Intersections Between the Academy and Practice

COLLABORATION: TECHNOLOGY, RESEARCH, PRACTICE

PAPERS FROM THE 2017 AIA/ACSA INTERSECTIONS SYMPOSIUM

SYMPOSIUM CO-CHAIRS:
John Folan, Carnegie Mellon University
Julie Ju-Youn Kim, Georgia Institute of Technology

AIA

ACSA

ABOUT THE SYMPOSIUM

2017 AIA/ACSA INTERSECTIONS SYMPOSIUM
Intersections Between the Academy and Practice:
Collaboration: Technology, Research, Practice

SYMPOSIUM CO-CHAIRS
John Folan, Carnegie Mellon University
Julie Ju-Youn Kim, Georgia Institute of Technology

The American Institute of Architects (AIA) and the Association of Collegiate Schools of Architecture (ACSA) are pleased to partner on this 3rd annual symposium, dedicated to the integration of education, research and practice of technologies at the 2017 AIA National Conference in Orlando. This symposium focuses on COLLABORATION. New technology and ways of working are helping to break down barriers between the different players in the construction process. Industry leaders already do a great deal to encourage collaboration among their teams, providing vision, collaboration-conducive work environments, collaboration technology, and by removing obstacles. Collaboration also happens outside the traditional AEC industry, between architects and cognitive scientists, between architects and community organizers, etc.

LEARN MORE
acsa-arch.org/intersections
aia.org/pages/144686-academic-engagement

CONTACT
Nissa Dahlin-Brown, AIA Academic Engagement, academicengagement@aia.org
Eric Wayne Ellis, ACSA Operations and Programs, info@acsa-arch.org

CONTENTS

The Building Genome Project:
Indentify faults in building energy performance

GRANT MOSEY
School of Architecture, University of Illinois at Urbana-Champaign

BRIAN DEAL
Department of Landscape Architecture: University of Illinois at Urbana-Champaign

This paper explores the use of new tools for the creation of novel methods of identifying faults in building energy performance remotely. With the rise in availability of interval utility data and the proliferation of machine learning processes, new methods are arising which promise to bridge the gap between architects, engineers, auditors, operators, and utility personnel. Utility use information, viewed with sufficient granularity, can offer a sort of "genome," that is a set of "genes" which are unique to a given building and can be decoded to provide information about the building's performance. The applications of algorithms to a large data set of these "genomes" can identify patterns across many buildings, providing the opportunity for identifying mechanical faults in a much larger sample of buildings that could previously be evaluated using traditional methods.

INTRODUCTION

In terms of buildings sciences, the energy performance of a building represents an outcome that has been generated through a myriad of disciplinary interactions. For example, performance is dependent on the shape of the building and the materials chosen by the architect, the mechanical systems are designed by the mechanical engineer, the maintenance of systems by the building operators, and the way the building is used – control settings etc., is determined by individual end users. With so many and varying pieces, expertise, and interests controlling a single outcome, it is perhaps no wonder that much of our existing building stock suffers from faults in energy performance design, operation, or construction.

The traditional method for uncovering building energy performance faults has been through the energy audit, a process by which a skilled analyst thoroughly surveys the building and relevant information about the building, both on and offsite. The general purpose of the audit is to identify problem areas and ascertain changes that can be made to enhance energy performance. While this tried and true method has doubtlessly produced measurable gains in the performance of individual buildings, it is not without its limitations. Audits are time consuming and labor intensive, requiring specially trained

professions capable of evaluating the interdisciplinary complexity of building energy systems – one building at a time. As a result only a very small portion of the total building stock can receive this valuable service in any given year. Audits are also something of a "snapshot" of a building, in that the auditor generally visits only for a short time and is able to observe only very limited operations first hand and follow up visits are even more time consuming and expensive. This means that many audited building have faults that go unidentified and continue to damage ongoing building energy performance even after close examination.

In order to improve the range and persistence of building energy performance analysis, new tools are needed that can evaluate much larger portions of the stock much more efficiently. Ideally, such tools would be capable of producing the type of analysis previously generated by energy audits in much more expedient and cost effective way. It would also be capable of seeing building performance for a far longer period of time, identifying patterns and thereby exposing faults which may be visible only over many months for example.

BUILDING ENERGY IN THE AGE OF BIG DATA

The simplest utility meters are those which measure use continuously and generate some value for usage which can be read and recorded manually one time for each utility billing cycle. For the problem of documenting usage in a given month to allow for accurate billing, these simple meters may prove adequate. However, advances in metering technology over the past decades have significantly expanded the scope and amount of data which can be collected and analyzed.

For example, electrical meters which measure consumption on much shorter "intervals" (15 seconds in some cases) are now becoming commonplace. These interval measurements allow for the monitoring of energy consumption on a functionally real-time basis, improving both the feedback and the ability to find common faults in usage patterns that previously were available only to expensive, smart building automation systems (BAS).

An example of collected electrical Interval utility data can be seen in Figure 1. This data belongs to a school in Illinois, and has been recorded every half-hour. The red line on the graph represents electrical use for a single day, in this case September 15, 2010. On this day, it
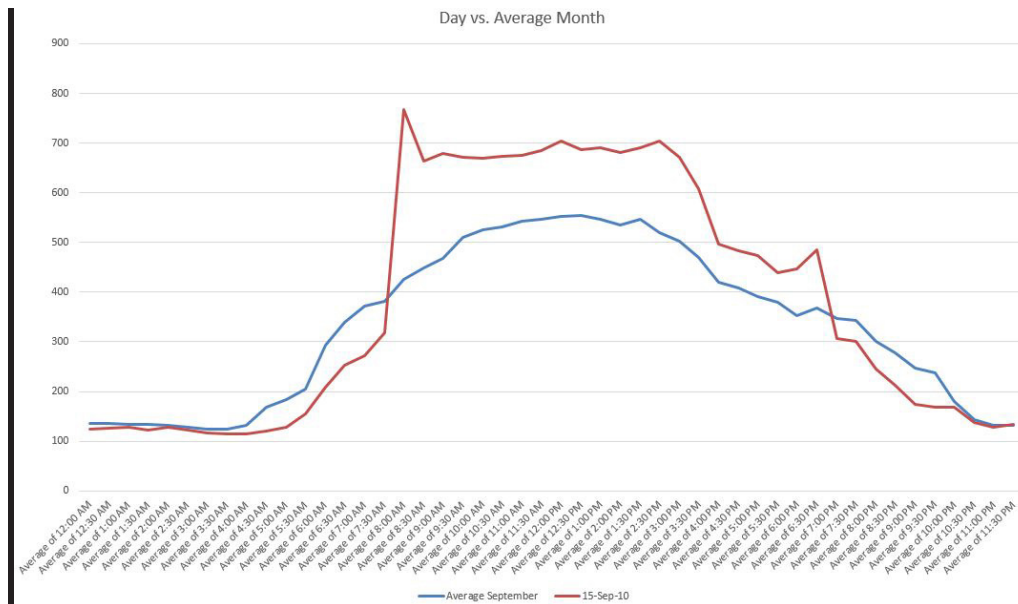
Figure 01: Eletrical Utility Load Profile for a School Building for the month of Semptember (Blue) and the day of September 15 (Red)

is visible that the early morning hours exhibit relatively low and consistent electricity usage. This likely represents a 'base load'. The term base load denotes the electricity which is constantly and consistently used by a building for always-on functions The blue line, represents the average electricity use throughout the day for the entire month of September. It paints a slightly different, although not incongruous picture. The blue, monthly line is slightly lower, which could be caused by changes in weather over the course of the month, or could simply reflect that weekends, when energy use is low, are included in this daily use curve. The curve is a bit smoother, with less of a plateau, which is not uncommon, as monthly curves tend to "average out" outlier values. For example, it may have been rainy in the morning with the skies clearing later in the day. This is captured by fine grained interval data, but gets averaged out in monthly data.

In this case, the "utility curves" reflect approximately what we might expect from a school. That is, the school uses little energy at night, has a period of very high usage while school is in session, and then a period of lower usage as evening activities take place, gradually reducing until all activities are over. One may note that usage appears higher than one would expect in the late evening, but there are a number of possible explanations for this, from late-night custodial work to errantly scheduled mechanical systems.

In the case of a single building, this type of analysis is straightforward and easy to see and quickly digest. The single day utility curve was generated with only 48 values. The monthly utility curve was generated with a more robust but still manageable 1,440 values. As one "zooms out," however, the picture becomes much more complicated. If this meter was installed in September 2010, then at the date of this writing, it has been operational for 76 months, or something like 2,300 days, meaning it has collected well over 100,000 readings. Multiply this number by the thousands of buildings with smart meters installed,

and the data stream quickly overwhelms our ability to provide this type of analysis for the entire building stock.

In this paper, we propose applying machine learning and genomic testing techniques to the problem of identifying energy performance faults in large numbers of buildings quickly and continuously – without human intervention. In the following sections we first review the literature on remote building energy modeling, we then propose expanding on that work with the introduction of a new remote energy analysis approach. We follow with information regarding the potential of the tool and some of the obstacles to its creation. We then provide our conceptualization of the way forward including genomic analysis. We conclude with thoughts on future research and next steps.

PREVIOUD WORK ON INTERVAL UTILITY DATA

[In recent years, there has been some academic interest in interval utility data and what it can reveal about buildings from a distance. Aksoezen et al (2014) focuses on the relationship between interval utility data and buildings with known parameters in an effort to find a correlation between building attributes and energy performance. Other research, such as that by Edwards et al (2012), has had a focus which is more predictive in nature, in this case using interval utility data to try and predict "next hour" consumption. Similarly, Espinoza et al (2005) considered utility data at the substation level with an eye toward predictive analytics. They were able to cluster substations into groups (e.g. residential, business) using utility profile analytics to work through end use analysis.

Still other work has focused on using interval data to discern building occupants building use patterns, rather than focusing on just the buildings themselves. Albert and Rajagopal (2006) use smart meter data to try and predict both occupancy and the characteristics of building users. Similarly, Kwac et al (2013) find some success using interval utility data from specific homes to segment utility customers by their lifestyles. The roll of occupant behavior as it relates to energy is also explored by Santin et al (2009), who try to evaluate how great of an

impact occupant behavior can have on heating and cooling energy use, again relying on residential hourly load profiles.

This type of data has even been used specifically to identify energy faults in buildings to some degree. For example, Brown et al (2009) monitor water, electricity, and gas use for 300 buildings over a seven year period. Using water as a proxy for occupancy and comparing this to the utility profiles, they were able to identify four common heating failure modes.

In terms of the private, building energy analysis sector, we find at least three companies who currently claim the ability to perform limited, auditing-type tasks with no building visits or minimal building visits, to wit FirstFuel, Agilis Energy, and Retroficiency (Lee et al 2014). Unfortunately, these are private entities with proprietary algorithms, so their methods and efficaciousness is difficult to ascertain.

Other research use a classification and regression tree (CART) algorithm to disaggregate energy usage using expert rules and then use statistical methods to identify outliers to identify faults. Lie et al (2010) examine this method for detecting abnormal electrical consumption for lighting in buildings. Using past electrical consumption records, occupancy, and time of day (as a surrogate for daylighting contribution), a decision tree is constructed using occupancy and time of day as independent variables. The analysis identified outliers when occupancy was low, yet electrical consumption was high. Khan et. al (2013) examined three different data mining techniques for detecting abnormal lighting energy consumption using hourly recorded energy consumption and peak demand (maximum power) data. CART, K-Means, and density-based spatial clustering of applications with noise (DBSCAN) were used. Interval meters can represent massive amounts of data which depend on some type of large-scale data analytics techniques.

Machine learning techniques that learn from data are now being developed. In one example, Lee et. al (2004) examine using a general regression neural-network (GRNN) model for on-line detection at the subsystem level. Energy fault detection techniques have also been embedded in building automation systems. They typically rely on a system of rules to determine a conditional probability for each of a plurality of possible fault causes given the detected fault (United States Patent Application, 2011, 2014). In this case the inputs to the system are embedded in the BAS.

While the academic literature would seem to fill in bits and pieces of what can be accomplished by applying data analytics to interval utility data, it seems as though no complete process has emerged that allows for remote auditing and fault detection. While the tools looked at in the private sector seem to have promise in the targeted evaluation of a client's building or portfolio of buildings, their inner-workings and exact capabilities remains opaque. Today, there exists no publically available means of evaluating interval utility data across wide numbers of buildings to detect energy performance faults in an efficient and low-cost way.

DEVELOPING THE TOOLS

Clearly, there exists a need for a tool which can put this data to its optimal use. With advances in machine learning, it would seem plausible that a program could be developed with would be capable of review huge amounts of utility data and finding anomalies. Unfortunately, the real problem of creating such a tool is not so straightforward.

First, there exists the problem of having a basis for comparison. If one imagines being given a utility profile and asked to find faults by comparing to some baseline, the question would quickly emerge as to what should represent the baseline. There seems to be, at minimum, three methods by which "standard behavior" could be established

Comparison to Self

Perhaps the simplest method of identifying faults would be comparing the current performance of a building to its past performance or performance under another condition. This method has the advantage of not requiring a large library of utility data for comparison, but would rather enable one to find anomalous data within the frame of just one building, if the building had been generating data for some period of time.

An example of this kind of analysis can be seen in Figure 02. Figure two represents the average, weekday, non-holiday energy consumption of a school for three months, June (blue), July (orange), and August (gray). Even without an outside basis for comparison, an anomaly appears immediately visible. Though the data appears phase-shifted, it is simple to see that in June and August, electricity seems to have a longer midday plateau. In contrast, July seems to have a sharp dip right at the center of the day. Considering that the summer schedule for this school was consistent throughout these three months, clearly something was changed operationally in this July, and then apparently changed back in August.

It could be that cooling equipment was set back in July, while it was allowed to cool an unoccupied building in part or all of June and August. It could be that lighting was turned on or off at different hours in this time. It could be that a piece of equipment was not operating and, therefore, not consuming power. Whatever the explanation, the anomaly is clear through comparing how the building operated at one point to how it operated at another.

This strategy has the downside of requiring some information about the building. Without knowing that the operation schedule had been static in this month, the anomaly detected could have been dismissed as a simple temporary change in use. Without know that the building was a school, its seasonality of operation would make little sense.

Comparison to a 'Real' Reference Building

If there is not a long history of interval utility operation, or information about context is unavailable, a more complicated method may be called for. Instead of comparing a buildings current operation to older operation, or comparing June to July, it may be more sensible to compare the operation of a school to the operation of all other schools in the same area or under the same climate conditions.
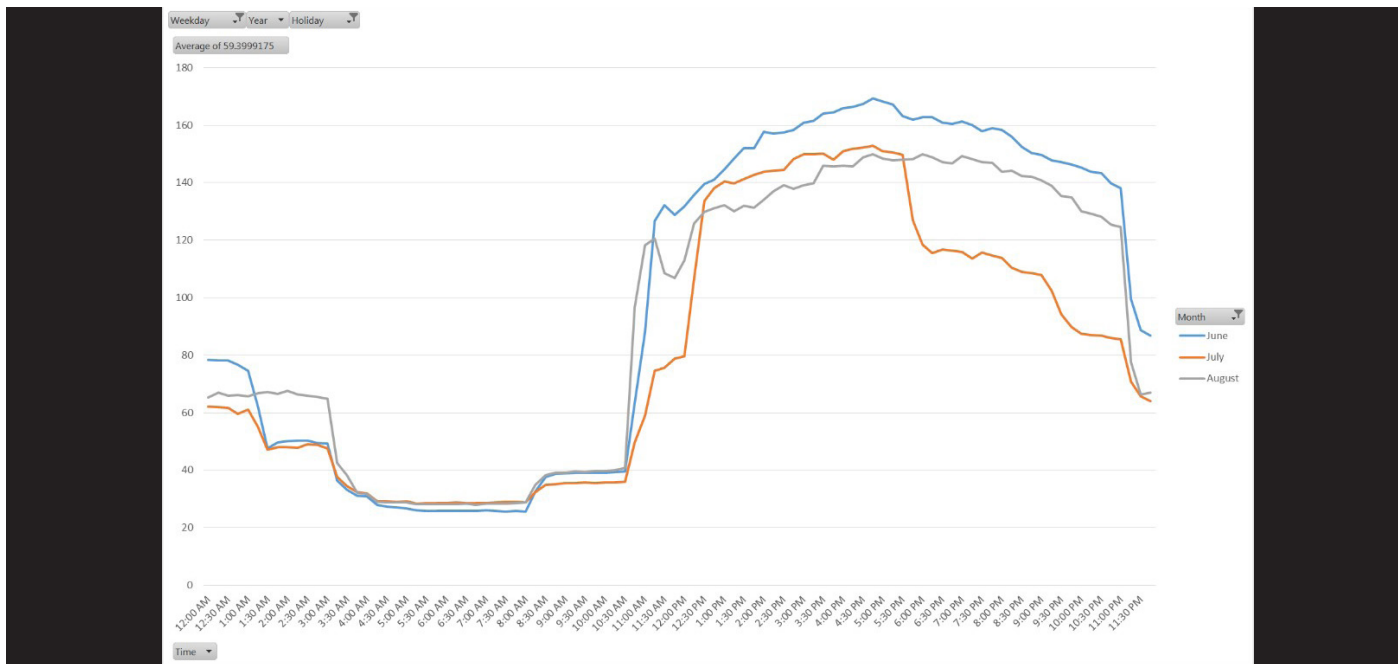
Figure 02: Electrical Utility Profiles for a school building in June (Blue), July (Orange) and August (Gray).

Of course, not all schools have similar utility curves. These can be affected by many variables including size, schedule, occupancy, and mechanical configuration. Yet, with a sufficient number of schools, one could ostensibly generate a "standard school profile" based on a library of data. Faults could then be detected in any one building by comparing it to the "standard" profile and recognizing where sharp differences occurred.

In a sense, this is not that different form the common practice of "benchmarking." Benchmarking compares the energy consumption of a building, usually on a normalized per unit area basis, to the consumption of a library of similar buildings. This technique has long been used to compare a subject building to the building stock as a whole. An interval utility comparison would be similar, except instead of making one comparison for one year, it would be capable of making thousands of comparisons at every half hour.

The downside to this method is that it requires a great deal of similar buildings to create a baseline for comparison. Because comparison of just a few buildings would be susceptible to "noise" in the data, a large sample of utility data would be required. It would also require knowing which utility streams belonged to buildings, or having an algorithm capable of making this distinction.

Comparison to a 'Simulated' Reference Building

Where a database of similar buildings is unavailable, or where the buildings are too distinct from one another to create a true average profile for comparison, the most sensible method may be comparing the real-world utility profile of a given building to a simulated utility profile.

Consider the example building shown in Figure 03. In these two graphs, the real utility profile of an academic building on the University of Illinois Campus (called the "Bill" profile, i.e. from utility bills) is compared to an output of a simulated version of the building (called the "Model" profile, i.e. from an energy model). Because the model is impervious to things like equipment breaking or errors of operation, it can represent how the building "should" perform. In looking at this example, we see the building is using more steam that would be expected during the spring shoulder season. Likewise, the building is using less electricity than expected in June, and more in September.

METHODOLOGICAL CONSTRUCTION

Our methodology for defining acceptable energy use patterns for a particular building includes using real and simulated reference data to detect degradations in energy efficiency performance of a particular building, and to diagnose probable faults. A preliminary flow chart for this process is shown in Figure 04. It shows the relationship between data management, analytical process, and reporting necessary to accomplish this goal. The high-level logic for our automated process includes using energy consumption data along with building classification information to first identify the correct building reference and then to diagnose potential energy-wasting faults in individual buildings in a continuous analytical process.

Building Screening

Continuous fault detection and diagnostic analysis provides information at a rate much faster than an organization's ability to respond with appropriate follow-up analysis and physical repairs. We propose a selection process that prioritizes buildings with the greatest savings potential. The process parameters can be adjusted to select for the desired number of buildings. Several metrics could be used as an initial screening process when selecting buildings for fault detection and diagnostic analysis. Calculation and comparison of total energy use,
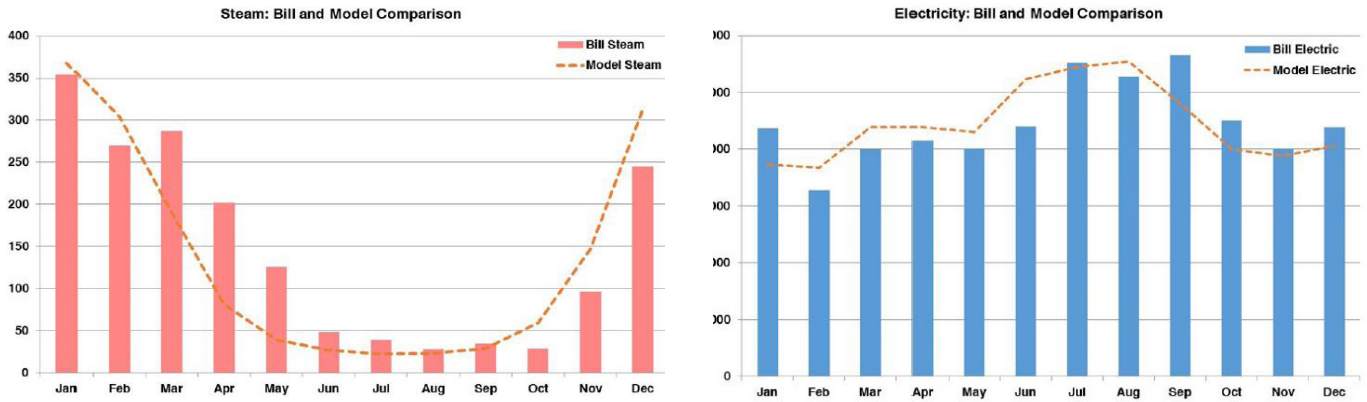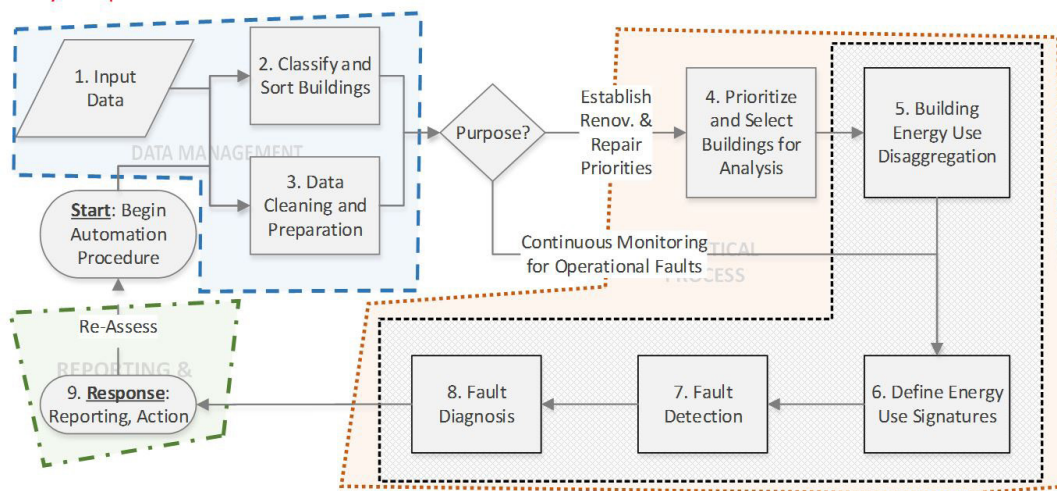
Figure 03: Comparision of Bill and Model Profiles on an annualized basis.

[Figure 04: A flow chart describing the automated process that uses energy consumption data and limted building information to identify and diagnose energy-wasting faults
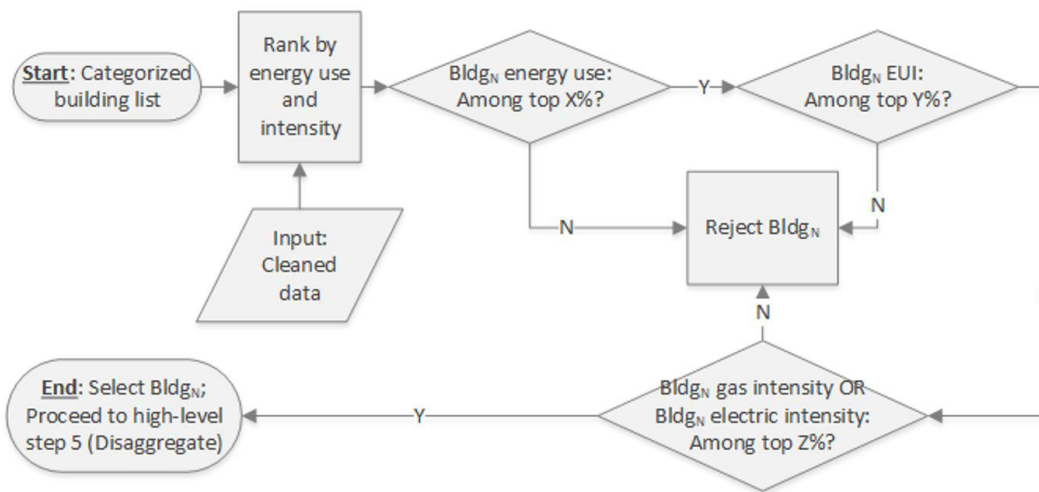

Figure 05: Example of a possible screening process that prioritizes and selects buildings for fault detection and diagnostic analysis.

energy use intensity (kBtu/ft2-year), electricity use intensity (kWh/ft2-year), and natural gas use intensity (therms/ft2-year) are among the parameters used in the screening process. Figure XX provides an example of a possible screening process.

Genomic Optimization Techniques.

One technique which has been explored is genomic modeling. Genomic modeling is an iterative, minimum-seeking algorithm. While it was originally designed for optimization and not classification problems, the challenge of matching utility profiles to a set of building "traits" which most likely generated them makes for a novel application.

In this process, a solution space of buildings with random "traits" are generated, then the "fittest" solutions move on to the next "generation." Fitness is determined, in this case, by the cost of function of aggregate difference from the given utility profile. Using this method, it is possible to find the set of building faults (in a fault detection example) or parameters (in a classification example), which are most likely to generate utility profiles like a given subject utility profile. While this approach is nascent, its early returns are promising.

CONCLUSION

The development and testing of such a technology would involve the cooperation of multiple independent actors. It requires Architects that can classify buildings; Building Energy Specialist who can correctly diagnose energy profiles; utility customers to share their data; computer scientists to automate the processes described - thousands of times per second.

Yet, if successfully developed, the technology has potential to allow all of these stakeholders to cooperate in new and interesting ways. Building engineers would be able to look up anomalous performance in their buildings and check them against a huge dataset to figure out what is likely causing the anomaly. Utility companies could send out annual reports to building owners showing potential faults in operations and potential ways to 'fix' them. Architects could specify systems by looking at the most common failure modalities in specific types of buildings in specific locations or working to address these potential faults in the design phase – to effectively 'nip it in the bud'. Policy officials could optimally allocate retrofitting incentives where their return on investment in terms of energy savings could be optimized. Monitoring companies could alert building engineers within seconds of a system fault or failure.

Decoding the building genome would require an immense amount of interdisciplinary cooperation and understanding, just as was the case in decoding the human genome. Yet, the collaboration required to create these tools would be sure to spur more collaboration, as the barriers that exist between project stakeholders, each with his/her own incentives and aims, could be substantially reduced.

WORKS CITED

Aksoezen, Mehmet, Magdalena Daniel, Uta Hassler, and Niklaus Kohler. "Building Age as an Indicator for Energy Consumption." Energy and Buildings 87 (2015): 74–86.

Albert, Adrian, and Ram Rajagopal. "Smart Meter Driven Segmentation: What Your Consumption Says about You." Power Systems, IEEE Transactions on 28, no. 4 (2013): 4019–4030.

Brown, Neil, A. J. Wright, Ashish Shukla, and Graeme Stuart. "Longitudinal Analysis of Energy Metering Data from Non-Domestic Buildings." Building Research & Information 38, no. 1 (2010): 80–91.

Edwards, Richard E., Joshua New, and Lynne E. Parker. "Predicting Future Hourly Residential Electrical Consumption: A Machine Learning Case Study." Energy and Buildings 49 (2012): 591–603.

Espinoza, Marcelo, Caroline Joye, Ronnie Belmans, and Bart De Moor. "Short-Term Load Forecasting, Profile Identification, and Customer Segmentation: A Methodology Based on Periodic Time Series." Power Systems, IEEE Transactions on 20, no. 3 (2005): 1622–1630.

Khan, Imran, Alfonso Capozzoli, Stefano Paolo Corgnati, Tania Cerquitelli. "Fault Detection Analysis of Building Energy Consumption Using Data Mining Techniques." Energy Procedia 42 (2013): 557-566.

Kwac, Jungsuk, Chin-Woo Tan, Nicole Sintov, June Flora, and Ram Rajagopal. "Utility Customer Segmentation Based on Smart Meter Data: Empirical Study." In Smart Grid Communications (SmartGridComm), 2013 IEEE International Conference on, 720–725. IEEE, 2013. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6688044.

Lee, Sang Hoon, Tianzhen Hong, and Marry Ann Piette. "Review of Existing Energy Retrofit Tools." Berkley, CA: Lawrence Berkley National Laboratories, 2014.

Lee, W-Y, JM House, N-H Kyong. "Subsystem level fault diagnosis of a building's air-handling unit using general regression neural networks." Applied Energy 77, no. 2 (2004): 153-170.

Liu, Dandan, Qijun Chen, Kazuyuki More, Yukio Kida. "A Method for Detecting Abnormal Electricity Consumption in Buildings." Journal of Computational Information Systems 6, no. 14 (2010): 4887-4895.

Santin, Olivia Guerra, Laure Itard, and Henk Visscher. "The Effect of Occupancy and Building Characteristics on Energy Use for Space and Water Heating in Dutch Residential Stock." Energy and Buildings 41, no. 11 (2009): 1223–1232.

United States Patent Application Publication Pub. No.: US 2011/0178977 A1, Drees, Pub. Date: Jul. 21, 2011, Building Management System with Fault Analysis, Johnson Controls Technology Company

United States Patent Application Publication Spivey et al. Pub. No.: US 2014/0324386 Al, Pub. Date: Oct. 30, 2014, BUILDING MANAGEMENT SYSTEM FALSE-POSITIVE FAULT INDICATIONS REDUCTION MECHANISM, ENERNOC INC., Boston' MA