figure 6.16: Heatmap of in-class specificity with p=168, a=6, w=14

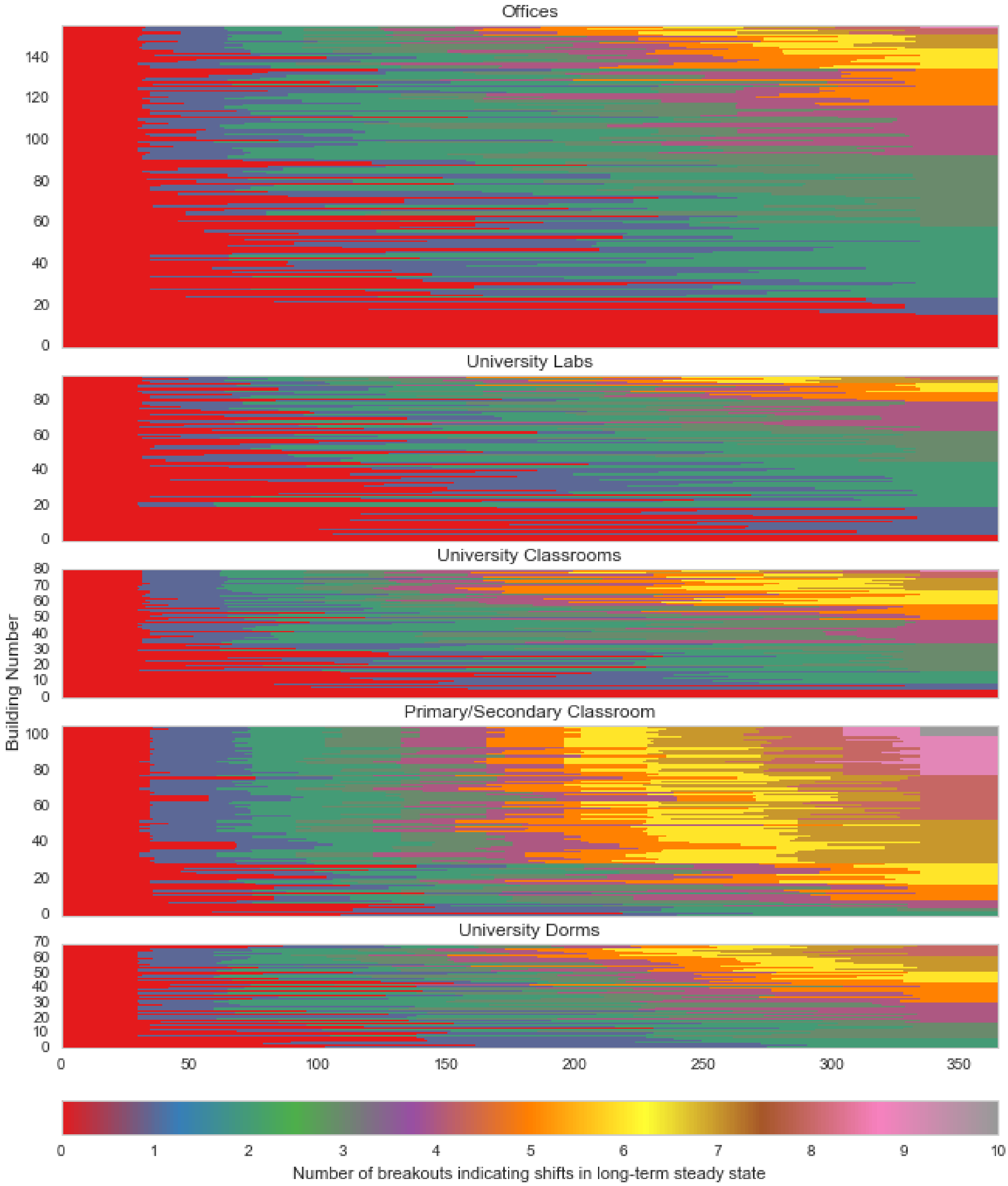

Figure 6.17: Heatmap of breakout detection on all case studies

7 Characterization of Building Use, Performance, and Operations

Visualization of temporal features on their own is a means of understanding the range of values of the various phenomenon across a time range. This situation gives an analyst the basis to begin understanding what discriminates a building based on different objectives. The next step is to utilize the features to predict whether a building falls into a particular category and test the importance of various elements in making that prediction. Understanding which features are most characteristic to a particular objective is the fundamental tenet of this study. In this section, three classification objectives are tested:

1. Principle Building Use - The primary use of the building is designated for the principal activity conducted by percentage of space designated for that activity. It is rare for a building to be devoted specifically to a single task, and mixed-use buildings pose a specific challenge to prediction.
2. Performance Class - Each building is assigned to a particular performance class according to whether its area-normalized consumption in the bottom, middle, or top 33% percentiles within its principle building use-type class.
3. General Operation Strategy - Buildings that are controlled by the same entity, such as those on a University campus, often have similar schedules, operating parameters, and use patterns. This objective tests to understand how distinct these differences are between different campuses.

7.1 Principal Building Use

The first scenario investigated is the characterization of primary building use type. The goal of this effort is to quantify what temporal behavior *is most characteristic in a building being used for a certain purpose*. For example, what makes the electrical consumption patterns of an office building unique as compared to other purposes such as a convenience store, airport, or laboratory. This objective is necessary to understand who are the *peers* of a building. Whatever category a building is assigned determines what benchmark is used to determine the performance level of a building. The EnergyStar Portfolio Manager is the most common benchmarking platform in the United States and the first step in its evaluation is identifying the property type. There are 80 *property types* in portfolio manager and each one is devoted to a particular primary building use type. Twenty-one of those property types are available for submission to achieve a 1-100 ENERGYSTAR score in the United States. These property types are seen in Figure 7.1.

In the United States:	In Canada:
Bank branch	Financial office
Barracks	K-12 school
Courthouse	Hospitals
Data center	Medical office
Distribution center	Office
Financial office	Residential care facility
Hospital (general medical & surgical)	Supermarket/Grocery store (covers supermarket/grocery store, food sales, and convenience store with or without gas station)
Hotel	
K-12 school	
Medical office	
Multifamily housing	
Non-refrigerated warehouse	
Office	
Refrigerated warehouse	
Residence hall/ dormitory	
Retail store	
Senior care community	
Supermarket/grocery store	
Wastewater treatment plant	
Wholesale club/supercenter	
Worship facility	

Figure 7.1: EnergyStar building use-types available for 1-100 rating

Allocation of the primary use type of a building is often considered a trivial activity when analyzed from a smaller set of buildings. As the number of building being analyzed grows, so does the complexity of space use evaluation. The use of buildings changes over time and these changes are not always documented. In several of the case studies, this topic was discussed and highlighted as an issue concerning benchmarking a building.

Discriminatory features have already been visualized extensively in Section ?? and the differences between the primary use types are apparent in the overview heat maps of each feature. In this and the following sections, a quantification of the impact of each feature will be evaluated using a random forest model and its associated variable importance methods. Figure 7.2 is the first such example of the output results of the classification model in predicting the building's primary use type using the temporal features created in this study. This visualization is a kind of error matrix, or confusion matrix, that illustrates the performance of a supervised classification algorithm. The *y-axis* represents the correct label of each classification input and the *x-axis* is the predicted label. An accurate classification would fall on the left-to-right diagonal of the grid. This grid is normalized according to the percentage of buildings within each class. The model was built using the scikit-learn Python library[1] with the number of estimators set to 100 and the minimum samples per leaf set to 2. The overall general accuracy of the model is 67.8% as compared to a baseline model of 22.2%. The baseline model using a stratified strategy in which categories are chosen randomly based on the percentage of each class occurring in the training set. Based on the analysis, university dormitories and primary/secondary classrooms are the best-characterized use types overall with precisions of 92% and 96% respectively and accuracies of 74% and 75%. The office category is easy confused with university classrooms and laboratories. This situation is not surprising as many of these facilities are quite similar and uses within these categories often overlap.

The most important features contributing to the accuracy of the classification model are found in Figure 7.3. These features are ranked according to their importance in designating the difference between all of the building types. Three of the top fifteen most important features are from the *stl* decomposition process. This fact shows the importance that normalized weekly patterns play in differentiation, in particular for dormitories. Eight of the fifteen are statistical metrics, either ratios or consumption statistics. The second highest variable importance is related to the correlation output from the loadshape model. And the remaining three variables pertain to the number of long-term breakouts, and thus, volatility.

7.1.1 University Dormitory and Laboratory Comparison

The random forest classification model and variable importance metrics provide an indication of how the features characterize a building's use. A deeper investigation of the

[1] http://scikit-learn.org/

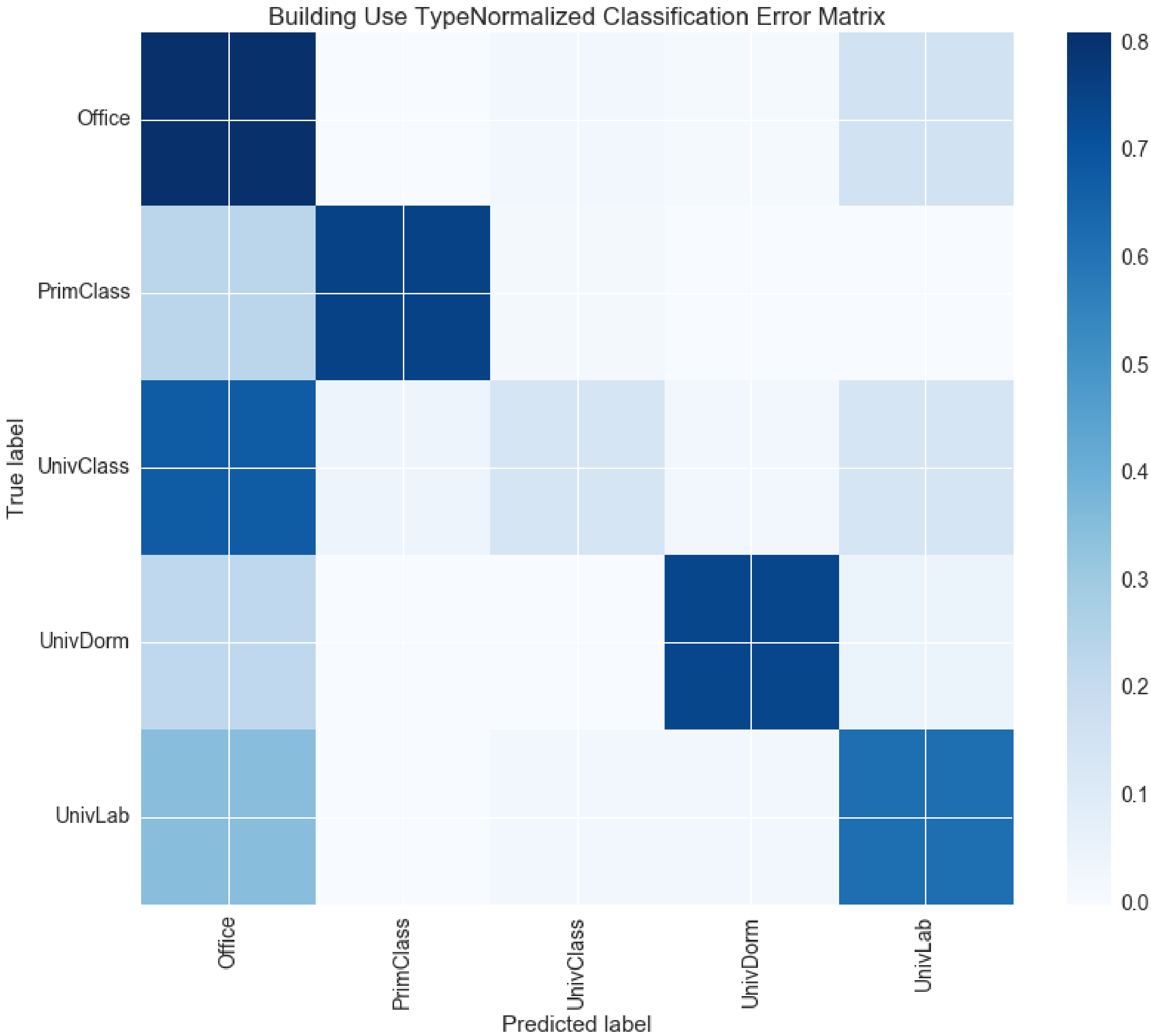

Figure 7.2: Classification error matrix for prediction of building use type using a random forest model

features with a comparison between two use types is useful to understand the characterization potential of various subsets of features. For this example, two building type classifications are compared that showed sharp distinction from each other in the random forest model: university laboratories and dormitories. For this comparison, the highly comparative time-series analysis (hctsa) code repository is used as a toolkit for analysis of the generated temporal features in this study Fulcher *et al.* (2013). This toolkit has various visualization tools that enable analysis of the predictive capabilities of temporal features. Figure 7.4 shows the top forty features in differentiating university laboratories and dormitories using a simple linear classifier model. These features are clustered ac-

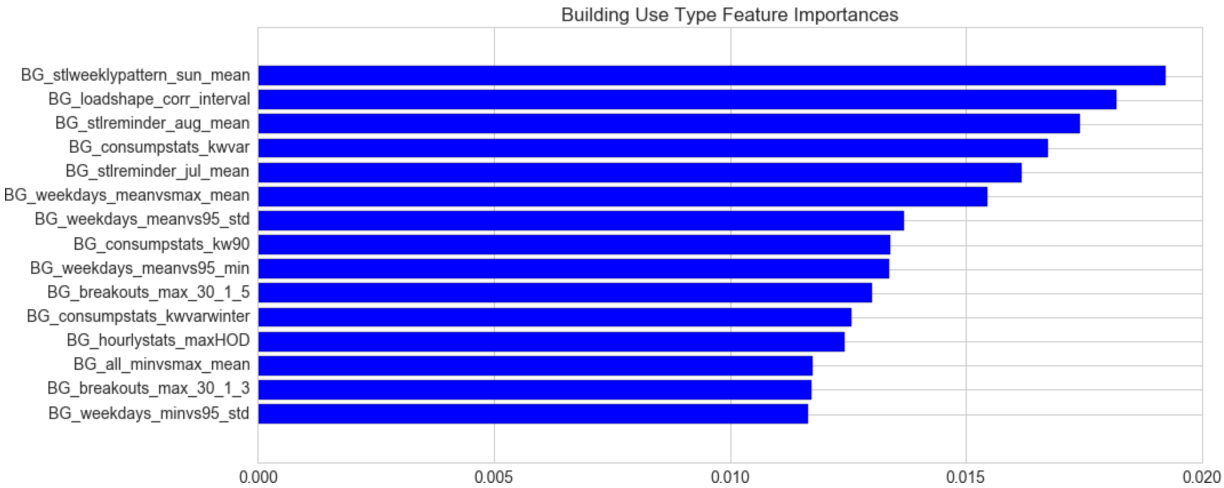

Figure 7.3: Importance of features in prediction of building use type

cording to their absolute correlation coefficients to understand how many unique sets of informative features are present. Groups of features in the same cluster are essentially giving the same type of information about the differences between a certain set of tested classes. In the case of laboratories and dormitories, there are eight sets of clusters giving information about this distinction. The first, fourth and fifth clusters contain a couple of breakout metrics representing volatility. The second and third clusters represent magnitudes of cooling energy and consumption statistics. The sixth cluster represents seasonal metrics. The seventh cluster is a collection of fourteen features that are highly correlated, with most being related to daily ratios and consumption-related metrics. The eighth and last cluster include fifteen features with several also representing consumption metrics and ratios, but also several related to daily pattern frequencies.

Figure 7.5 illustrates the probability distributions of the top five differentiating features for distinguishing laboratories from dormitories. The probability density of each of the features is relatively similar in shape and distribution. This situation is because most of the features are from clusters seven and eight which are highly correlated within the cluster and between the clusters as well.

Figure 7.6 shows a distribution of the library of features on the data set compared to a benchmark of nulls generated by randomly selecting the class. This visualization indicates that there is a clear statistical difference in discriminating these two categories for a significant number of the input features. The real mean is approximately 62%, while the null mean is slightly above 50%. The ability to distinguish between these two classes is relatively high.

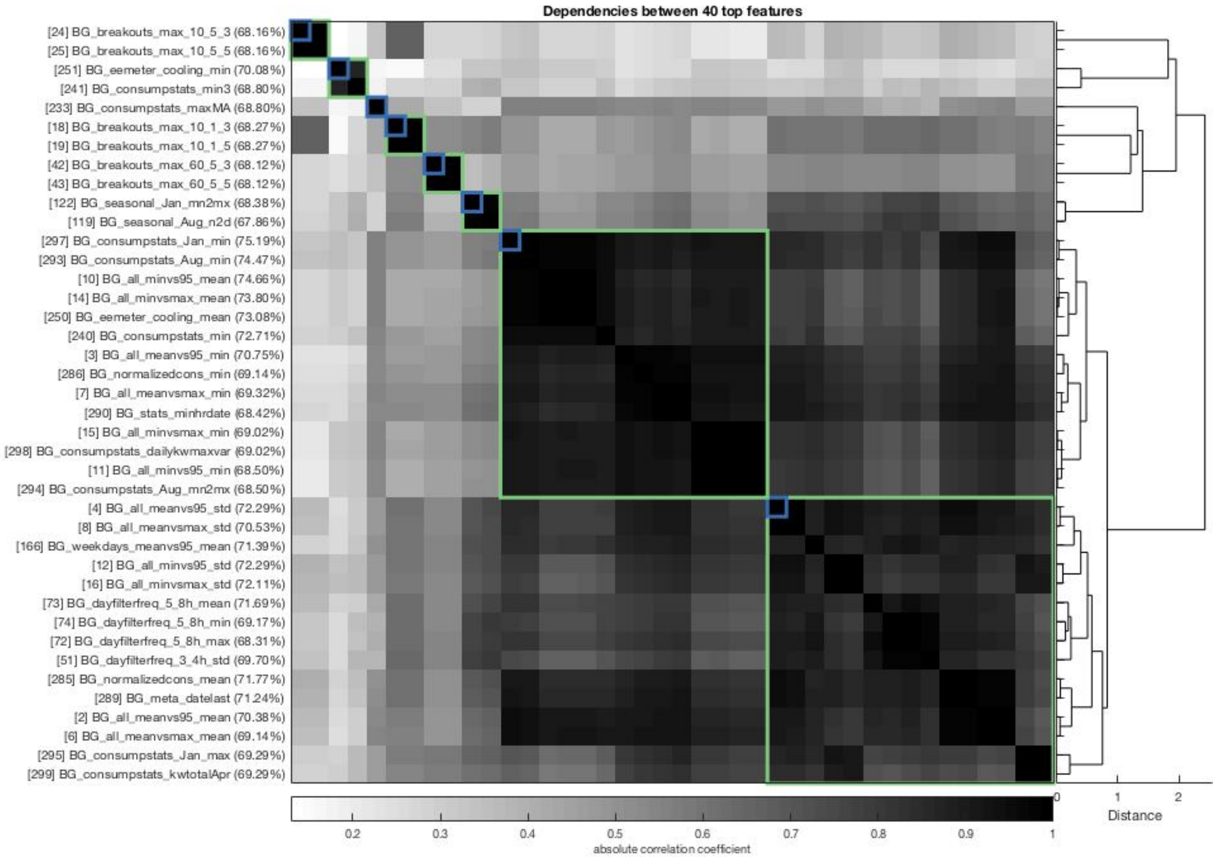

Figure 7.4: Clustering of dominant features in the comparison of university dormitories and laboratories

7.1.2 Discussion with Campus Case Study Subjects

Previously, an example of how to characterize building use type was illustrated using a random forest model and various feature importance techniques. In this subsection, a discussion is presented of how this sort of characterization can be useful in a practical setting. In the case study interviews, the topic of benchmarking of buildings was discussed. One of the issues presented to the operations teams was the concept of not having a complete understanding of the way the buildings on their campus were being used. For example, several of the campuses have a spreadsheet outlines various metadata about the facilities on campus. This worksheet, in many cases, includes the *primary use type* of the building. It was found that this primary use type designation is often loosely based on information from when the building was constructed or through informal site survey. In other situations, the building has an accurate breakdown of all the sub-spaces in the building and approximately what the spaces are being used for. In these discussions, the

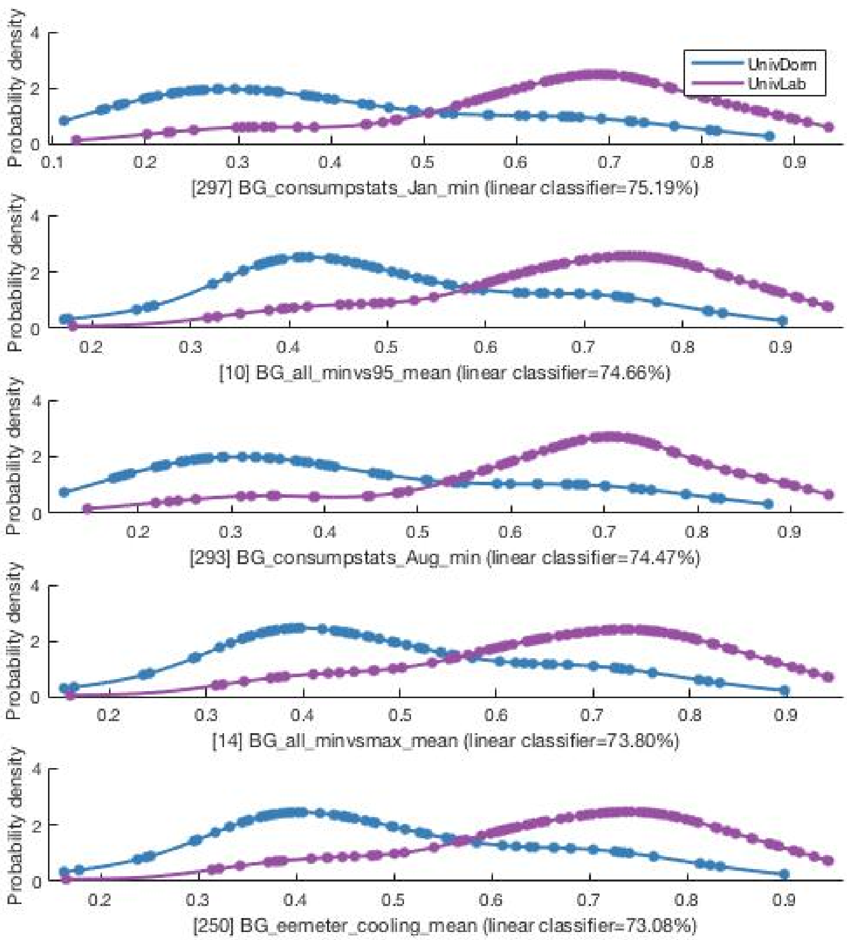

Figure 7.5: Probability density distribution of top five features in characterizing the difference between university dormitories and laboratories

idea was presented that building use type characterization could be used to determine automatically whether the labels within these spreadsheets aligned with the patterns of use characterization using the temporal feature extraction. This proposal was met some positive feedback, albeit there was a hesitation to confirm fully that this process would be entirely necessary if labor were directed to do the same task.

Many of the case study subjects then were shown a series of graphics designed to tell the story of building use type characterization in an automated way. Figure 7.7 is the first graphic shown to the subjects. This figure illustrates several of the most easily understood temporal features and how they break down across the various building use types. This graphic was created using the data for a particular case study; therefore more separation between the classes exist than in the prediction of classes found in the previous section.

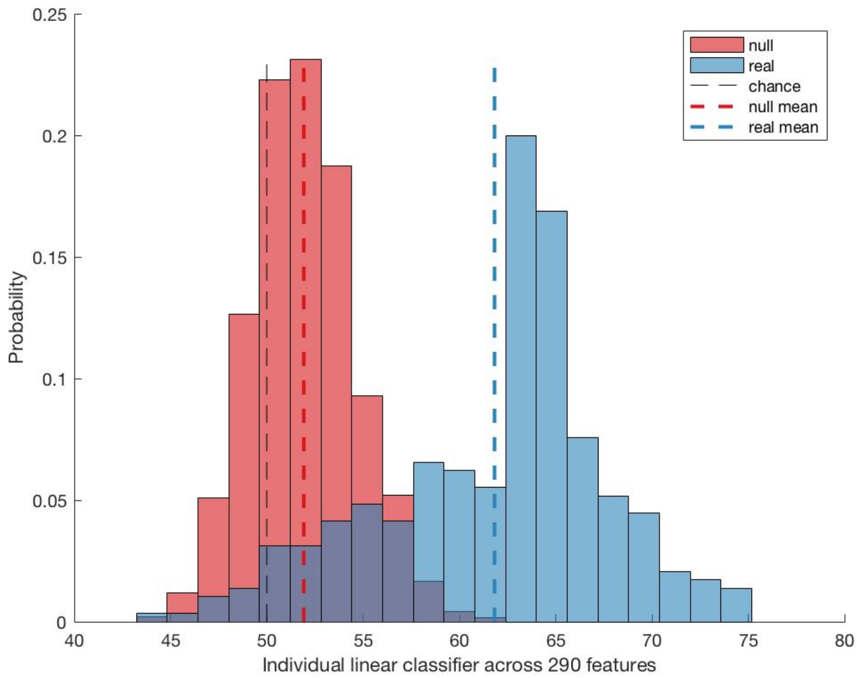

Figure 7.6: Ability of temporal features to distinguish between dormitories and laboratories as compared to the null hypothesis

Discussions using this graphic first centered around the first feature: *Daily Magnitude per Area*. It was intuitive to most participants that a university laboratory has more and primary/secondary schools have less consumption per area than the other use types. It is more surprising, however, that certain building use types are characterized well by other features, such as a number of breakouts with primary/secondary schools and daily and weekly specificity with university dormitories.

After a discussion of how different use types of buildings are characterized using temporal features, the concept of misclassified buildings was introduced. Misclassification of buildings pertains to when the primary use type of the building doesn't match the temporal features of the electrical consumption, particularly the daily and weekly patterns of use. Figure 7.8 was designed to illustrate this concept. This figure contains a subset of the case study buildings within the office, university classroom, and university laboratory categories. The pattern specificity for offices, classrooms and laboratories were calculated for each building as shown in the first three columns of the graphic. They are clustered according to their similarity with red indicating low values and blue indicating high values. The column on the far right indicates the use type classification for each building. The laboratories are yellow, classrooms are blue, and offices are green. It can be seen that there are distinct clusters of building types and a few regions in which there is a mix of

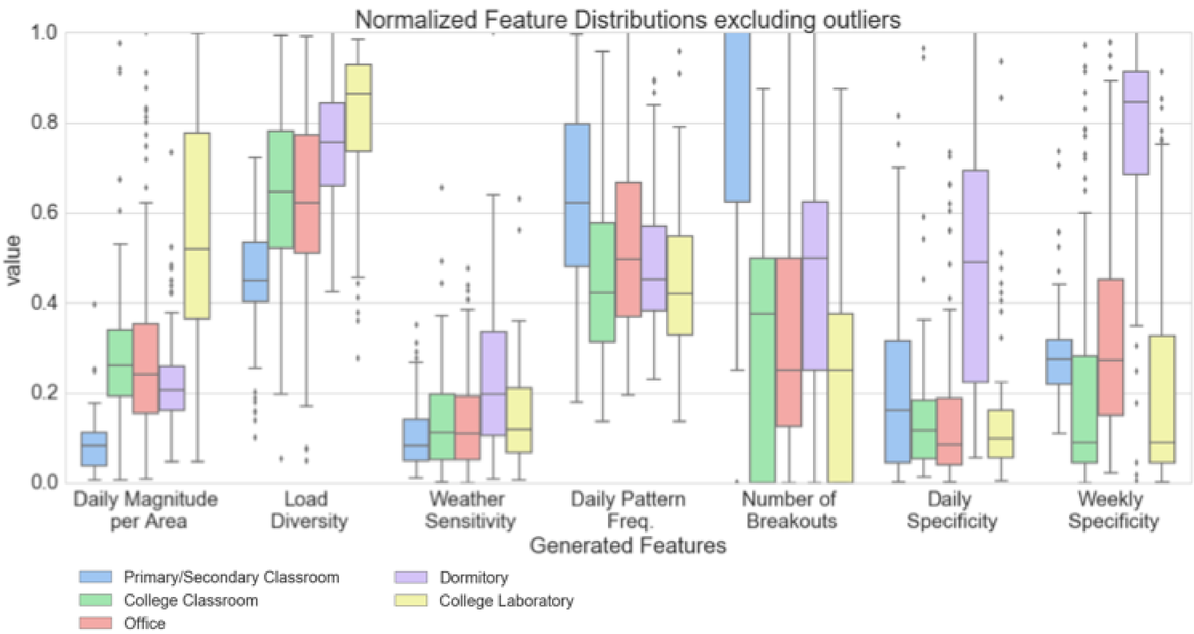

Figure 7.7: Simplified breakdowns of general features according to building use type that were presented to case study subjects

building use types in the final column.

Figure 3.7 shows the same diagram zoomed in on a certain subsection of a cluster that contains mostly buildings that identify as *classrooms*. Interspersed amongst these classrooms are several buildings labeled as *offices*. These offices can be potentially thought of as *misfits* in that they are not members of more consistently homogeneous clusters. Discussions with members of the case study groups revealed that this information is *interesting*, but immediately there wasn't a clear understanding of how this information would influence decision-making. It was suggested that this information could be used to supplement the results of the benchmarking process by giving more insight into potentially *why* a building is not performing well within its class. The situation may actually be that the building is more a member of a different class and therefore may not be comparable to those particular *peers*.

7.2 Characterization of Building Performance Class

The second objective targeted in this study is the ability for temporal features to characterize whether a building performs well or not within it use-type class. Consumption

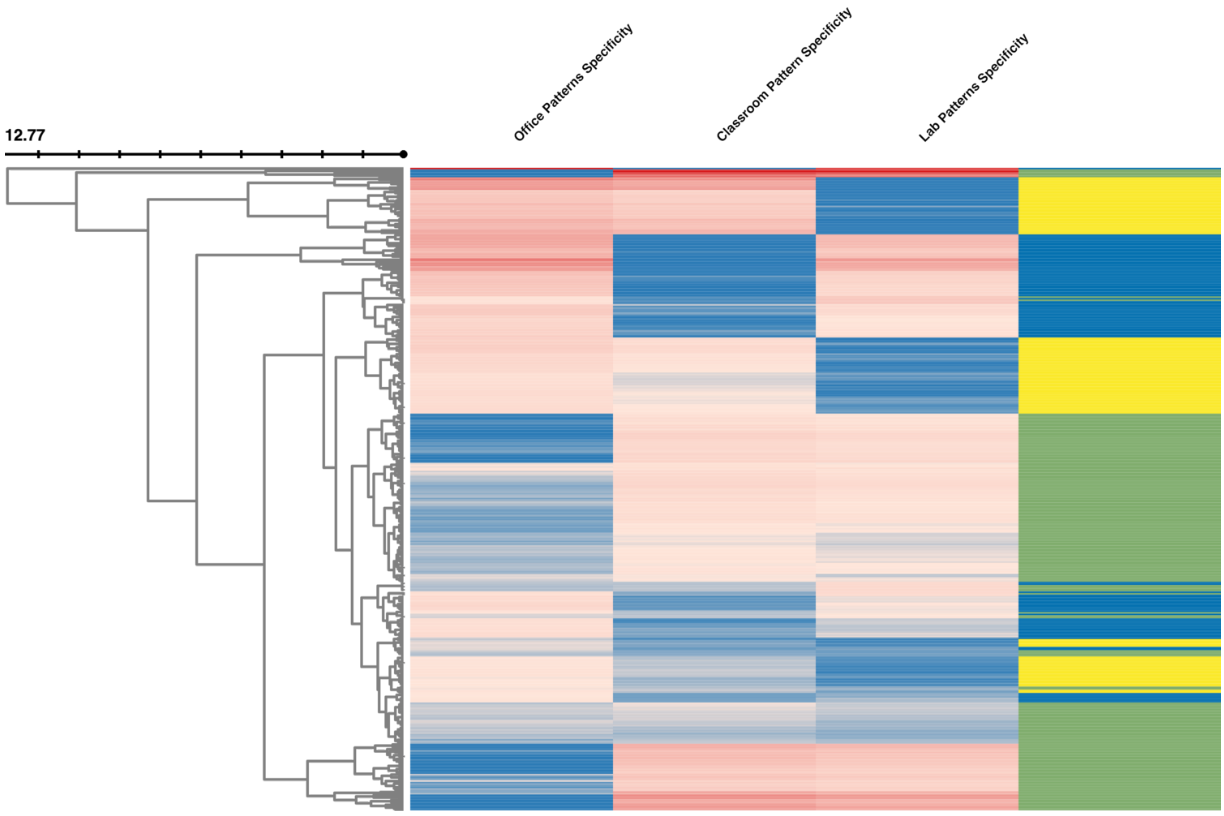

Figure 7.8: Hierarchical clustering of buildings according to laboratory (yellow), office (green), and classroom (blue) specificity

is the metric being measured; therefore it's not the goal of this analysis to predict the performance of a building, its to determine which temporal characteristics are correlated with good or poor performance. This effort is related to the process of benchmarking buildings. Using the insight gained through characterization of building use type, it is possible to inform whether a building's behavior matches its peers. Once a building is part of a peer group, its necessary to understand how well that building performs within that group. In this section, the case study buildings are divided according to which percentile each fits within in its in-class performance. The buildings are divided according to percentiles, with those in the lowest 33% are classified as "Low", the 33 to 66% percentile are "Intermediate", and the top 33% are classified as "High". As in the previous section, these classifications and a subset of temporal features are implemented into a random forest model to understand how well the features are at characterizing the different classes. Since this objective is related to consumption, all input features with known correlations to consumption were removed from the training set. These include the obvious features of consumption per area, but also include many of the statistical metrics such as maximum

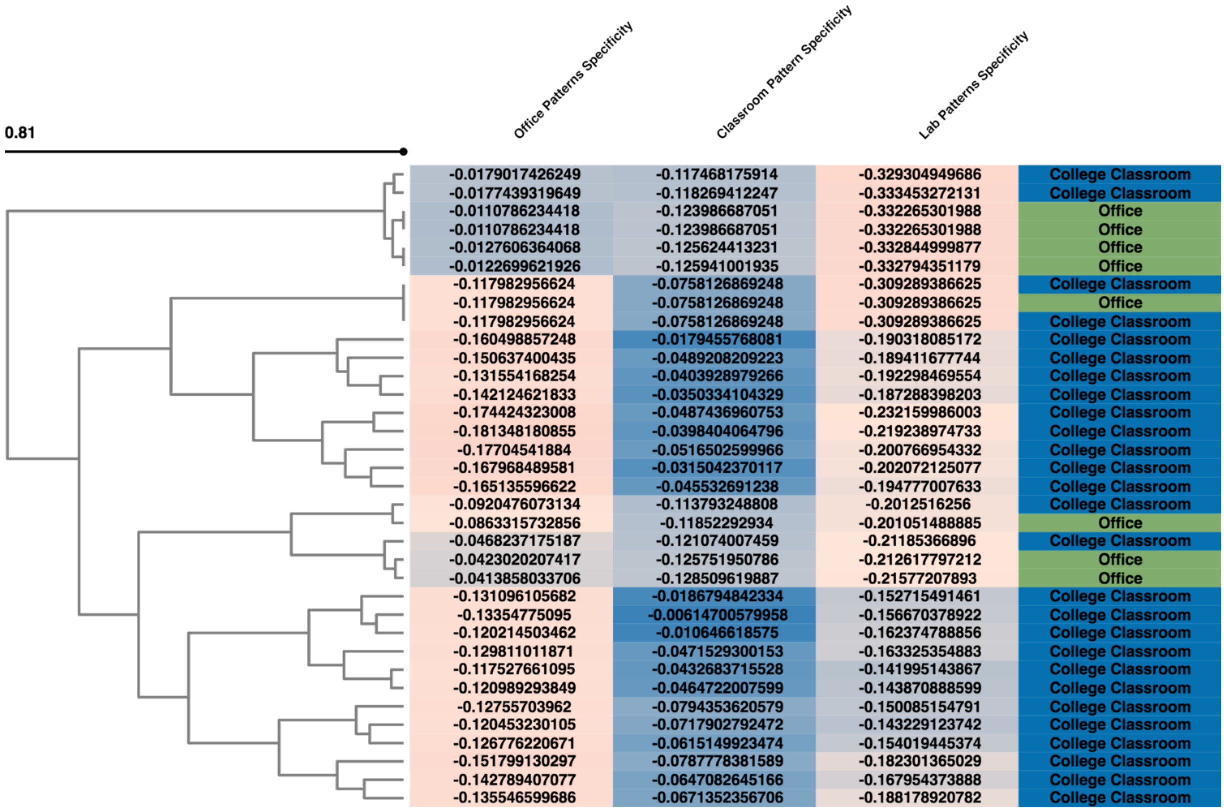

Figure 7.9: Hierarchical clustering of buildings according to laboratory, office, and classroom specificity zoomed in on a cluster with illustrates *misfits*

and minimum values. Most of the daily ratio input features remain in the analysis as they are not directly correlated with total consumption. Figure 7.10 illustrates the results of the model in an error matrix. It can be seen that *high* and *low* consuming buildings are well characterized. The *intermediate* buildings have higher error rates and are often misclassified with the other two classes. The overall accuracy of the model for classification is 62.3% as compared to a baseline of 38%.

Figure 7.11 shows the variable importance calculation as it relates to classification for all three classes. The top features for this model are a mix of statistical features and model-based features. Within the statistical features category, the seasonal range for both winter and summer are top features in addition to several daily ratios. For model-based features, the loadshape model errors, the *stl* model residuals, and the eemeter residuals are all present.

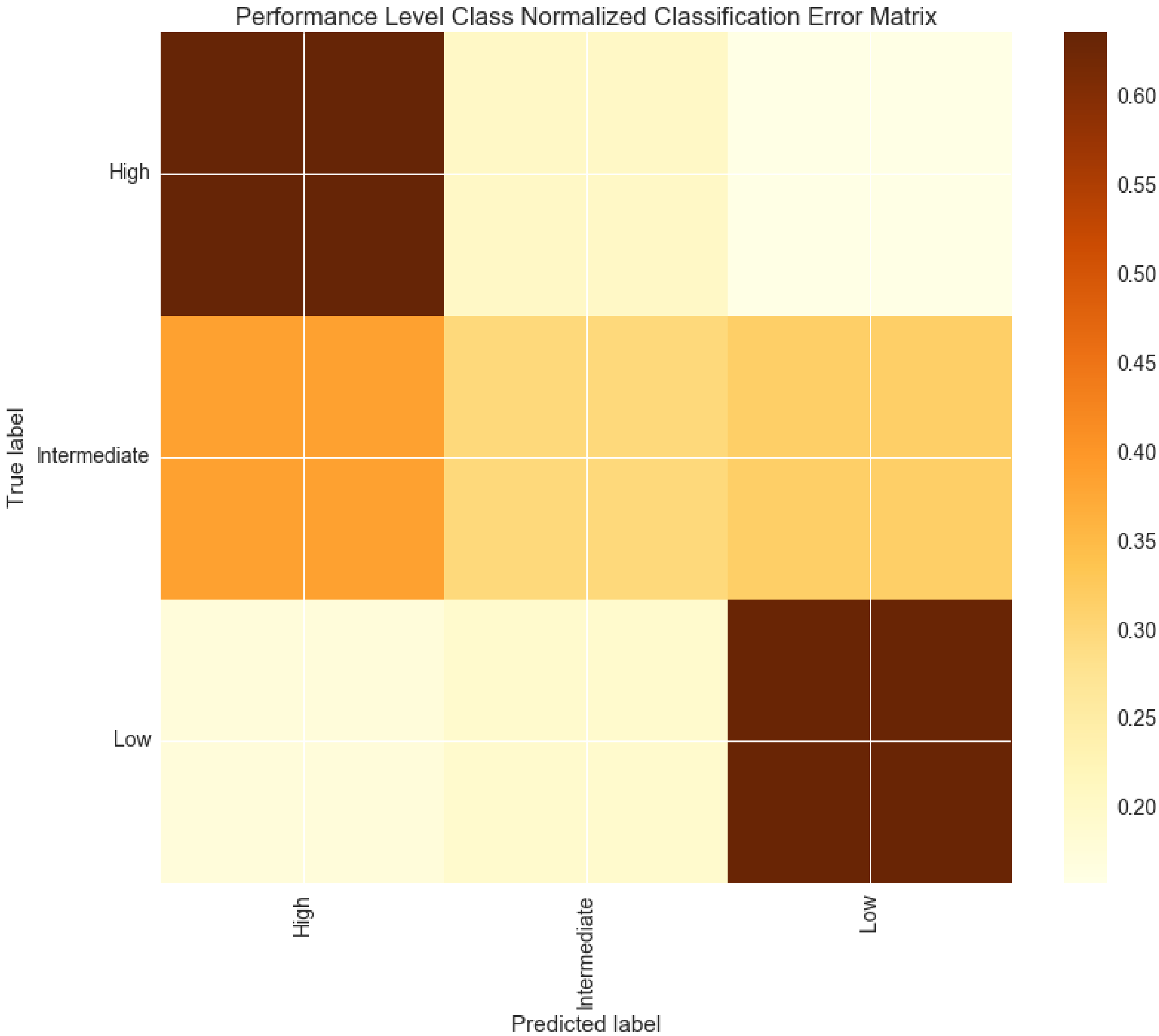

Figure 7.10: Classification error matrix for prediction of performance class using a random forest model

7.2.1 High versus Low Consumption Comparison

The two classifications chosen for this objective are intuitively the *high* and *low* consuming buildings. This part of the analysis gives a more in-depth perspective of exactly which features are most important in the differentiation between these two types of buildings. This understanding provides insight on potentially what behavior in a building results in good or poor performing buildings. Once again, the highly comparative time-series analysis (hctsa) code repository is used for this process. Figure 7.12 is a correlation matrix showing the top forty features as determined by hctsa according to the in-sample

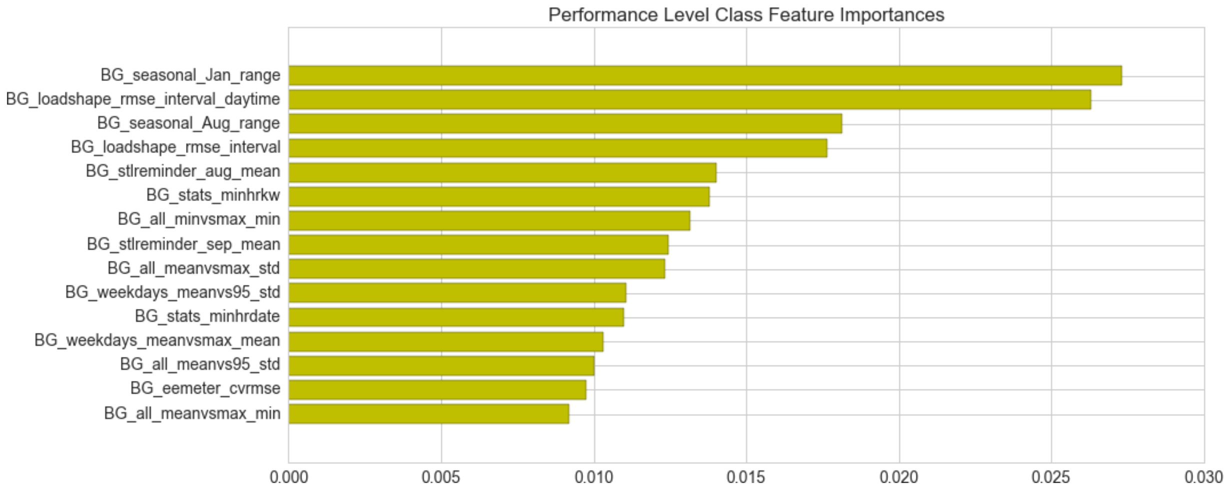

Figure 7.11: Importance of features according to random forest model in prediction of building performance class

linear classification performance. Eight clusters of features are detected on discriminating between high and low consumption. The first set of correlated features seen in the upper left corner of the figure contains a mix of statistical and daily pattern-based features. The second cluster includes a set of four features related to daily ratios. The third and largest group is mostly statistical and daily ratio-based features. The fourth, sixth, seventh, and eighth clusters all contain mostly in-class similarity and temporal features created using *jmotif*. These features are an indicator of how well a building's patterns fit within its own class. An interesting aspect of these features is their lack of correlation with the rest of the larger set. This situation indicates that they are capturing unique behavior, not picked up by others in the set. These clusters are also relatively small with only one to four members. The sixth cluster contains a set of features that are mostly generated by the *stl* decomposition models.

Figure 7.13 shows the probability distributions of the top five performing features in predicting high versus low consumption. The number one top feature for differentiating between these classes is the daily in-class similarity feature that is generated by the *jmotif* process. This feature informs us that buildings from all classes that have the highest average daily pattern similarity to their peers are often also amongst the highest consuming buildings in their class. Buildings that are on average less similar in their daily patterns to their class are often in a lower percentile of consumption. This fact suggests that many buildings that are misclassified are lower consumers of electricity. The second and fourth features are daily statistical ratios. Buildings with higher consumption tend to have more *flat* profiles, likely due to a higher base load during unoccupied periods. The third top

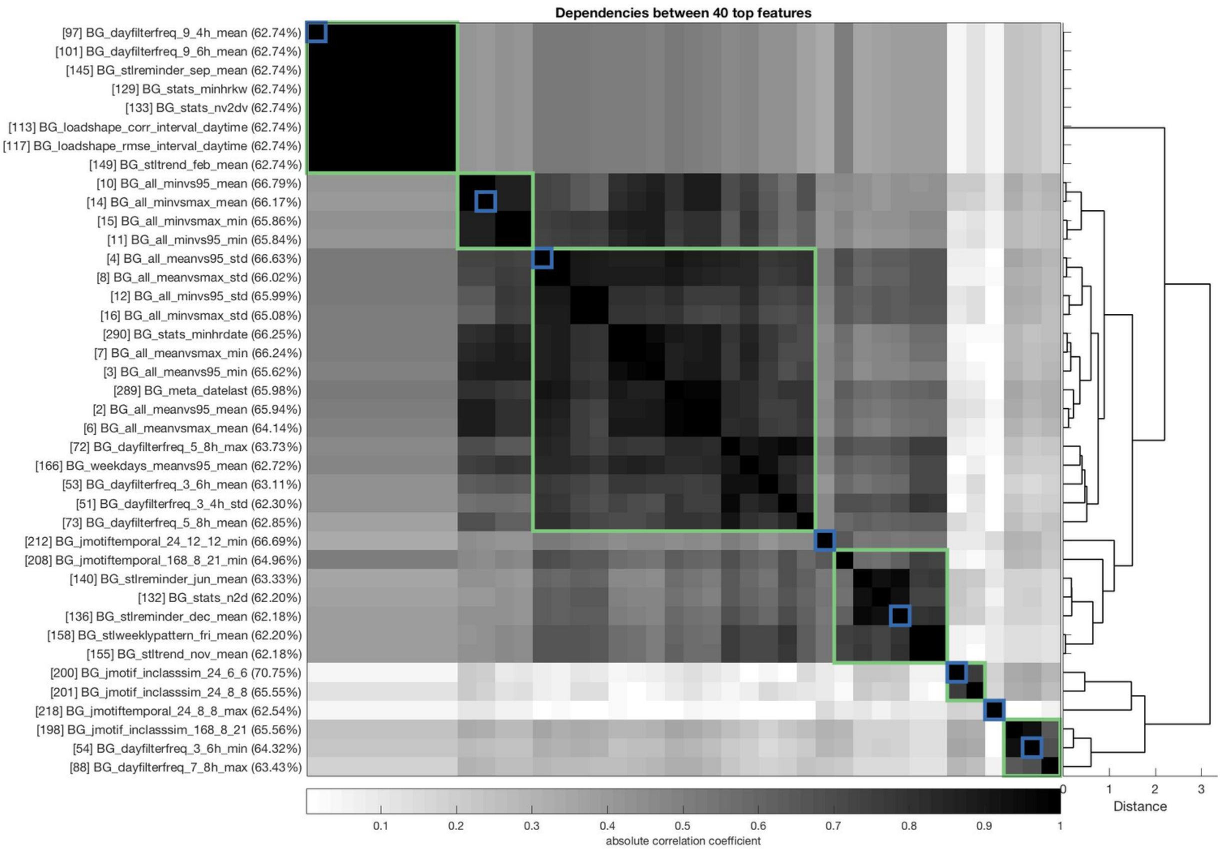

Figure 7.12: Clustering of dominant features in the comparison of high and low consumption performance classes

classifier is also created using the *jmotif* library and it suggests that a building that whose minimum daily pattern specificity across the year is an indicator of higher than average consumption.

Figure 7.14 shows the probability distribution of the features in their ability to distinguish between high and low consumption as compared to a baseline. The mean of the created features is approximately 58%, while the null mean is 51%. This situation indicates that the generated temporal features have a significant impact on the prediction and evaluation of whether a building performs well or not.

7.2.2 Discussion with Campus Case Study Subjects

In a situation similar to the discussion about building use type, participants in the case studies were guided through the process of analysis using a subset of features from buildings

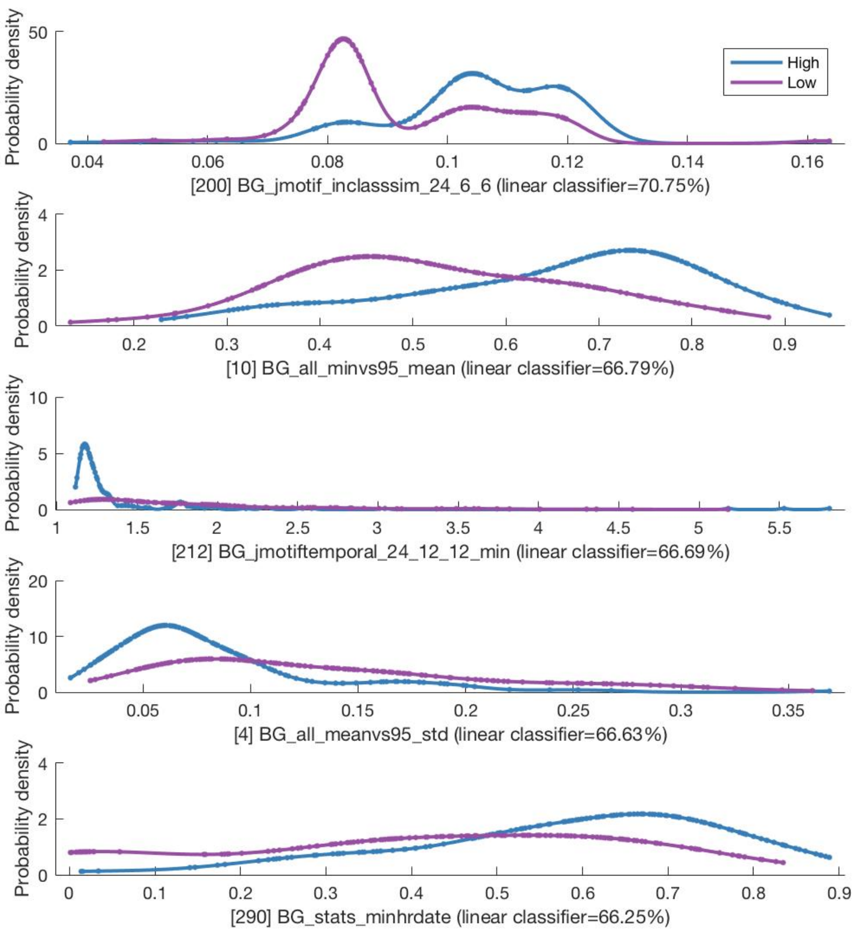

Figure 7.13: Probability density distribution of top five features in characterizing the difference between high and low consumption

on their campus. Figure 7.15 illustrates a graphic that was shown to the groups. In this case, the buildings are divided into two classes: *Good* and *Bad*. These categories are based on whether the building falls in the upper or lower 50% within its class. The first observation by the case study participants is that the load diversity, or the daily maximum versus minimum, is a strong indicator of the performance class. This fact is not surprising as this metric indicates the magnitude of the base load consumption as compared to the peak. Other relatively strong differentiators, in this case, are cooling energy, seasonal changes, and weekly specificity. The discussions related to this graphic centered around the potential for the temporal features to inform *why* a building is performing well or not. The results of Section 7.2.1 also include such clues on why a building may be in a high or low performing state.

Figure 7.16 illustrates another graphic related to building consumption classes that were

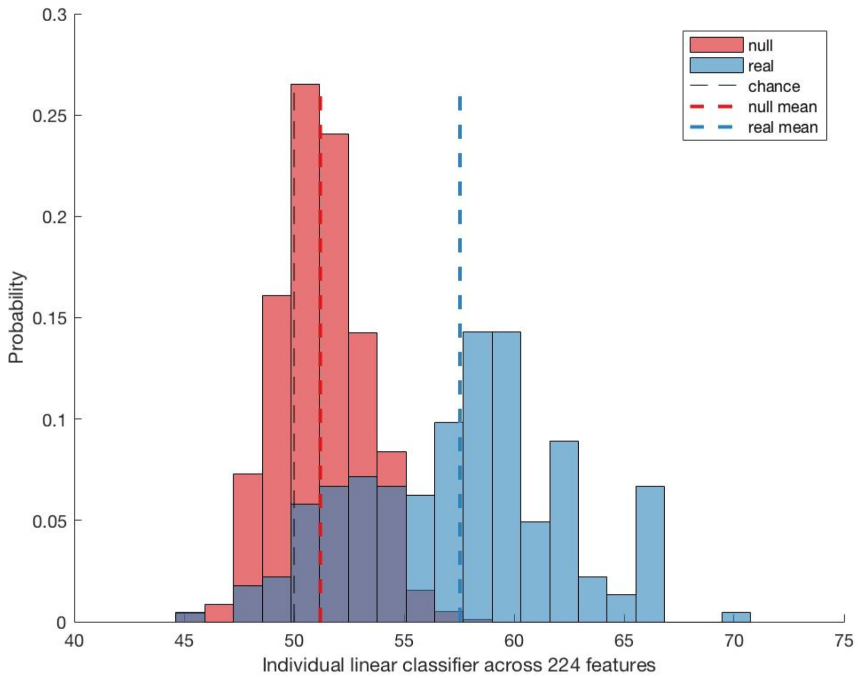

Figure 7.14: Ability of temporal features to distinguish between high and low consumers as compared to the null hypothesis

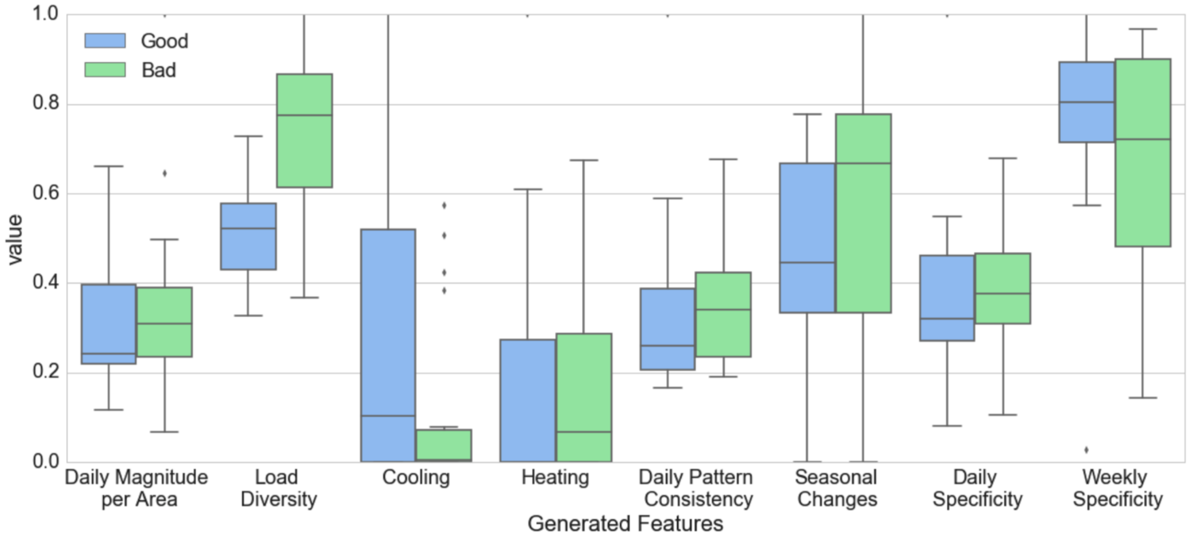

Figure 7.15: Simplified breakdowns of general features according to performance level that were presented to case study subjects

discussed with case study participants. This graphic is an overview of the distributions of the simplified set of features for a certain campus as compared to the entire set of case study buildings. This graphic shows where the buildings on this campus stand as compared to their peers. In this case, the buildings are on the higher end of the normalized consumption, which could likely be because they're also almost all in the highest 20% of buildings for heating energy consumption. The buildings also have a relatively high load diversity, thus the base loads for this campus are likely higher than average and interventions could be designed to reduce this unoccupied load. Many of the case study participants saw this insight as useful as it *supplements* the information from benchmarking.

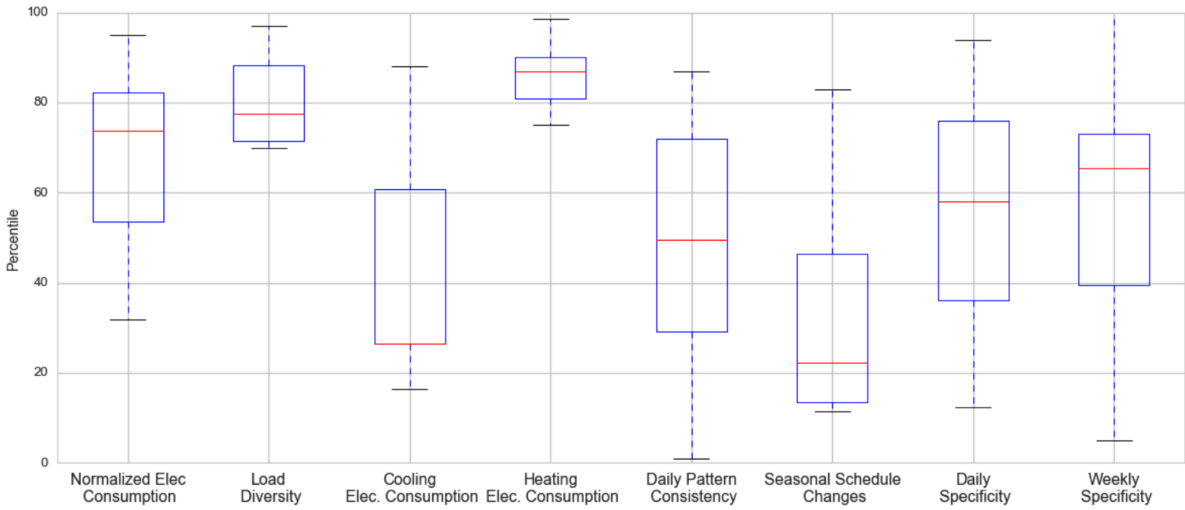

Figure 7.16: Feature distributions of a single campus as compare to all other case study buildings

7.3 Characterization of Operations Strategies

The final characterization objective for the case studies is the ability for the temporal features to classify buildings from the same campus, and thus buildings that are being operated in similar ways. This characterization takes to into account the similarity in occupancy schedules, patterns of use, and other factors related to how a building operates. Like the performance classes, this type of classification is more important in understanding the features that contribute to the differentiation, rather than the classification itself. Seven campuses were selected from the 507 buildings to create seven *groups* of buildings to characterize the difference between their operating behavior. Features were removed for this objective that are indicators of weather sensitivity as these would be related to the location of the buildings, and thus, the campus that they're located. Figure 7.17

illustrates the results from the random forest model trained on these data. The accuracy of this model is 80.5% as compared to a baseline of 16.9%. The model is excellent at predicting some of the groups, such as groups 1-4, which more deficient in others, such as 5-7. The high accuracy of this prediction is surprising and lends itself to the ability of the temporal features and the random forest model to predict the operational normalities of these buildings.

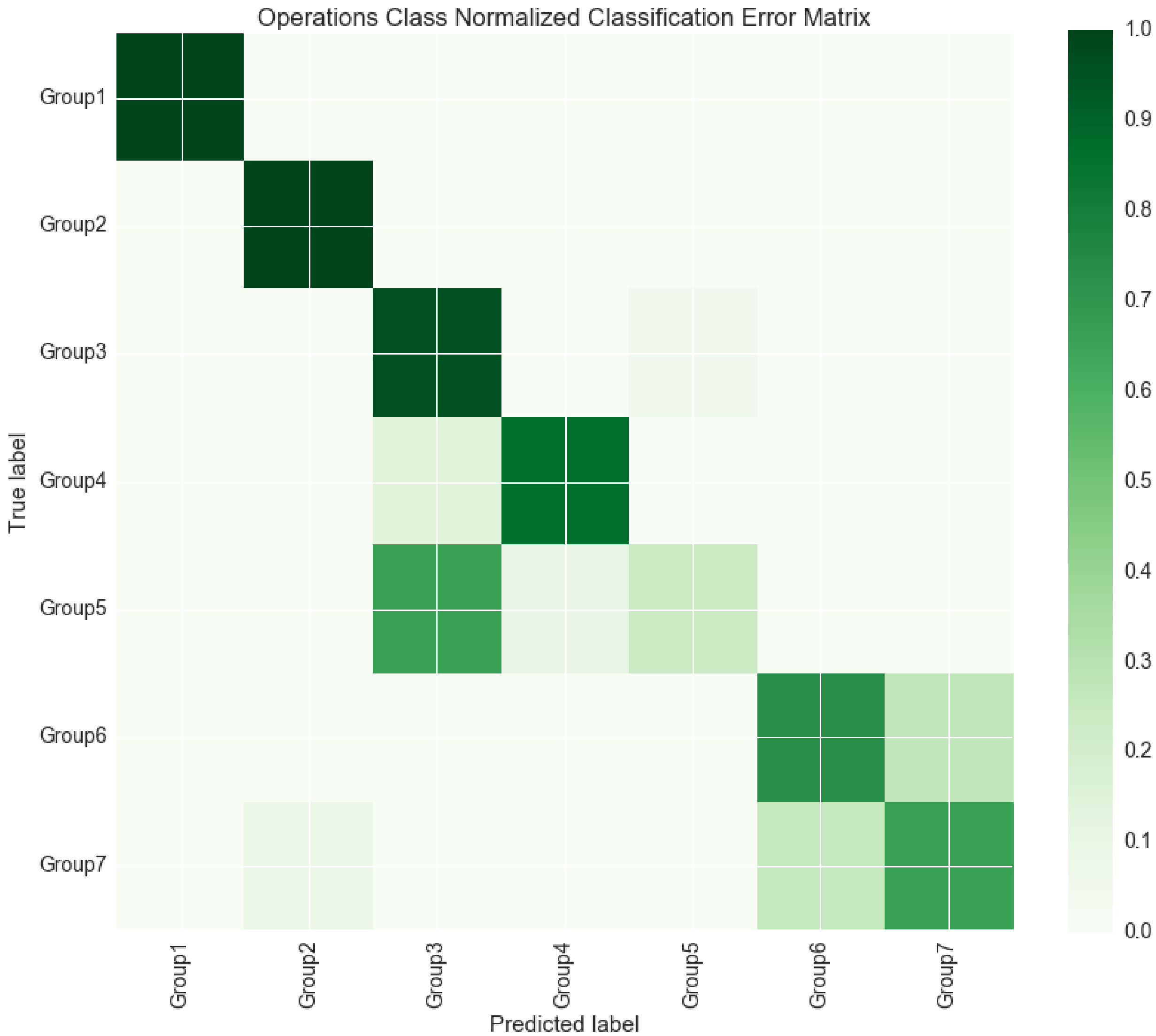

Figure 7.17: Classification error matrix for prediction of operations group type using a random forest model

Figure 7.18 illustrates the temporal features identified by the random forest model as the most important in class differentiation. One can observe several daily pattern-based features in addition to statistical and daily ratio-based features. This finding lends weight to the assumption that similarity in daily scheduling is a key discriminator between the operations of various campuses.

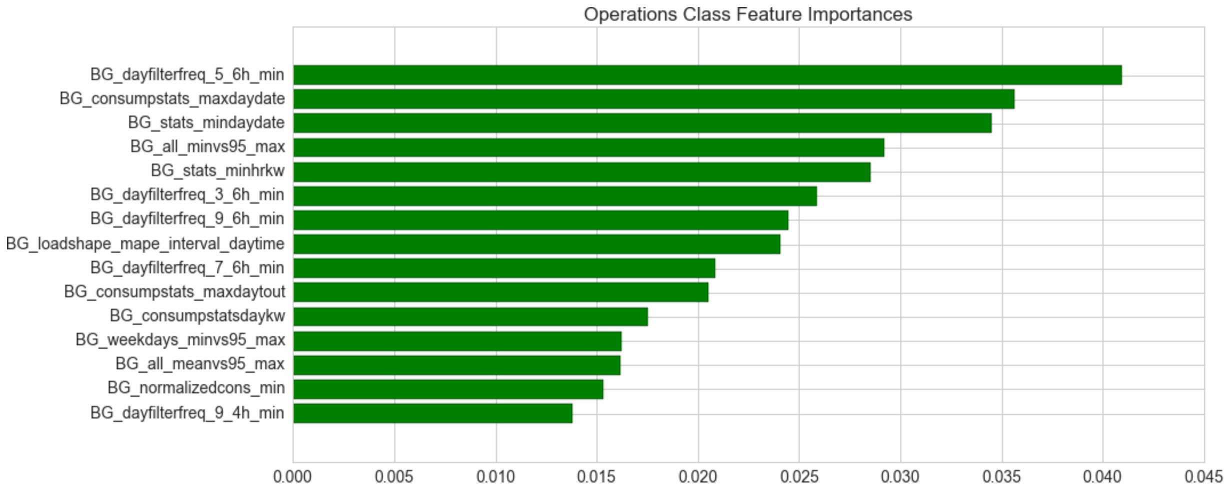

Figure 7.18: Importance of features in prediction of operations type

7.3.1 Group 1 versus Group 2 Comparison

Groups 1 and 2 were selected to undertake a deeper analysis using the highly comparative time-series analysis library. Figure 7.19 shows the top forty features and their correlated clusters. The first and largest cluster of features, in this case, are from the breakout detection process, a calculation of long-term volatility. This insight suggests that breakouts are a key discriminatory aspect of seasonal patterns that would exist for buildings being operated in the same way. The third cluster includes a diverse set of features including a few from the loadshape library and several statistics-based metrics. The fourth cluster contains features from the jmotif library, including both in-class similarity and specificity metrics. The remaining clusters are all quite small, only containing one or two features, and are made up of both pattern and motif-based features.

Figure 7.20 illustrates the top five features in the comparison of Group 1 and 2. The first three features are variations of in-class similarity. This indication shows that the buildings from these two particular groups are differentiated by how much the buildings fit within their designated class. The fourth and fifth dominant features are associated with the number of breakouts and long-term volatility.

Figure 7.21 illustrates how well all of the features can discriminate the difference between these two groups of buildings. The separation for a majority of the features is not much

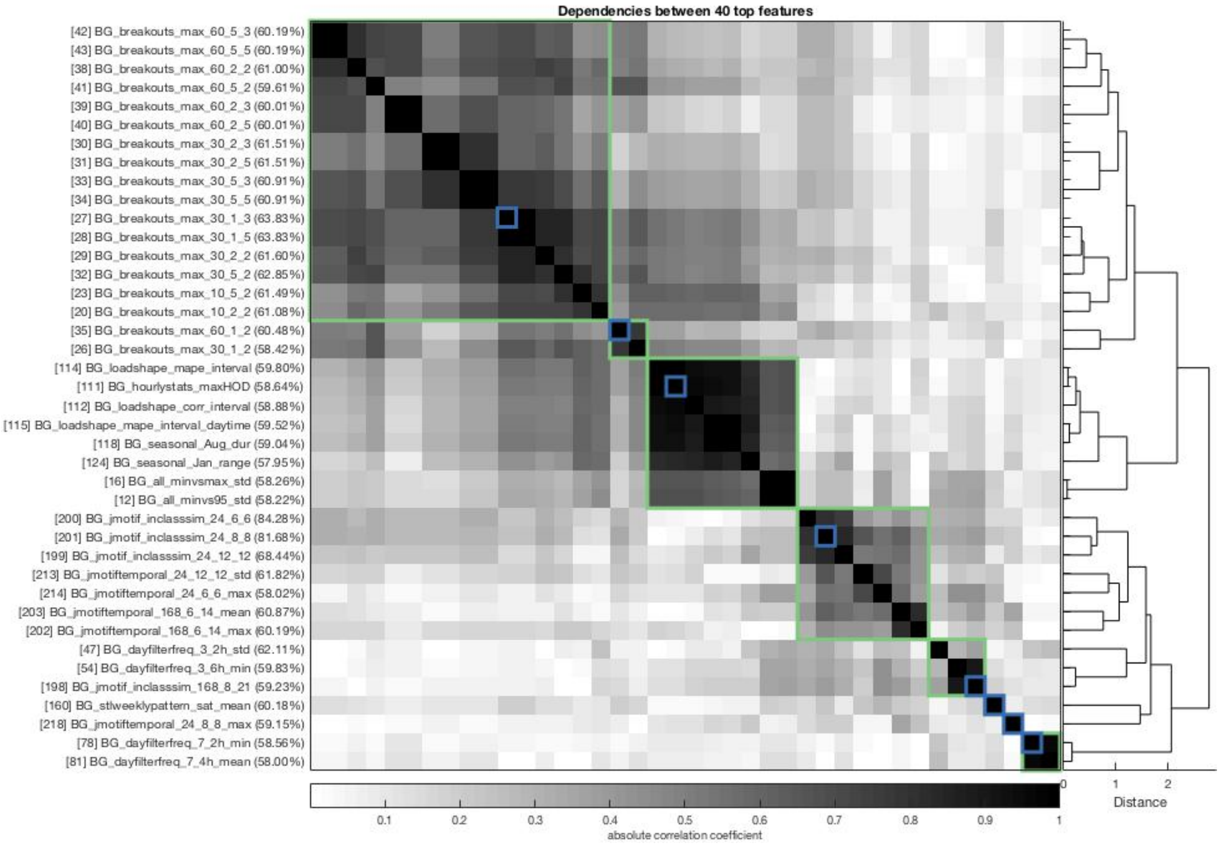

Figure 7.19: Clustering of dominant features in the comparison of operations group 1 and 2

1752 greater than the null mean, but the top differentiators are quite prominent.

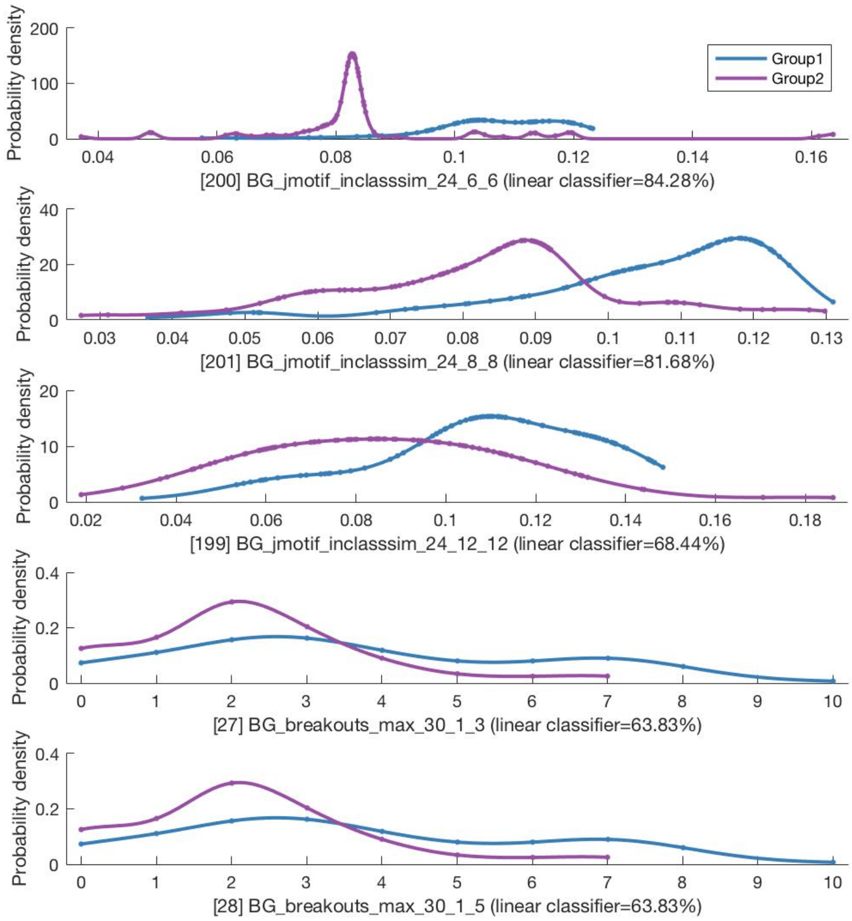

Figure 7.20: Probability density distribution of top five features in characterizing the difference between Group 1 and 2 operations classes

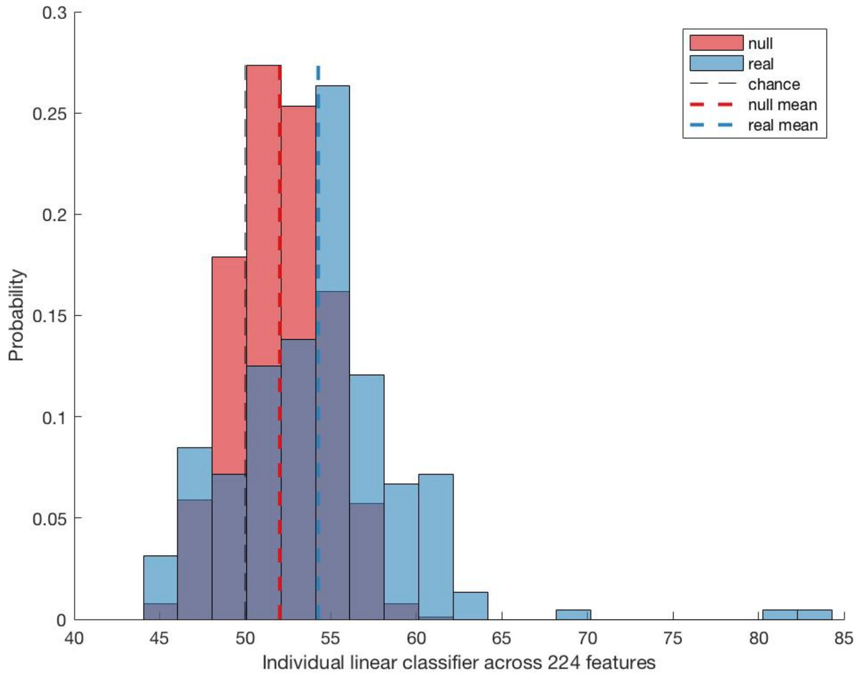

Figure 7.21: Ability of temporal features to distinguish between group 1 and 2 operations types as compared to the null hypothesis

8 Characterization of Energy-Savings Measure Implementation Success

In the previous sections, the process of temporal feature extraction and interpretation is implemented on a test set of 507 buildings. One of the key pieces of feedback from the case study interviews was that conventional analysis and meta-data collection for a set of buildings at this level is reasonable if the resources are allocated. This assumption quickly becomes untenable when discussing the analysis of the millions of buildings with smart meter data. These data are also known as Advanced Metering Infrastucture (AMI) data. In this section, execution of a subset of the temporal feature extraction process is applied to a data set of close to 10,000 buildings that have been aggregated by a third-party organization on behalf of electrical utilities. The utilization goal of these data is to supplement a process of targeting buildings for energy savings implementation measures. Utilization of temporal features is discussed in the context of assisting to label the approximate building use type and predicting measure success implementation through a combination of smart meter data and past project experience meta-data. These objectives are common in situations with large amounts of AMI data as often the only meta-data available for these buildings is related to the location and demographic characteristics of a building.

8.1 Predicting General Industry Membership

The first task that the features are used for is to characterize the general industry for which a building is being used. This task is a first step in using temporal features to predict necessary conventional features that can be used for more conventional targeting processes. As a proof-of-concept about this task, temporal data is used to build a classification model to predict the most common meta-data attribute of a building: its general use type. In this case, the label for use type is the Standard Industrial Class (SIC) one digit classification is used. The breakdown of the number of buildings within each of the SIC code categories is found in Figure 8.1.

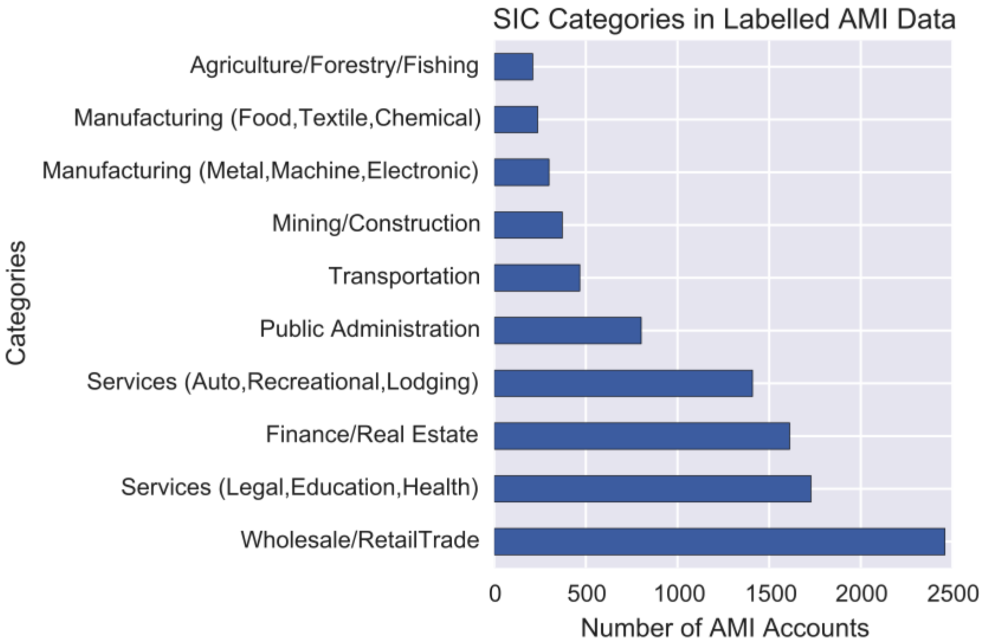

Figure 8.1: Building Type Classification of the Labeled AMI Accounts

Four classification models are then created to predict the general SIC Category of each account:

- Baseline model - using the distributions of the input samples to guess the category
- Non-Temporal Features Model – using the old non-temporal features containing monthly data and zip code/location information
- Temporal Features – using the new features generated from the AMI data
- Combined Features – using all the features, temporal and non-temporal

Once again a random forest model was implemented using Python's Scikit-Learn library. The models were executed an out-of-bag error to calculate mean model accuracy of a multi-label classification. Figure 8.2 illustrates the results of the models with respect to percent mean accuracy improvement over the baseline.

The baseline model correctly predicts the labels with a 18.1% accuracy, while the features influenced models were 38.5% for Non-Temporal, 45.3% for the Temporal and 45.7% for the combined model. The baseline model represents common practice in which a class is chosen based on the probability distribution of that class occurring in the labeled dataset.

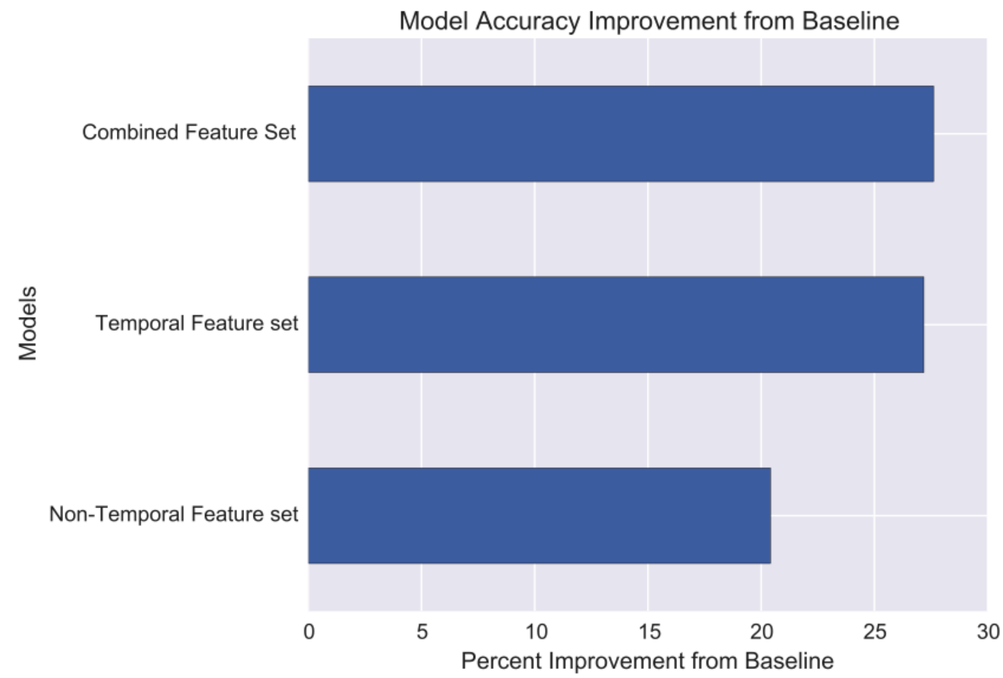

Figure 8.2: Mean Model Accuracy Improvement from Baseline

The combined feature set more than doubles the probability of predicting this piece of meta-data.

Mean accuracy of multi-label classification models as done in this analysis is a harsh metric as it forces the model to make a single choice for labeling each sample. In practice, it is not desired for a model that completely makes this decision; but instead to simply want the model to inform what the probability that a sample fits within a class. For example, there could be 45% chance an unlabeled account is an office, a 35% chance it is a school and 20% chance it is a grocery store. The reason to choose mean model accuracy in this report it to communicate a simplified message of the techniques and the progress made thus far. The fact that the overall classification model accuracy is around 40-60% for a classification model with ten classes is not discouraging. It is the improvement in mean accuracy from baselines that is the focus and this has been demonstrated so far in the project.

It can also be seen in detail how the model predicts the classes for each by creating and analyzing a classification confusion matrix. Figure 8.3 illustrates this matrix for the combined model. It is observed that two of the largest classes, Retail and Finance, have the highest accuracy rates at over 55-60% with several other categories being misclassified within them. This issue is common with imbalanced classification models and further feature development would improve the model by better characterizing the difference between each class.

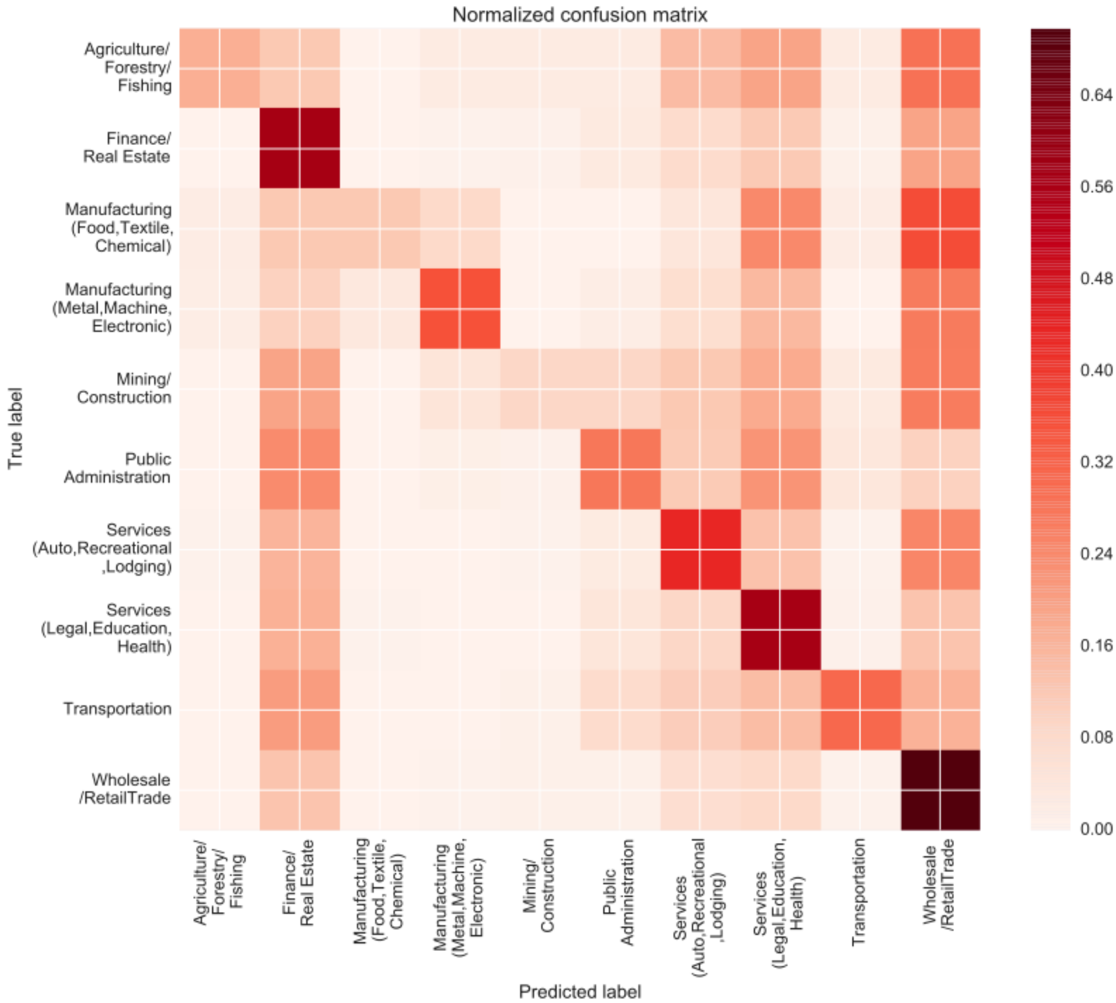

Figure 8.3: Classification error matrix for prediction of standard industry class (SIC) using a random forest model

8.2 Energy Efficiency Measure Implementation Success Prediction

The next example of using the temporal features is predicting the success of future measure implementation events using the past data. For this proof-of-concept, Pre and Post-measure implementation data are utilized from close to 1,600 buildings that had one or more measures implemented. The difference in mean daily consumption before an after the measure implementation is calculated to achieve a rough indication of measure success. The measures into three classifications is divided according to where the difference in daily consumption for each account fits in the range of values. In this analysis, the accounts in the lowest 33% were considered "Poor", while the 66% percentile were "Average" and the top 33% are considered "Good". Simple difference in mean daily consumption is not a perfect metric for success, as it is not normalized for weather or occupancy changes; although it is adequate for this step as we are already arbitrarily choosing the thresholds for class difference anyway and are looking for a simple metric at this point.

Figure 8.4 illustrates a breakdown of the measure categories within the tested dataset.

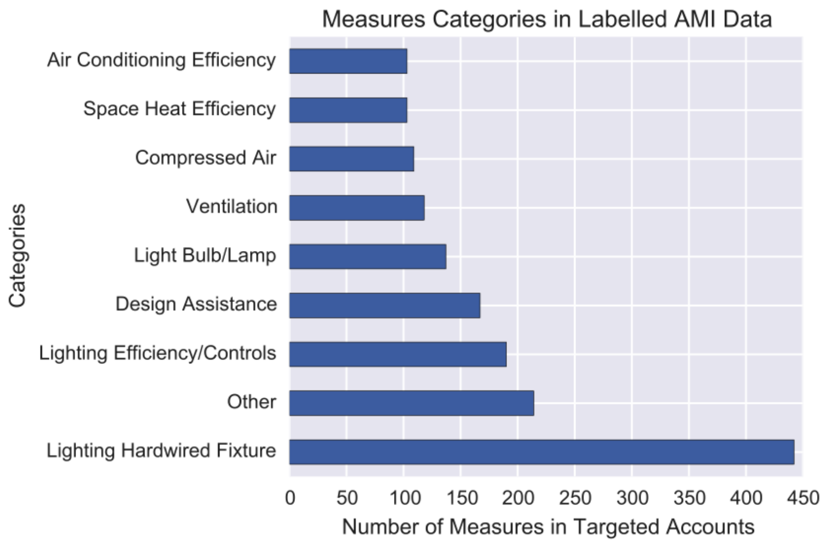

Figure 8.4: Breakdown of Measure Categories included in the Dataset

A Random Forest algorithm was implemented to use the temporal features to predict the class of potential measure success (Good, Average or Poor). Figure 8.5 illustrates the classification error matrix for this model.

The baseline model with this data is able to predict the success within this set of classification at 32.8% accuracy, while the model based on temporal features achieved 51.1% accuracy. The more important aspect to pay attention to is that the misclassification rate between "Good" and "Poor" is less than 20% – a promising fact that motivates further investigation using the existing temporal data-set.

Figure 8.5: Classification error matrix for prediction of measure implementation success using a random forest model

8.3 Discussion

This section discusses the creation of additional information about smart meter by extracting characteristics from the high-frequency time-series measurements. Based on a classification test using almost 9,600 labeled smart meter accounts, the accuracy of predicting building type is improved (based on SIC 1-Digit category) by over 27% over a conventional baseline.

Data about energy efficiency measures implementation and classified almost 1,600 accounts was aggregated into Good, Average, and Poor performing classes according to pre and post-measure consumption. A classification model is developed that improves the ability to predict measure implementation class success by 18% over a baseline. Additionally, there was only a 20% error rate in differentiating between Good and Poor performing measures.

The biggest opportunity ahead is to apply to unlabeled smart meter accounts to help characterize missing meta-data and predict measure implementation success for future projects. Much work is also yet to be done to improve the models and input information to bring the overall prediction accuracies higher in absolute terms. Model prediction can also be improved incrementally as the AMI and measures implementation data are better integrated.

9 Conclusion and Outlook

This dissertation was undertaken with objectives related to the characterization of building behavior using temporal feature extraction and variable importance screening. The primary goal of the effort is to automate the process of predicting various types of meta-data. This process was implemented on two sets of case study buildings and the key quantitative conclusions include:

- The framework can characterize primary building use type with a general accuracy of 67.8% as compared to a baseline model of 22.2% based on five use type classes. Temporal features enable a three-fold increase in building use prediction. Pattern-based features are the most common category in the top ten in the characterization of use-type, thus are important differentiators as compared to more traditional features.
- Building performance class overall accuracy of the model for classification is 62.3% as compared to a baseline of 38%. The top indicator of high versus low building in-class performance was temporal features pattern specificity. Once again, pattern-based temporal features were found to be significant in distinguishing between different types of behavior.
- For operations class, the accuracy of this model is 80.5% as compared to a baseline of 16.9%, a four-fold increase in accuracy. Daily scheduling of buildings was captured using the *DayFilter* features, accounting for half of the entire input features.
- The ability to assist in the targeting of buildings based on how well they respond to energy savings measures is enhanced significantly using this process. An experiment was conducted in which prediction of whether a building fits within three classes of energy savings success. In the baseline model, there was only an 18.1% accuracy in predicting whether a building will be good or bad with regards to an energy-saving measure implementation. The temporal features developed and implemented were able to predict a 45.3% accuracy of prediction, more than double the performance.

It should be noted that the quantitative analysis portion of this study seeks to illustrate the accuracy of characterization. This success metric is as compared to the quantity of energy saved, the percentage of savings due to implementation, and other building performance metrics. This shift in focus is deliberate as the framework is designed as a step between raw data and other techniques that target the decision-making process.

Several insights were gathered from the qualitative research approaches on the case study interviews. This insight can be found in Section 7. The first key issue was that the two-step framework was seen as *interesting and insightful* regarding the results. Participants were generally engaged with the content and results, but little concrete decision-making power was extracted from them. Guidelines for further work in the utilization of the framework for practical applications was discussed.

9.1 Outlook

A major future effort to build upon this work is expansion and enhancement of both the building data library and the applied techniques. The more meta-data collected for each building, the more detailed understanding of what temporal behavior is correlated with those data. Thus, a more detailed characterization of each building and correlations between the meta-data can occur. Additionally, increasing the number and scope of the buildings in the data set enhances the ability to generalize the results across the wider building stock. This repository could grow into something of a *Building Data Genome* that enables researchers to download, make generalizations and infer information from the data set in addition to comparing it to buildings from their portfolios. This idea draws inspiration from the field of bioinformatics and the study of genomes in the biological world. These genomes were sequenced from raw data (DNA) and are used to find patterns or correlations related to certain meta-data about a specific organism. The release of the data and code generated to create this framework is announced in 9.2.

The first major area of influence that the framework outlined in this dissertation is within the domain of building performance benchmarking. This focus was discussed in Section 7.1 in the ability for the framework to predict what the primary use type of a building based on its temporal data. With the increased availability of high-frequency data, soon building owners will have the ability to submit their fifteen-minute frequency performance data directly from their utility or energy management systems. Extracting information about how well each building performs as compared to its peers can be enhanced through the use of this high-frequency data. This dissertation has illustrated the use of temporal features for the purpose of building use and performance class prediction; both concepts that are very relevant to this application. The next steps in this effort include fine-tuning the algorithms such that meta-data about a potential input building is checked against the temporal features generated from the raw data.

Another promising field of research is in the automated targeting of buildings amongst vast portfolios for various objectives such as retrofit opportunities. This field is emerging as large numbers of AMI data sets become available. As discussed in the introduction, there is an under-supply of qualified data analytics experts to extract patterns and information from these data to make decisions on which buildings to prioritize on various objectives. The framework outlined identifies an initial step in the direction of characterizing energy savings measures. Further work is necessary to develop these models into a tool that automatically determines the applicability of various energy savings measures based on temporal data from past projects and training data from potential targeted buildings. These types of tools could act as screening process in how well a building fits within the category its being benchmarked against. This process could also provide feedback as to *why* a building did or didn't perform well within its class based on

The effort in this dissertation also works to reduce the ambiguity of algorithm applicability in commercial building research. This phenomenon is observed in the wider data mining community as a whole (Keogh & Kasetty 2003). In this study, Keogh et al. describe a scenario in which "Literally hundreds of papers have introduced new algorithms to index, classify, cluster, and segment time series." They go on to state, "Much of this work has very little utility because the contribution made (speed in the case of indexing, accuracy in the case of classification and clustering, model accuracy in the case of segmentation) offer an amount of improvement that would have been completely dwarfed by the variance that would have been observed by testing on many real world datasets, or the variance that would have been observed by changing minor (unstated) implementation details." They make the case that time series benchmarking data sets should be used to evaluate whether a new proposed algorithm is more beneficial as compared to previous work. The use of benchmark data sets reduces the impact of implementation bias, the disparity in the quality of implementation of a proposed approach versus its competitors, and data bias, the use of a particular set of testing data to confirm the desired finding. These biases were proven common amongst popular data mining publications, and it is suspected that they may be prevalent in the papers in this review. Benchmarking data sets for building performance analysis could be developed and promoted for use in papers similar to what was used in the *Great Building Energy Predictor Shootout* competition that was held in the mid-1990's (**?**). In this competition, standardized training and testing data sets were provided to numerous participants to determine who could create the most accurate model to predict future consumption. A modern-day *energy predictor shootout* could be held to incorporate the numerous advances made in machine learning since then. In addition to the ability to compare accuracy of algorithms, publications should also include more detailed explanations of the effort required to implement the proposed techniques such that a third-party could evaluate whether the effort-to-accuracy balance is right for their application.

Regarding outlook, the techniques outlined in this study are also applicable to other domains with temporal data and daily, weekly and seasonal patterns from fields such as transportation or finance. For example, finding the specificity or long-term volatility of the driving habits of cars on the road may also

9.2 Reproducible Research Outputs

A primary goal of this dissertation was the creation of a repository of building performance data and techniques that can be implemented by other researchers and professionals. The 507 building case study data set and much of the data analysis behind the temporal feature extraction and classification has been combined into a GitHub repository that is open and accessible online (https://github.com/architecture-building-systems/the-building-data-genome). The release of specific data sets for data science publications could become the norm, thus facilitating the ability for a third-party to recreate the results. The repository includes a set of Jupyter notebooks that can be downloaded and used to replicate the results of those studies easily. The Jupyter notebook website states that it is "an open source, web application-based document that combines live code, equations, visualizations, and explanatory text."[1] The use of these types of formats is an opportunity to enhance the interdisciplinary communication further through the sharing and utilization of publication data.

[1] https://jupyter.org/

A Complete List of Generated Temporal Features

This appendix section provides a list of the library of temporal features developed or utilized in this dissertation. The last three columns indicate whether the feature was used as an input in each of the sections of Chapter 7: Use Type (U), Consumption Type (C), and Operations Type (O).

Feature Code	Description	Category	Type	U	C	O
consumpstats dailykwminvar	Daily minimum variance	Stats.	Cons. Stats	X	X	
consumpstats dailykwvar	Daily variance	Stats.	Cons. Stats	X	X	
consumpstats kw90	Ninety percentile	Stats.	Cons. Stats	X	X	
consumpstats kwmean	Mean	Stats.	Cons. Stats	X	X	
consumpstats kwmeanannual	Annual mean	Stats.	Cons. Stats	X	X	
consumpstats kwmeansummer	Annual summer	Stats.	Cons. Stats	X	X	
consumpstats kwmeanwinter	Annual winter	Stats.	Cons. Stats	X	X	
consumpstats kwtotal	Total	Stats.	Cons. Stats	X	X	
consumpstats kwvar	Variance	Stats.	Cons. Stats	X	X	
consumpstats max	Max	Stats.	Cons. Stats	X	X	
consumpstats max97	Max percentile	Stats.	Cons. Stats	X	X	
consumpstats maxMA	Max MA	Stats.	Cons. Stats	X	X	
consumpstats maxdaydate	Day of max use	Stats.	Cons. Stats	X	X	
consumpstats maxdaypct	Day of max as a pct.	Stats.	Cons. Stats	X	X	
consumpstats maxdaytout	Day of max output	Stats.	Cons. Stats	X	X	
consumpstats maxhrkw	Max hour	Stats.	Cons. Stats	X	X	
consumpstats maxhrtout	Outdoor air temp on max day	Stats.	Cons. Stats	X	X	
consumpstats mean	Mean	Stats.	Cons. Stats	X	X	
consumpstats min	Minimum	Stats.	Cons. Stats	X	X	
consumpstats min3	Minimum percentile	Stats.	Cons. Stats	X	X	
consumpstats range	Range	Stats.	Cons. Stats	X	X	
consumpstats t10kw	Most common hour in top ten percent	Stats.	Cons. Stats	X	X	
consumpstatsdaykw	Total on max day	Stats.	Cons. Stats	X	X	
consumpstatsdaytout	Outdoor air temp	Stats.	Cons. Stats	X	X	
consumpstatsmaxdaykw	Day with max cons.	Stats.	Cons. Stats	X	X	
consumpstatst 90kw	Max percentile	Stats.	Cons. Stats	X	X	
normalizedcons max	Area normalized stats	Stats.	Cons. Stats	X	X	
normalizedcons mean	Area normalized stats	Stats.	Cons. Stats	X	X	
normalizedcons min	Area normalized stats	Stats.	Cons. Stats	X	X	
normalizedcons std	Area normalized stats	Stats.	Cons. Stats	X	X	
consumpstats Aug max	Aug stats	Stats.	Cons. Stats	X		

consumpstats Aug mean	Aug stats	Stats.	Cons. Stats	X		
consumpstats Aug min	Aug stats	Stats.	Cons. Stats	X		
consumpstats Aug mn2mx	Aug stats	Stats.	Cons. Stats	X		
consumpstats Jan max	Jan stats	Stats.	Cons. Stats	X		
consumpstats Jan mean	Jan stats	Stats.	Cons. Stats	X		
consumpstats Jan min	Jan stats	Stats.	Cons. Stats	X		
consumpstats dailykwmaxvar	Daily max variance	Stats.	Cons. Stats	X		
consumpstats kwtotalApr	Monthly totals	Stats.	Cons. Stats	X		
consumpstats kwtotalAug	Monthly totals	Stats.	Cons. Stats	X		
consumpstats kwtotalDec	Monthly totals	Stats.	Cons. Stats	X		
consumpstats kwtotalFeb	Monthly totals	Stats.	Cons. Stats	X		
consumpstats kwtotalJan	Monthly totals	Stats.	Cons. Stats	X		
consumpstats kwtotalJul	Monthly totals	Stats.	Cons. Stats	X		
consumpstats kwtotalJun	Monthly totals	Stats.	Cons. Stats	X		
consumpstats kwtotalMar	Monthly totals	Stats.	Cons. Stats	X		
consumpstats kwtotalMay	Monthly totals	Stats.	Cons. Stats	X		
consumpstats kwtotalNov	Monthly totals	Stats.	Cons. Stats	X		
consumpstats kwtotalOct	Monthly totals	Stats.	Cons. Stats	X		
consumpstats kwtotalSep	Monthly totals	Stats.	Cons. Stats	X		
consumpstats kwvarsummer	Summer variance	Stats.	Cons. Stats	X		
consumpstats kwvarwinter	Winter variance	Stats.	Cons. Stats	X		
consumpstats maxhrdate	Timestamp of max cons.	Stats.	Cons. Stats	X		
consumpstats t10t	Temp at percentile	Stats.	Cons. Stats	X		
consumpstats t90t	Temp at percentil	Stats.	Cons. Stats	X		
all meanvs95 max	Ratio of daily	Stats.	Daily Ratios	X	X	X
all meanvs95 mean	Ratio of daily	Stats.	Daily Ratios	X	X	X
all meanvs95 min	Ratio of daily	Stats.	Daily Ratios	X	X	X
all meanvs95 std	Ratio of daily	Stats.	Daily Ratios	X	X	X
all meanvsmax max	Ratio of daily	Stats.	Daily Ratios	X	X	X
all meanvsmax mean	Ratio of daily	Stats.	Daily Ratios	X	X	X
all meanvsmax min	Ratio of daily	Stats.	Daily Ratios	X	X	X
all meanvsmax std	Ratio of daily	Stats.	Daily Ratios	X	X	X
all minvs95 max	Ratio of daily	Stats.	Daily Ratios	X	X	X
all minvs95 mean	Ratio of daily	Stats.	Daily Ratios	X	X	X
all minvs95 min	Ratio of daily	Stats.	Daily Ratios	X	X	X
all minvs95 std	Ratio of daily	Stats.	Daily Ratios	X	X	X
all minvsmax max	Ratio of daily	Stats.	Daily Ratios	X	X	X
all minvsmax mean	Ratio of daily	Stats.	Daily Ratios	X	X	X
all minvsmax min	Ratio of daily	Stats.	Daily Ratios	X	X	X
all minvsmax std	Ratio of daily	Stats.	Daily Ratios	X	X	X
weekdays meanvs95 max	Ratio of weekday	Stats.	Daily Ratios	X	X	X
weekdays meanvs95 mean	Ratio of weekday	Stats.	Daily Ratios	X	X	X
weekdays meanvs95 min	Ratio of weekday	Stats.	Daily Ratios	X	X	X
weekdays meanvs95 std	Ratio of weekday	Stats.	Daily Ratios	X	X	X
weekdays meanvsmax max	Ratio of weekday	Stats.	Daily Ratios	X	X	X
weekdays meanvsmax mean	Ratio of weekday	Stats.	Daily Ratios	X	X	X
weekdays meanvsmax min	Ratio of weekday	Stats.	Daily Ratios	X	X	X
weekdays meanvsmax std	Ratio of weekday	Stats.	Daily Ratios	X	X	X
weekdays minvs95 max	Ratio of weekday	Stats.	Daily Ratios	X	X	X
weekdays minvs95 mean	Ratio of weekday	Stats.	Daily Ratios	X	X	X
weekdays minvs95 min	Ratio of weekday	Stats.	Daily Ratios	X	X	X
weekdays minvs95 std	Ratio of weekday	Stats.	Daily Ratios	X	X	X
weekdays minvsmax max	Ratio of weekday	Stats.	Daily Ratios	X	X	X
weekdays minvsmax mean	Ratio of weekday	Stats.	Daily Ratios	X	X	X

weekdays minvsmax min	Ratio of weekday	Stats.	Daily Ratios	X	X	X
weekdays minvsmax std	Ratio of weekday	Stats.	Daily Ratios	X	X	X
weekend meanvs95 max	Ratio of weekend	Stats.	Daily Ratios	X	X	X
weekend meanvs95 mean	Ratio of weekend	Stats.	Daily Ratios	X	X	X
weekend meanvs95 min	Ratio of weekend	Stats.	Daily Ratios	X	X	X
weekend meanvs95 std	Ratio of weekend	Stats.	Daily Ratios	X	X	X
weekend meanvsmax max	Ratio of weekend	Stats.	Daily Ratios	X	X	X
weekend meanvsmax mean	Ratio of weekend	Stats.	Daily Ratios	X	X	X
weekend meanvsmax min	Ratio of weekend	Stats.	Daily Ratios	X	X	X
weekend meanvsmax std	Ratio of weekend	Stats.	Daily Ratios	X	X	X
weekend minvs95 max	Ratio of weekend	Stats.	Daily Ratios	X	X	X
weekend minvs95 mean	Ratio of weekend	Stats.	Daily Ratios	X	X	X
weekend minvs95 min	Ratio of weekend	Stats.	Daily Ratios	X	X	X
weekend minvs95 std	Ratio of weekend	Stats.	Daily Ratios	X	X	X
weekend minvsmax max	Ratio of weekend	Stats.	Daily Ratios	X	X	X
weekend minvsmax mean	Ratio of weekend	Stats.	Daily Ratios	X	X	X
weekend minvsmax min	Ratio of weekend	Stats.	Daily Ratios	X	X	X
weekend minvsmax std	Ratio of weekend	Stats.	Daily Ratios	X	X	X
hourlystats maxHOD	Hour of day stat	Stats.	Hourly Stats.	X	X	X
hourlystats HODmean1	Hour of day stat	Stats.	Hourly Stats.	X	X	
hourlystats HODmean10	Hour of day stat	Stats.	Hourly Stats.	X	X	
hourlystats HODmean11	Hour of day stat	Stats.	Hourly Stats.	X	X	
hourlystats HODmean12	Hour of day stat	Stats.	Hourly Stats.	X	X	
hourlystats HODmean13	Hour of day stat	Stats.	Hourly Stats.	X	X	
hourlystats HODmean14	Hour of day stat	Stats.	Hourly Stats.	X	X	
hourlystats HODmean15	Hour of day stat	Stats.	Hourly Stats.	X	X	
hourlystats HODmean16	Hour of day stat	Stats.	Hourly Stats.	X	X	
hourlystats HODmean17	Hour of day stat	Stats.	Hourly Stats.	X	X	
hourlystats HODmean18	Hour of day stat	Stats.	Hourly Stats.	X	X	
hourlystats HODmean19	Hour of day stat	Stats.	Hourly Stats.	X	X	
hourlystats HODmean2	Hour of day stat	Stats.	Hourly Stats.	X	X	
hourlystats HODmean20	Hour of day stat	Stats.	Hourly Stats.	X	X	
hourlystats HODmean21	Hour of day stat	Stats.	Hourly Stats.	X	X	
hourlystats HODmean22	Hour of day stat	Stats.	Hourly Stats.	X	X	
hourlystats HODmean23	Hour of day stat	Stats.	Hourly Stats.	X	X	
hourlystats HODmean24	Hour of day stat	Stats.	Hourly Stats.	X	X	
hourlystats HODmean3	Hour of day stat	Stats.	Hourly Stats.	X	X	
hourlystats HODmean4	Hour of day stat	Stats.	Hourly Stats.	X	X	
hourlystats HODmean5	Hour of day stat	Stats.	Hourly Stats.	X	X	
hourlystats HODmean6	Hour of day stat	Stats.	Hourly Stats.	X	X	
hourlystats HODmean7	Hour of day stat	Stats.	Hourly Stats.	X	X	
hourlystats HODmean8	Hour of day stat	Stats.	Hourly Stats.	X	X	
hourlystats HODmean9	Hour of day stat	Stats.	Hourly Stats.	X	X	
seasonal Aug dur	Seasonal stats	Stats.	Simple Stats.	X	X	X
seasonal Aug n2d	Seasonal stats	Stats.	Simple Stats.	X	X	X
seasonal Aug range	Seasonal stats	Stats.	Simple Stats.	X	X	X
seasonal Jan dur	Seasonal stats	Stats.	Simple Stats.	X	X	X
seasonal Jan mn2mx	Seasonal stats	Stats.	Simple Stats.	X	X	X
seasonal Jan n2d	Seasonal stats	Stats.	Simple Stats.	X	X	X
seasonal Jan range	Seasonal stats	Stats.	Simple Stats.	X	X	X
stats dur	Duration	Stats.	Simple Stats.	X	X	X
stats kwtoutcor	Temp and cons. Correlation	Stats.	Simple Stats.	X	X	X
stats mindaydate	Minimum cons day	Stats.	Simple Stats.	X	X	X
stats mindaypct	Min day percentage	Stats.	Simple Stats.	X	X	X

stats minhrkw	Min hour	Stats.	Simple Stats.	X	X	X
stats minhrtout	Temp at min. hour	Stats.	Simple Stats.	X	X	X

Feature Code	Description	Category	Type	U	C	
eemeter coolbalpt	Cooling balance point	Model	EEMeter Model	X	X	X
eemeter cvrmse	Model fit coefficient	Model	EEMeter Model	X	X	X
eemeter heatbalpt	Heating balance point	Model	EEMeter Model	X	X	X
eemeter baseload	Baseload	Model	EEMeter Model	X	X	
eemeter cooling max	Maximum cooling cons.	Model	EEMeter Model	X	X	
eemeter cooling mean	Mean cooling cons.	Model	EEMeter Model	X	X	
eemeter cooling min	Min cooling cons.	Model	EEMeter Model	X	X	
eemeter cooling std	Std. Dev. Cooling cons.	Model	EEMeter Model	X	X	
eemeter coolslope	Slope of cooling linear regression	Model	EEMeter Model	X	X	
eemeter heating max	Maximum heating cons.	Model	EEMeter Model	X	X	
eemeter heating mean	Mean heating cons.	Model	EEMeter Model	X	X	
eemeter heating min	Min. heating cons.	Model	EEMeter Model	X	X	
eemeter heating std	Std. Dev. Heaint cons.	Model	EEMeter Model	X	X	
eemeter heatslope	Slope of heatig linear regression	Model	EEMeter Model	X	X	
eemeter nmbe	Model fit coefficient	Model	EEMeter Model	X	X	
loadshape corr interval	Model fit coefficient	Model	Loadshape Model	X	X	X
loadshape corr interval daytime	Model fit coefficient	Model	Loadshape Model	X	X	X
loadshape mape interval	Model fit coefficient	Model	Loadshape Model	X	X	X
loadshape mape interval daytime	Model fit coefficient	Model	Loadshape Model	X	X	X
loadshape rmse interval	Model fit coefficient	Model	Loadshape Model	X	X	X
loadshape rmse interval daytime	Model fit coefficient	Model	Loadshape Model	X	X	X
stlreminder apr mean	Model fit remainder	Model	STL Model	X	X	X
stlreminder aug mean	Model fit remainder	Model	STL Model	X	X	X
stlreminder dec mean	Model fit remainder	Model	STL Model	X	X	X
stlreminder feb mean	Model fit remainder	Model	STL Model	X	X	X
stlreminder jan mean	Model fit remainder	Model	STL Model	X	X	X
stlreminder jul mean	Model fit remainder	Model	STL Model	X	X	X
stlreminder jun mean	Model fit remainder	Model	STL Model	X	X	X
stlreminder mar mean	Model fit remainder	Model	STL Model	X	X	X

stlreminder may mean	Model fit remainder	Model	STL Model	X	X	X
stlreminder nov mean	Model fit remainder	Model	STL Model	X	X	X
stlreminder oct mean	Model fit remainder	Model	STL Model	X	X	X
stlreminder sep mean	Model fit remainder	Model	STL Model	X	X	X
stltrend apr mean	Model trend mean	Model	STL Model	X	X	X
stltrend aug mean	Model trend mean	Model	STL Model	X	X	X
stltrend dec mean	Model trend mean	Model	STL Model	X	X	X
stltrend feb mean	Model trend mean	Model	STL Model	X	X	X
stltrend jan mean	Model trend mean	Model	STL Model	X	X	X
stltrend jul mean	Model trend mean	Model	STL Model	X	X	X
stltrend jun mean	Model trend mean	Model	STL Model	X	X	X
stltrend mar mean	Model trend mean	Model	STL Model	X	X	X
stltrend may mean	Model trend mean	Model	STL Model	X	X	X
stltrend nov mean	Model trend mean	Model	STL Model	X	X	X
stltrend oct mean	Model trend mean	Model	STL Model	X	X	X
stltrend sep mean	Model trend mean	Model	STL Model	X	X	X
stlweeklypattern fri mean	Model trend mean	Model	STL Model	X	X	X
stlweeklypattern mon mean	Model trend mean	Model	STL Model	X	X	X
stlweeklypattern sat mean	Model trend mean	Model	STL Model	X	X	X
stlweeklypattern sun mean	Model trend mean	Model	STL Model	X	X	X
stlweeklypattern thur mean	Model trend mean	Model	STL Model	X	X	X
stlweeklypattern tue mean	Model trend mean	Model	STL Model	X	X	X
stlweeklypattern wed mean	Model trend mean	Model	STL Model	X	X	X

Feature Code	Description	Category	Type	U	C	
breakouts max 10 1 2	Number of breakouts (various inputs)	Pattern	Breakout	X	X	X
breakouts max 10 1 3	Number of breakouts (various inputs)	Pattern	Breakout	X	X	X
breakouts max 10 1 5	Number of breakouts (various inputs)	Pattern	Breakout	X	X	X
breakouts max 10 2 2	Number of breakouts (various inputs)	Pattern	Breakout	X	X	X
breakouts max 10 2 3	Number of breakouts (various inputs)	Pattern	Breakout	X	X	X
breakouts max 10 2 5	Number of breakouts (various inputs)	Pattern	Breakout	X	X	X
breakouts max 10 5 2	Number of breakouts (various inputs)	Pattern	Breakout	X	X	X
breakouts max 10 5 3	Number of breakouts (various inputs)	Pattern	Breakout	X	X	X
breakouts max 10 5 5	Number of breakouts (various inputs)	Pattern	Breakout	X	X	X
breakouts max 30 1 2	Number of breakouts (various inputs)	Pattern	Breakout	X	X	X
breakouts max 30 1 3	Number of breakouts (various inputs)	Pattern	Breakout	X	X	X
breakouts max 30 1 5	Number of breakouts (various inputs)	Pattern	Breakout	X	X	X
breakouts max 30 2 2	Number of breakouts (various inputs)	Pattern	Breakout	X	X	X
breakouts max 30 2 3	Number of breakouts (various inputs)	Pattern	Breakout	X	X	X

breakouts max 30 2 5	Number of breakouts (various inputs)	Pattern	Breakout	X	X	X
breakouts max 30 5 2	Number of breakouts (various inputs)	Pattern	Breakout	X	X	X
breakouts max 30 5 3	Number of breakouts (various inputs)	Pattern	Breakout	X	X	X
breakouts max 30 5 5	Number of breakouts (various inputs)	Pattern	Breakout	X	X	X
breakouts max 60 1 2	Number of breakouts (various inputs)	Pattern	Breakout	X	X	X
breakouts max 60 1 3	Number of breakouts (various inputs)	Pattern	Breakout	X	X	X
breakouts max 60 1 5	Number of breakouts (various inputs)	Pattern	Breakout	X	X	X
breakouts max 60 2 2	Number of breakouts (various inputs)	Pattern	Breakout	X	X	X
breakouts max 60 2 3	Number of breakouts (various inputs)	Pattern	Breakout	X	X	X
breakouts max 60 2 5	Number of breakouts (various inputs)	Pattern	Breakout	X	X	X
breakouts max 60 5 2	Number of breakouts (various inputs)	Pattern	Breakout	X	X	X
breakouts max 60 5 3	Number of breakouts (various inputs)	Pattern	Breakout	X	X	X
breakouts max 60 5 5	Number of breakouts (various inputs)	Pattern	Breakout	X	X	X
dayfilterfreq 3 2h max	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 3 2h mean	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 3 2h min	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 3 2h std	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 3 4h max	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 3 4h mean	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 3 4h min	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 3 4h std	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 3 6h max	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 3 6h mean	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 3 6h min	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 3 6h std	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 3 8h max	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 3 8h mean	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 3 8h min	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 3 8h std	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 5 2h max	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 5 2h mean	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 5 2h min	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 5 2h std	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 5 4h max	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 5 4h mean	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 5 4h min	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 5 4h std	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 5 6h max	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 5 6h mean	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 5 6h min	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 5 6h std	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X

dayfilterfreq 5 8h max	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 5 8h mean	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 5 8h min	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 5 8h std	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 7 2h max	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 7 2h mean	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 7 2h min	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 7 2h std	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 7 4h max	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 7 4h mean	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 7 4h min	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 7 4h std	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 7 6h max	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 7 6h mean	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 7 6h min	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 7 6h std	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 7 8h max	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 7 8h mean	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 7 8h min	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 7 8h std	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 9 2h max	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 9 2h mean	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 9 2h min	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 9 2h std	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 9 4h max	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 9 4h mean	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 9 4h min	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 9 4h std	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 9 6h max	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 9 6h mean	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 9 6h min	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 9 6h std	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 9 8h max	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 9 8h mean	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 9 8h min	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
dayfilterfreq 9 8h std	Daily freq. (various inputs)	Pattern	DayFilter Freq.	X	X	X
jmotif inclasssim 168 6 14	jMotif In-class similiarity	Pattern	jMotif Pattern	X	X	
jmotif inclasssim 168 8 21	jMotif In-class similiarity	Pattern	jMotif Pattern	X	X	
jmotif inclasssim 24 12 12	jMotif In-class similiarity	Pattern	jMotif Pattern	X	X	
jmotif inclasssim 24 6 6	jMotif In-class similiarity	Pattern	jMotif Pattern	X	X	
jmotif inclasssim 24 8 8	jMotif In-class similiarity	Pattern	jMotif Pattern	X	X	
jmotiftemporal 168 6 14 max	jMotif temporal specificity	Pattern	jMotif Pattern	X	X	
jmotiftemporal 168 6 14 mean	jMotif temporal specificity	Pattern	jMotif Pattern	X	X	
jmotiftemporal 168 6 14 min	jMotif temporal specificity	Pattern	jMotif Pattern	X	X	
jmotiftemporal 168 6 14 std	jMotif temporal specificity	Pattern	jMotif Pattern	X	X	
jmotiftemporal 168 8 21 max	jMotif temporal specificity	Pattern	jMotif Pattern	X	X	
jmotiftemporal 168 8 21 mean	jMotif temporal specificity	Pattern	jMotif Pattern	X	X	
jmotiftemporal 168 8 21 min	jMotif temporal specificity	Pattern	jMotif Pattern	X	X	
jmotiftemporal 168 8 21 std	jMotif temporal specificity	Pattern	jMotif Pattern	X	X	
jmotiftemporal 24 12 12 max	jMotif temporal specificity	Pattern	jMotif Pattern	X	X	
jmotiftemporal 24 12 12 mean	jMotif temporal specificity	Pattern	jMotif Pattern	X	X	
jmotiftemporal 24 12 12 min	jMotif temporal specificity	Pattern	jMotif Pattern	X	X	
jmotiftemporal 24 12 12 std	jMotif temporal specificity	Pattern	jMotif Pattern	X	X	
jmotiftemporal 24 6 6 max	jMotif temporal specificity	Pattern	jMotif Pattern	X	X	

jmotiftemporal 24 6 6 mean	jMotif temporal specificity	Pattern	jMotif Pattern	X	X	
jmotiftemporal 24 6 6 min	jMotif temporal specificity	Pattern	jMotif Pattern	X	X	
jmotiftemporal 24 6 6 std	jMotif temporal specificity	Pattern	jMotif Pattern	X	X	
jmotiftemporal 24 8 8 max	jMotif temporal specificity	Pattern	jMotif Pattern	X	X	
jmotiftemporal 24 8 8 mean	jMotif temporal specificity	Pattern	jMotif Pattern	X	X	
jmotiftemporal 24 8 8 min	jMotif temporal specificity	Pattern	jMotif Pattern	X	X	
jmotiftemporal 24 8 8 std	jMotif temporal specificity	Pattern	jMotif Pattern	X	X	

Bibliography

2014. Code share. *Nature*, **514**(7524), 536–536.

2014. Journals unite for reproducibility. *Nature*, **515**(7525), 7–7.

Adamopoulou, Anna A., Tryferidis, Athanasios M., & Tzovaras, Dimitrios K. 2015. A context-aware method for building occupancy prediction. *Energy and Buildings.*

Agarwal, Yuvraj, Weng, Thomas, & Gupta, Rajesh K. 2009. The energy dashboard: improving the visibility of energy consumption at a campus-wide scale. *Page 55 of: Proceedings of the First ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Buildings (BuildSys 2009).* Berkeley, CA, USA: ACM Press.

Ahn, Ki-Uhn, & Park, Cheol-Soo. 2016. Correlation between occupants and energy consumption. *Energy and Buildings*, **116**(mar), 420–433.

An, Lianjun, Horesh, Raya, Chae, Young Tae, & Lee, Young M. 2012 (aug). Estimation of Thermal Parameters of Buildings Through Inverse Modeling and Clustering for a Portfolio of Buildings. *In: Proceedings of the Fifth National Conference of IBPSA-USA (SimBuild 2012).*

Augello, Agnese, Ortolani, Marco, Re, G Lo, & Gaglio, Salvatore. 2011. Sensor mining for user behavior profiling in intelligent environments. *Advances in Distributed Agent-Based Retrieval Tools*, jan, 143–158.

Basseville, M, & Nikiforov, I V. 1993. *Detection of abrupt changes: theory and application.*

Bellala, Gowtham, Marwah, Manish, Arlitt, Martin, Lyon, Geoff, & Bash, Cullen E. 2011. Towards an understanding of campus-scale power consumption. *Page 73 of: Proceedings of the Third ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Buildings (BuildSys 2011).* Seattle, WA, USA: ACM Request Permissions.

Bidoki, S M, Mahmoudi-Kohan, N, Sadreddini, M H, Zolghadri Jahromi, M, & Moghaddam, M P. 2010. Evaluating different clustering techniques for electricity customer classification. *Pages 1–5 of: Transmission and Distribution Conference and Exposition, 2010 IEEE PES.* New Orleans, LA, USA: IEEE.

Bogen, A C, Rashid, M, East, E W, & Ross, J. 2013. Evaluating a data clustering approach for life-cycle facility control. *Journal of Information Technology in Construction*, **Vol. 18**(apr), 99–118.

Borgeson, Sam. 2013. *Targeted Efficiency: Using Customer Meter Data to Improve Efficiency Program Outcomes.* PhD, University of California, Berkeley, Berkeley, CA, USA.

Borgeson, Sam, & Kwac, Jungsuk. 2015. *visdom: R package for energy data analytics.* R package version 0.9.

Breiman, Leo, Last, Michael, & Rice, John. Random Forests: Finding Quasars. *Pages 243–254 of: Statistical Challenges in Astronomy.* Springer Science Business Media.

Cabrera, David F Motta, & Zareipour, Hamidreza. 2013. Data association mining for identifying lighting energy waste patterns in educational institutes. *Energy and Buildings*, **62**(jul), 210–216.

Capozzoli, Alfonso, Lauro, Fiorella, & Khan, Imran. 2015. Fault detection analysis using data mining techniques for a cluster of smart office buildings. *Expert Systems with Applications*, **42**(9), 4324–4338.

Chen, Bei, Sinn, Mathieu, Ploennigs, Joern, & Schumann, Anika. 2014. Statistical Anomaly Detection in Mean and Variation of Energy Consumption. *Pages 3570–3575 of: Proceedings of the 22nd International Conference on Pattern Recognition (ICPR).* Stockholm, Sweden: IEEE.

Chicco, Gianfranco. 2012. Overview and performance assessment of the clustering methods for electrical load pattern grouping. *Energy*, **42**(1), 68–80.

Chicco, Gianfranco, & Ilie, I S. 2009. Support vector clustering of electrical load pattern data. *IEEE Transactions on Power Systems*, **24**(3), 1619–1628.

Chicco, Gianfranco, Ionel, O-M, & Porumb, Radu. 2013. Electrical Load Pattern Grouping Based on Centroid Model With Ant Colony Clustering. *IEEE Transactions on Power Systems*, **28**(2), 1706–1715.

Chou, Jui-Sheng, & Telaga, Abdi Suryadinata. 2014. Real-time detection of anomalous power consumption. *Renewable and Sustainable Energy Reviews*, **33**(C), 400–411.

Cleveland, Robert B, Cleveland, William S, McRae, Jean E, & Terpenning, Irma. 1990. STL: A seasonal-trend decomposition procedure based on loess. *Journal of Official Statistics*, **6**(1), 3–73.

Coughlin, K, Piette, Mary Ann, Goldman, Charles, & Kiliccote, Sila. 2009. Statistical analysis of baseline load models for non-residential buildings. *Energy and Buildings*.

De Silva, Daswin, Yu, Xinghuo, Alahakoon, Damminda, & Holmes, Grahame. 2011. A Data Mining Framework for Electricity Consumption Analysis From Meter Data. *IEEE Transactions on Industrial Informatics*, **7**(3), 399–407.

Diong, B., Zheng, G., & Ginn, M. 2015 (apr). Establishing the foundation for energy management on university campuses via data analytics. *Pages 1–7 of: Proceedings of the IEEE SoutheastCon 2015.*

Domahidi, Alexander, Ullmann, Fabian, Morari, Manfred, & Jones, Colin N. 2014. Learning decision rules for energy efficient building control. *Journal of Process Control*, **24**(6), 763–772.

Dong, Bing, & Lam, Khee Poh. 2011. Building energy and comfort management through occupant behaviour pattern detection based on a large-scale environmental sensor network. *Journal of Building Performance Simulation*, **4**(4), 359–369.

Dounis, Anastasios I. 2010. Artificial intelligence for energy conservation in buildings. *Advances in Building Energy Research*, **4**(1), 267–299.

Duarte, Carlos, Acker, Brad, Grosshans, Ray, Manic, Milos, Van Den Wymelenberg, Kevin, & Rieger, Craig. 2011 (aug). Prioritizing and Visualizing Energy Management and Control System Data to Provide Actionable Information For Building Operators. *In: Proceedings from 2011 Western Energy Policy Research Conference.*

Duda, Richard O., Hart, Peter E., & Stork, David G. 2012. *Pattern classification.* John Wiley & Sons.

D'Oca, Simona, & Hong, Tianzhen. 2015. Occupancy schedules learning process through a data mining framework. *Energy and Buildings*, **88**(feb), 395–408.

Effinger, Mark, Friedman, Hannah, & Moser, Dave. 2010 (sep). *Building Performance Tracking in Large Commercial Buildings: Tools and Strategies - Subtask 4.2 Research Report: Investigate Energy Performance Tracking Strategies in the Market.* Tech. rept.

Fagiani, Marco, Squartini, Stefano, Severini, Marco, & Piazza, Francesco. 2015 (jul). A novelty detection approach to identify the occurrence of leakage in smart gas and water grids. *Pages 1–8 of: Proceedings of the 2015 International Joint Conference on Neural Networks (IJCNN).*

Fan, Cheng, Xiao, Fu, & Shengwei, Wang. 2013 (jan). Prediction of Chiller Power Consumption Using Time Series Analysis and Artificial Neural Networks. *In: Proceedings of the 8th International Conference on Indoor Air Quality, Ventilation and Energy Conservation in Buildings (CLIMA 2013).*

Fan, Cheng, Xiao, Fu, & Yan, Chengchu. 2015a. A framework for knowledge discovery in massive building automation data and its application in building diagnostics. *Automation in Construction*, **50**(feb), 81–90.

Fan, Cheng, Xiao, Fu, Madsen, Henrik, & Wang, Dan. 2015b. Temporal Knowledge Discovery in Big BAS Data for Building Energy Management. *Energy and Buildings.*

Fels, Margaret F. 1986. PRISM: an introduction. *Energy and Buildings*, **9**(1-2), 5–18.

Figueiredo, V, Rodrigues, F, Vale, Z, & Gouveia, J B. 2005. An Electric Energy Consumer Characterization Framework Based on Data Mining Techniques. *IEEE Transactions on Power Systems*, **20**(2), 596–602.

Florita, Anthony R, Brackney, Larry J, Otanicar, Todd P, & Robertson, Jeffrey. 2012 (jul). Classification of Commerical Building Electrical Demand Profiles for Energy Storage Applications. *In: Proceedings of ASME 2012 6th International Conference on Energy Sustainability & 10th Fuel Cell Science, Engineering and Technology Conference (ESFuelCell2012).*

Fontugne, Romain, Tremblay, Nicolas, Borgnat, Pierre, Flandrin, Patrick, & Esaki, Hiroshi. 2013a (mar). Mining Anomalous Electricity Consumption Using Ensemble Empirical Mode Decomposition. *Pages 5238–5242 of: Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on.*

Fontugne, Romain, Ortiz, Jorge, Tremblay, Nicolas, Borgnat, Pierre, Flandrin, Patrick, Fukuda, Kensuke, Culler, David, & Esaki, Hiroshi. 2013b. Strip, bind, and search: a method for identifying abnormal energy consumption in buildings. *Pages 129–140 of: Proceedings of the 12th ACM/IEEE Int Conference on Information Processing in Sensor Network (ISPN 2013).* Philadelphia, PA, USA: ACM.

Forlines, Clifton, & Wittenburg, Kent. 2010. Wakame: Sense Making of Multi-dimensional Spatial-temporal Data. *Pages 33–40 of: Proceedings of the International Conference on Advanced Visual Interfaces (AVI 2010).* AVI '10. Roma, Italy: ACM.

Fulcher, B. D., Little, M. A., & Jones, N. S. 2013. Highly comparative time-series analysis: the empirical structure of time series and their methods. *Journal of The Royal Society Interface*, **10**(83), 20130048–20130048.

Gaitani, N, Lehmann, C, Santamouris, M, Mihalakakou, G, & Patargias, P. 2010. Using principal component and cluster analysis in the heating evaluation of the school building sector. *Applied Energy*, **87**(6), 2079–2086.

Gayeski, Nicholas, Kleindienst, Sian, Gagne, Jaime, Werntz, Bradley, Cruz, Ryan, & Samouhos, Stephen. 2015 (jun). *Data and Interfaces for Advanced Building Operations and Maintenance - RP 1633 Final Report.* Tech. rept. ASHRAE.

Georgescu, M, & Mezic, I. 2014. Site-Level Energy Monitoring and Analysis using Koopman Operator Methods. *In: Proceedings of the 2014 ASHRAE/IBPSA-USA Building Simulation Conference (SimBuild 2014).* Atlanta, GA, USA: ASHRAE/IBPSA.

Geurts, Pierre, Irrthum, Alexandre, & Wehenkel, Louis. 2009. Supervised learning with decision tree-based methods in computational and systems biology. *Mol. BioSyst.*, **5**(12), 1593.

Geyer, Philipp, Schlüter, Arno, & Cisar, Sasha. 2016. Application of clustering for the development of retrofit strategies for large building stocks. *Advanced Engineering Informatics.*

Goldin, Dina Q, & Kanellakis, Paris C. 1995. On similiarity queries for time-series data: constraint specification and implementation. *Pages 137–153 of: Principles and Practice of Constraint Programming (CP '95).* Springer.

Granderson, Jessica, Piette, Mary Ann, & Ghatikar, Girish. 2010. Building energy information systems: user case studies. *Energy Efficiency*, **4**(1), 17–30.

Green, R., Staffell, I., & Vasilakos, N. 2014. Divide and Conquer K-Means Clustering of Demand Data Allows Rapid and Accurate Simulations of the British Electricity System. *IEEE Transactions on Engineering Management*, **61**(2), 251–260.

Greensfelder, Erik, Friedman, Hannah, & Crowe, Eliot. 2010 (nov). *Building Performance Tracking in Large Commercial Buildings: Tools and Strategies - Subtask 4.4 Research Report: Characterization of Building Performance Metrics Tracking Methodologies.* Tech. rept.

Gullo, Francesco, Ponti, Giovanni, Tagarelli, Andrea, Ruffolo, Massimiliano, & Labate, Diego. 2009. Low-voltage electricity customer profiling based on load data clustering. *Pages 330–333 of: Proceedings of the 2009 International Database Engineering & Applications Symposium (IDEAS 2009).* Calabria, Italy: ACM.

Habib, U., & Zucker, G. 2015 (dec). Finding the Different Patterns in Buildings Data Using Bag of Words Representation with Clustering. *Pages 303–308 of: 2015 13th International Conference on Frontiers of Information Technology (FIT).*

Hao, M., Marwah, M., Janetzko, H., Sharma, R., Keim, D. A., Dayal, U., Patnaik, D., & Ramakrishnan, N. 2011. Visualizing frequent patterns in large multivariate time series. *Pages 78680J–78680J–10 of: Proceedings of SPIE Conference on Visualization and Data Analysis 2011*, vol. 7868. San Francisco, CA, USA: SPIE.

Hastie, Trevor, Tibshirani, Robert, & Friedman, J. H. 2009. *The elements of statistical learning: data mining, inference, and prediction.* 2nd ed edn. Springer series in statistics. New York, NY: Springer.

Heidarinejad, Mohammad, Dahlhausen, Matthew, McMahon, Sean, Pyke, Chris, & Srebric, Jelena. 2014. Cluster analysis of simulated energy use for LEED certified U.S. office buildings. *Energy and Buildings*, **85**(dec), 86–97.

Hong, Dezhi, Ortiz, Jorge, Whitehouse, Kamin, & Culler, David. 2013 (nov). Towards Automatic Spatial Verification of Sensor Placement in Buildings. *In: Proceedings of the 5th ACM Workshop on Embedded Systems For Energy-Efficient Buildings (BuildSys 2013).*

Iglesias, Félix, & Kastner, Wolfgang. 2013. Analysis of Similarity Measures in Times Series Clustering for the Discovery of Building Energy Patterns. *Energies*, **6**(2), 579–597.

Ioannidis, Dimosthenis, Fotiadou, Angeliki, Krinidis, Stelios, Stavropoulos, George, Tzovaras, Dimitrios, & Likothanassis, Spiridon. 2015. Big Data and Visual Analytics for Building Performance Comparison. *Pages 421–430 of:* Chbeir, Richard, Manolopoulos, Yannis, Maglogiannis, Ilias, & Alhajj, Reda (eds), *Artificial Intelligence Applications and Innovations.* IFIP Advances in Information and Communication Technology, no. 458. Springer International Publishing. DOI: 10.1007/978-3-319-23868-5_30.

Jacob, Dirk, Dietz, Sebastian, Komhard, Susanne, Neumann, Christian, & Herkel, Sebastian. 2010. Black-box models for fault detection and performance monitoring of buildings. *Journal of Building Performance Simulation*, **3**(1), 53–62.

James, Gareth, Witten, Daniela, Hastie, Trevor, & Tibshirani, Robert. 2013. *An introduction to statistical learning.* Vol. 112. Springer.

James, Nicholas A, Kejariwal, Arun, & Matteson, David S. 2014. Leveraging Cloud Data to Mitigate User Experience from Breaking Bad. *arXiv preprint arXiv:1411.7955.*

Janetzko, Halldór, Stoffel, Florian, Mittelstädt, Sebastian, & Keim, Daniel A. 2013. Anomaly detection for visual analytics of power consumption data. *Computers & Graphics*, **38**(oct), 27–37.

Jang, Dongsik, Eom, Jiyong, Jae Park, Min, & Jeung Rho, Jae. 2016. Variability of electricity load patterns and its effect on demand response: A critical peak pricing experiment on Korean commercial and industrial customers. *Energy Policy*, **88**(jan), 11–26.

Jarrah Nezhad, Aylin, Wijaya, Tri Kurniawan, Vasirani, Matteo, & Aberer, Karl. 2014. SmartD: Smart Meter Data Analytics Dashboard. *Pages 213–214 of: Proceedings of the 5th International Conference on Future Energy Systems (e-Energy '14).* e-Energy '14. New York, NY, USA: ACM.

Kelly Kissock, J., & Eger, Carl. 2008. Measuring industrial energy savings. *Applied Energy*, **85**(5), 347–361.

Keogh, E, & Kasetty, S. 2003. On the need for time series data mining benchmarks: A survey and empirical demonstration. *Data Mining and Knowledge Discovery*, **7**(4), 349–371.

Keogh, Eamonn J, Lin, Jessica, & Fu, Ada. 2005. Hot sax: Efficiently finding the most unusual time series subsequence. *In: Proceedings of the Fifth IEEE International Conference on Data Mining (ICDM'05).* Houston, TX, USA: IEEE.

Kissock, J. Kelly, & Eger, Carl. 2008. Measuring industrial energy savings. *Applied Energy*, **85**(5), 347–361.

Krarti, Moncef. 2003. An Overview of Artificial Intelligence-Based Methods for Building Energy Systems. *Journal of Solar Energy Engineering*, **125**(3), 331.

Kusiak, Andrew, & Song, Zhe. 2008. Clustering-Based Performance Optimization of the Boiler Turbine System. *IEEE Transactions on Energy Conversion*, **23**(2), 651–658.

Lam, Joseph C, Wan, Kevin K W, Cheung, K L, & Yang, Liu. 2008. Principal component analysis of electricity use in office buildings. *Energy and Buildings*, **40**(5), 828–836.

Lange, B, Rodriguez, N, Puech, W, & Vasques, X. 2013. Discovering unexpected information using a building energy visualization tool. *Page 86500Q of:* Baskurt, Atilla M, & Sitnik, Robert (eds), *Proceedings of the IS&T/SPIE 2013 Conference on Electronic Imaging*, vol. 8650. San Francisco, CA, USA: SPIE.

Lange, Benoit, Rodriguez, Nancy, & Puech, William. 2012. Energy Consumption Improvement Through a Visualization Software. *Pages 161–184 of: Energy Efficiency - A Bridge to Low Carbon Economy.* Intech.

Lavin, Alexander, & Klabjan, Diego. 2014. Clustering time-series energy data from smart meters. *Energy Efficiency*, **8**(4), 681–689.

Le Cam, M, Zmeureanu, R, & Daoud, A. 2014. Application of Data Mining techniques for energy modeling of HVAC sub-systems. *In: Proceedings of eSim 2014 - IBPSA-Canada Conference.* Ottawa, Canada: IBPSA.

Lee, S H, Hong, Tianzhen, Piette, Mary Ann, & Taylor-Lange, S C. 2015. Energy retrofit analysis toolkits for commercial buildings: A review. *Energy*, jan.

Lehrer, David. 2009 (sep). *Research Scoping Report: Visualizing Information in Commercial Buildings.* Tech. rept. Center for the Built Environment - UC Berkeley.

Lehrer, David, & Vasudev, Janani. 2011 (nov). *Visualizing Energy Information in Commercial Buildings: A Study of Tools, Expert Users, and Building Occupants.* Tech. rept. Center for the Built Environment - UC Berkeley.

Li, Shun, & Wen, Jin. 2014. A model-based fault detection and diagnostic methodology based on PCA method and wavelet transform. *Energy and Buildings*, **68**(jan), 63–71.

Lin, Jessica, Keogh, Eamonn J, Lonardi, Stefano, & Chiu, Bill. 2003. A symbolic representation of time series, with implications for streaming algorithms. *Pages 2–11 of: Proceedings of the 8th ACM SIGMOD Workshop on Research Issues in Data Mining and Knowledge Discovery (DMKD '03).* San Diego, CA, USA: ACM.

Lin, Jessica, Keogh, Eamonn J, Lonardi, Stefano, Lankford, Jeffrey P, & Nystrom, Donna M. 2004. Visually mining and monitoring massive time series. *Pages 460–469 of: Proceedings of the 10th ACM SIGKDD international conference on Knowledge discovery and data mining (KDD '04).* Seattle, WA, USA: ACM.

Lin, Jessica, Keogh, Eamonn J, & Lonardi, Stefano. 2005. Visualizing and discovering non-trivial patterns in large time series databases. *Information Visualization*, **4**(2), 61–82.

Lin, Jessica, Keogh, Eamonn J, Wei, Li, & Lonardi, Stefano. 2007. Experiencing SAX: a novel symbolic representation of time series. *Data Mining and Knowledge Discovery*, **15**(2), 107–144.

Linda, O, Wijayasekara, D, Manic, M, & Rieger, C. 2012 (aug). Computational intelligence based anomaly detection for Building Energy Management Systems. *Pages 77–82 of: Proceedings of the 5th International Symposium on Resilient Control Systems (ISRCS 2012).*

Liu, Dandan, Chen, Qijun, Mori, Kazuyuki, & Kida, Yukio. 2010. A method for detecting abnormal electricity energy consumption in buildings. *Journal Of Computational Information Systems*, **14**(jan), 4887–4895.

Liu, Xiufeng, Golab, L., & Ilyas, I.F. 2015 (apr). SMAS: A smart meter data analytics system. *Pages 1476–1479 of: Proceedings of the 2015 IEEE 31st International Conference on Data Engineering (ICDE).*

Louppe, Gilles, Wehenkel, Louis, Sutera, Antonio, & Geurts, Pierre. 2013. Understanding variable importances in forests of randomized trees. *Pages 431–439 of: Advances in Neural Information Processing Systems.*

Manning, Christopher D., Raghavan, Prabhakar, & Schutze, Hinrich. Probabilistic information retrieval. *Pages 201–217 of: Introduction to Information Retrieval.* Cambridge University Press (CUP).

Mansur, Vitor, Carreira, Paulo, & Arsenio, Artur. 2015. A Learning Approach for Energy Efficiency Optimization by Occupancy Detection. *Pages 9–15 of: Internet of Things. User-Centric IoT*, vol. 150. Cham: Springer International Publishing.

Mathieu, Johanna L., Price, Phillip N., Kiliccote, Sila, & Piette, Mary Ann. 2011. Quantifying changes in building electricity use, with application to demand response. *Smart Grid, IEEE Transactions on*, **2**(3), 507–518.

May-Ostendorp, Peter, Henze, Gregor P, Corbin, Charles D, Rajagopalan, Balaji, & Felsmann, Clemens. 2011. Model-predictive control of mixed-mode buildings with rule extraction. *Building and Environment*, **46**(2), 428–437.

May-Ostendorp, Peter T, Henze, Gregor P, Rajagopalan, Balaji, & Corbin, Charles D. 2013. Extraction of supervisory building control rules from model predictive control of windows in a mixed mode building. *Journal of Building Performance Simulation*, **6**(3), 199–219.

Miller, Clayton, & Schlueter, Arno. 2015. Forensically discovering simulation feedback knowledge from a campus energy information system. *Pages 33–40 of: Proceedings of the 2015 Symposium on Simulation for Architecture and Urban Design (SimAUD 2015)*. Washington DC, USA: SCS.

Miller, Clayton, Nagy, Zoltan, & Schlueter, Arno. 2014 (jun). A seed dataset for a public, temporal data repository for energy informatics research on commercial building performance. *In: Proceedings of the 3rd Conf. on Future Energy Business & Energy Informatics*.

Miller, Clayton, Nagy, Zoltán, & Schlueter, Arno. 2015. Automated daily pattern filtering of measured building performance data. *Automation in Construction*, **49, Part A**(jan), 1–17.

Miller, Clayton, Nagy, Zoltan, & Schlueter, Arno. Submitted for publication. A review of unsupervised statistical learning and visual analytics techniques applied to performance analysis of non-residential buildings. *Renewable and Sustainable Energy Reviews*.

Mills, Evan. 2011. Building commissioning: a golden opportunity for reducing energy costs and greenhouse gas emissions in the United States. *Energy Efficiency*, **4**(2), 145–173.

Mirkin, Boris. 2012. *Clustering: A Data Recovery Approach, Second Edition*. CRC Computer Science & Data Analysis. Chapman & Hall.

Mitsa, Theophano. 2010. *Temporal Data Mining*. Chapman and Hall/CRC.

Morais, Jefferson, Pires, Yomara, Cardoso, Claudomir, & Klautau, Aldebaro. 2009. An Overview of Data Mining Techniques Applied to Power Systems. *Data Mining and Knowledge Discovery in Real Life Applications*, jan.

Morán, A, Fuertes, J J, Domínguez, M, Prada, M A, Alonso, S, & Barrientos, P. 2013. Analysis of electricity bills using visual continuous maps. *Neural Computing and Applications*, **23**(3-4), 645–655.

Nikolaou, Triantafyllia G, Kolokotsa, Dionysia S, Stavrakakis, George S, & Skias, Ioannis D. 2012. On the Application of Clustering Techniques for Office Buildings' Energy and Thermal Comfort Classification. *IEEE Transactions on Smart Grid*, **3**(4), 2196–2210.

Pan, Erte, Li, Husheng, Song, Lingyang, & Han, Zhu. 2015 (mar). Kernel-based non-parametric clustering for load profiling of big smart meter data. *Pages 2251–2255 of: Proceedings of the 2015 IEEE Wireless Communications and Networking Conference (WCNC)*.

Panapakidis, Ioannis, Alexiadis, Minas, & Papagiannis, Grigoris. 2015. Evaluation of the performance of clustering algorithms for a high voltage industrial consumer. *Engineering Applications of Artificial Intelligence*, **38**(C), 1–13.

Patel, P, Keogh, E, Lin, J, & Lonardi, S. 2002. Mining motifs in massive time series databases. *Pages 370–377 of: Proceedings of 2002 IEEE International Conference on Data Mining (ICDM 2002)*.

Patnaik, Debprakash, Marwah, Manish, Sharma, Ratnesh, & Ramakrishnan, Naren. 2009. Sustainable Operation and Management of Data Center Chillers Using Temporal Data Mining. *Pages 1305–1314 of: Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2009)*. KDD '09. New York, NY, USA: ACM.

Patnaik, Debprakash, Marwah, Manish, Sharma, Ratnesh K, & Ramakrishnan, Naren. 2010. Data Mining for Modeling Chiller Systems in Data Centers. *Pages 125–136 of: Advances in Intelligent Data Analysis IX*, vol. 6065. Berlin, Heidelberg: Springer Berlin Heidelberg.

Petcharat, Siriwarin, Chungpaibulpatana, Supachart, & Rakkwamsuk, Pattana. 2012. Assessment of potential energy saving using cluster analysis: A case study of lighting systems in buildings. *Energy and Buildings*, **52**(sep), 145–152.

Pieri, Stella Panayiota, IoannisTzouvadakis, & Santamouris, Mat. 2015. Identifying energy consumption patterns in the Attica hotel sector using cluster analysis techniques with the aim of reducing hotels' CO2 footprint. *Energy and Buildings*, **94**(may), 252–262.

Ploennigs, Joern, Chen, Bei, Schumann, Anika, & Brady, Niall. 2013 (jan). Exploiting Generalized Additive Models for Diagnosing Abnormal Energy Use in Buildings. *In: Proceedings of the 5th ACM Workshop On Embedded Systems For Energy-Efficient Buildings (Buildsys '13).*

Ploennigs, Joern, Chen, Bei, Palmes, Paulito, & Lloyd, Raymond. 2014. e2-Diagnoser: A System for Monitoring, Forecasting and Diagnosing Energy Usage. *Pages 1231–1234 of: Data Mining Workshop (ICDMW), 2014 IEEE International Conference on*. Shenzhen, China: IEEE.

Price, Philip. 2010. Methods for Analyzing Electric Load Shape and its Variability. *Lawrence Berkeley National Laboratory*, aug.

Ramos, Sérgio, Duarte, JMM, Soares, João, Vale, Zita, & Duarte, Fernando J. 2012. Typical load profiles in the smart grid context—A clustering methods comparison. *Pages 1–8 of: Proceedings of the Power and Energy Society General Meeting, 2012 IEEE*. San Diego, CA, USA: IEEE.

Reddy, T. Agami. 2011. *Applied data analysis and modeling for energy engineers and scientists*. Springer Science & Business Media.

Reinhardt, Andreas, & Koessler, Sebastian. 2014. PowerSAX: Fast motif matching in distributed power meter data using symbolic representations. *Pages 531–538 of: Proceedings of 9th IEEE International Workshop on Practical Issues in Building Sensor Network Applications (SenseApp 2014)*. Edmonton, Canada: IEEE.

Ruparathna, Rajeev, Hewage, Kasun, & Sadiq, Rehan. 2016. Improving the energy efficiency of the existing building stock: A critical review of commercial and institutional buildings. *Renewable and Sustainable Energy Reviews*, **53**(jan), 1032–1045.

Räsänen, Teemu, & Kolehmainen, Mikko. 2009. Feature-based clustering for electricity use time series data. *Pages 401–412 of: Adaptive and Natural Computing Algorithms: 9th International Conference, ICANNGA 2009*. Kuopio, Finland: Springer.

Räsänen, Teemu, Ruuskanen, Juhani, & Kolehmainen, Mikko. 2008. Reducing energy consumption by using self-organizing maps to create more personalized electricity use information. *Applied Energy*, **85**(9), 830–840.

Santamouris, M, Mihalakakou, G, Patargias, P, Gaitani, N, Sfakianaki, K, Papaglastra, M, Pavlou, C, Doukas, P, Primikiri, E, Geros, V, Assimakopoulos, M N, Mitoula, R, & Zerefos, S. 2007. Using intelligent clustering techniques to classify the energy performance of school buildings. *Energy and Buildings*, **39**(1), 45–51.

Schlueter, A., Geyer, P., & Cisar, S. 2016. Analysis of Georeferenced Building Data for the Identification and Evaluation of Thermal Microgrids. *Proceedings of the IEEE*, **104**(4), 713–725.

Sedano, Javier, Villar, José Ramón, Curiel, Leticia, De La Cal, Enrique, & Corchado, Emilio. 2009. Improving energy efficiency in buildings using machine intelligence. *Pages 773–782 of: Intelligent Data Engineering and Automated Learning-IDEAL 2009*. Springer.

Seem, John E. 2005. Pattern recognition algorithm for determining days of the week with similar energy consumption profiles. *Energy and Buildings*, **37**(2), 127–139.

Seem, John E. 2006. Using intelligent data analysis to detect abnormal energy consumption in buildings. *Energy and Buildings*, **39**(1), 52–58.

Senin, P., & Malinchik, S. 2013a (dec). SAX-VSM: Interpretable Time Series Classification Using SAX and Vector Space Model. *Pages 1175–1180 of: 2013 IEEE 13th International Conference on Data Mining (ICDM).*

Senin, Pavel, & Malinchik, Sergey. 2013b. SAX-VSM: Interpretable Time Series Classification Using SAX and Vector Space Model. *In: 2013 IEEE 13th International Conference on Data Mining*. Institute of Electrical & Electronics Engineers (IEEE).

Shahzadeh, Abbas, Khosravi, Abbas, & Nahavandi, Saeid. 2015 (jul). Improving load forecast accuracy by clustering consumers using smart meter data. *Pages 1–7 of: Proceedings of the 2015 International Joint Conference on Neural Networks (IJCNN).*

Shao, Huijuan, Marwah, Manish, & Ramakrishnan, Naren. 2013 (jul). A Temporal Motif Mining Approach to Unsupervised Energy Disaggregation: Applications to Residential and Commercial Buildings. *Page 2250 of: Proceedings of the Twenty-Seventh AAAI Conference on Artificial Intelligence (AAAI 2013)*, vol. 250.

Shneiderman, B. 1996 (sep). The eyes have it: a task by data type taxonomy for information visualizations. *Pages 336–343 of: , IEEE Symposium on Visual Languages, 1996. Proceedings.*

Sun, Yan, Wu, Tin-Yu, Zhao, Guotao, & Guizani, M. 2015. Efficient Rule Engine for Smart Building Systems. *IEEE Transactions on Computers*, **64**(6), 1658–1669.

Taylor, James W., De Menezes, Lilian M., & McSharry, Patrick E. 2006. A comparison of univariate methods for forecasting electricity demand up to a day ahead. *International Journal of Forecasting*, **22**(1), 1–16.

Thanayankizil, L V, Ghai, S K, Chakraborty, D, & Seetharam, D P. 2012. Softgreen: Towards energy management of green office buildings with soft sensors. *Pages 1–6 of: Proceedings of the Fourth International Conference on Communication Systems and Networks (COMSNETS 2012).* Bangalore, India: IEEE.

The White House. 2016. FACT SHEET: Cities, Utilities, and Businesses Commit to Unlocking Access to Energy Data for Building Owners and Improving Energy Efficiency. jan.

Ulickey, Joy, Fackler, Tim, Koeppel, Eric, & Soper, Jonathan. 2010 (sep). *Building Performance Tracking in Large Commercial Buildings: Tools and Strategies - Subtast 4.3 Characterization of Fault Detection and Diagnostic (FDD) and Advanced Energy Information System (EIS) Tools.* Tech. rept.

Vale, Zita A, Ramos, Carlos, Ramos, Sérgio, & Pinto, Tiago. 2009. Data mining applications in power systems—Case-studies and future trends. *Pages 1–4 of: IEEE Transmission & Distribution Conference & Exposition: Asia and Pacific, 2009.* Seoul, South Korea: IEEE.

Verdu, S V, Garcia, M O, Senabre, C, Marin, A G, & Franco, F J G. 2006. Classification, Filtering, and Identification of Electrical Customer Load Patterns Through the Use of Self-Organizing Maps. *IEEE Transactions on Power Systems*, **21**(4), 1672–1682.

Wang, Shengwei, & Cui, J. 2005. Sensor-fault detection, diagnosis and estimation for centrifugal chiller systems using principal-component analysis method. *Applied Energy*, jan.

Wang, Shengwei, Zhou, Qiang, & Xiao, Fu. 2010. A system-level fault detection and diagnosis strategy for HVAC systems involving sensor faults. *Energy and Buildings*, **42**(4), 477–490.

Weiner, Peter. 1973. Linear pattern matching algorithms. *Pages 1–11 of: Switching and Automata Theory, 1973. SWAT '08. IEEE Conference Record of 14th Annual Symposium on.*

Wijayasekara, D., Linda, O., Manic, M., & Rieger, C. 2014. Mining Building Energy Management System Data Using Fuzzy Anomaly Detection and Linguistic Descriptions. *IEEE Transactions on Industrial Informatics*, **10**(3), 1829–1840.

Wrinch, Michael, El-Fouly, Tarek HM, & Wong, Steven. 2012. Anomaly detection of building systems using energy demand frequency domain analysis. *Pages 1–6 of: Power and Energy Society General Meeting, 2012 IEEE.* San Diego, CA, USA: IEEE.

Xiao, Fu, & Fan, Cheng. 2014. Data mining in building automation system for improving building operational performance. *Energy and Buildings*, **75**(jun), 109–118.

Yarbrough, I., Sun, Q., Reeves, D. C., Hackman, K., Bennett, R., & Henshel, D. S. 2015. Visualizing building energy demand for building peak energy analysis. *Energy and Buildings*, **91**(mar), 10–15.

Yoshida, Keigo, Inui, Minoru, Yairi, Takehisa, Machida, Kazuo, Shioya, Masaki, & Masukawa, Yoshio. 2008 (jan). Identification of Causal Variables for Building Energy Fault Detection by Semi-supervised LDA and Decision Boundary Analysis. *Pages 164–173 of: Data Mining Workshops, 2008. ICDMW '08. IEEE International Conference on.*

Yu, Zhun Jerry, Haghighat, Fariborz, Fung, Benjamin C M, & Zhou, Liang. 2012. A novel methodology for knowledge discovery through mining associations between building operational data. *Energy and Buildings*, **47**(apr), 430–440.

Yu, Zhun Jerry, Fung, Benjamin C M, & Haghighat, Fariborz. 2013. Extracting knowledge from building-related data — A data mining framework. *Building Simulation*, **6**(2), 207–222.

Zhou, Kai-le, Yang, Shan-lin, & Shen, Chao. 2013. A review of electric load classification in smart grid environment. *Renewable and Sustainable Energy Reviews*, **24**(C), 103–110.

Zhu, Yonghua, Jin, Xinqiao, & Du, Zhimin. 2012. Fault diagnosis for sensors in air handling unit based on neural network pre-processed by wavelet and fractal. *Energy and Buildings*, **44**(jan), 7–16.

Çakmak, Hüseyin Kemâl, Maa\s s, Heiko, Bach, Felix, & Kühnapfel, Uwe G. 2014. *A new framework for the analysis of large scale multi-rate power data*. Karlsruhe Institute of Technology.