Streamlining Building Efficiency Evaluation with DOE's Asset Score Preview

Supriya Goel, Nora Wang and Juan Gonzalez, Pacific Northwest National Laboratory
Henry Horsey and Nicholas Long, National Renewable Energy Laboratory

ABSTRACT

Building Energy Asset Score (Asset Score), developed by the Pacific Northwest National Laboratory (PNNL) for the U.S. Department of Energy (DOE), is a tool to help building owners and managers assess the efficiency of a building's energy-related systems and encourage investment in cost-effective improvements. The Asset Score uses an EnergyPlus model to provide a quick assessment of building energy performance with minimum user inputs of building characteristics and identifies upgrade opportunities. Even with a reduced set of user inputs, data collection remains a challenge for wide-spread adoption, especially when evaluating a large number of buildings. To address this, Asset Score Preview was developed to allow users to enter as few as seven building characteristics to quickly assess their buildings before a more in-depth analysis. A streamlined assessment from Preview to full Asset Score provides an easy entry point and also enables users who manage a large number of buildings to screen and prioritize buildings that can benefit most from a more detailed evaluation and possible energy efficiency upgrades without intensive data collection.

Introduction

The Building Energy Asset Scoring Tool is a web-based tool to help building owners, operators, and tenants understand the current and potential performance of a building's assets, which include building envelope (roof, walls and windows), mechanical systems, and electrical systems. The Asset Score is calculated based on standardized building operation conditions independent of an individual building's operational choices. The Asset Score is a simple 10-point scale where the higher the score, the more efficient the building asset relative to its population of peers. The population is constructed via robust building stock simulations for each building use type (Wang et al. 2016).

Asset Score Preview (Preview), a newly added capability within the Asset Scoring Tool, allows users to assess energy efficiency of a building with as few as seven inputs. Due to the limited inputs, Preview analysis is more appropriate for simple buildings with a single predominant use type, typical HVAC systems (for example, Preview doesn't analyze complex systems such as ground source heat pumps) and rectangular geometry. The seven building characteristics form the basis for inferring the remaining parameters to complete a full energy model. Preview then uses a regression model to calculate the possible range of the building's energy use based on stock simulations and uncertainty analysis. The estimated energy use range is mapped to the Asset Score scale to provide users with an estimated score range. To facilitate analysis of a large number of buildings, Preview also supports batch analysis, which allows users to upload a group of buildings through a spreadsheet and generate Preview scores all at once. Users can choose to convert their building from the Preview mode into a full Asset Score assessment mode by entering additional building characteristics. The Preview is intended to provide an easy, streamlined approach for quickly screening and evaluating building efficiency

with limited available information. The capability to convert a Preview building to a full Asset Score building would allow the Asset Scoring Tool and associated energy efficiency analysis to be more accessible to a broader range of users.

Asset Score Preview

Structure of Asset Scoring Tool

The Asset Scoring Tool is modular in design for clean separation of functionalities, easier testing and development. The core components of the Asset Scoring Tool application (Wang et al. 2015) are functionally separated into the following five subsystems: (i) User Interface, which offers two input modes – Preview and Full assessment (ii) Asset Score Application that translates all user inputs (for Preview and Full) into the Asset Score schema, (iii) Analytical Engine (using Facility Energy Decision System [PNNL 2014]) that infers additional building information based on user inputs (for Preview and Full), (iv) Modeling Engine that runs an EnergyPlus simulation via OpenStudio for a full assessment or an uncertainty analysis for a Preview analysis, and (v) Report Generator that post processes the analysis results.

Figure 1 shows the tool structure of the Preview component. The User Interface collects the minimal data points required for a Preview building analysis. The Asset Score Application populates the remaining data points required for a complete Asset Score model through the defaults database, which are displayed in the user interface for verification. The default values along with any user modifications are then passed to the Analytical Engine to infer the remaining building information such as thermal properties for envelope components and efficiencies for HVAC equipment. A complete Asset Score data model, with the defaults and inferences, is then sent to the Modeling Engine, which runs the regression analysis and estimates the energy use intensity (EUI) range for the building being analyzed. The estimated EUIs are then passed to the Report Generator to create a report (in PDF format), which includes a score range and estimated savings.
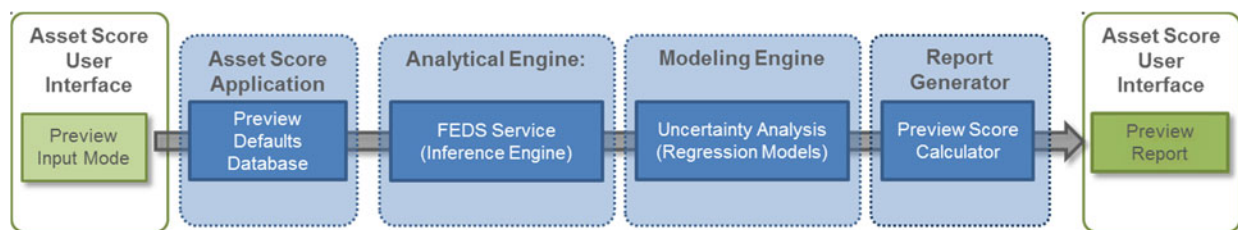

Figure 1. Asset Score Tool Components for Preview Analysis

Preview Inputs and Defaults Database

Asset Score Preview provides a platform for quick analysis of a building using limited user inputs. The intent of this functionality is to allow users to assess their building energy efficiency with limited investment in data collection and data entry. Preview requires high level inputs (Table 1) including building name, location, year of construction or major renovation, conditioned floor area, predominant use type, building orientation, and number of floors. Additional inputs of year of retrofits for lighting, HVAC and service hot water systems can be entered if applicable.

Based on the above data provided by the user, Preview's defaults database populates the remaining data required for creating a complete Asset Score model, which is then analyzed for its

energy efficiency and upgrade potential. The inputs provided by Preview's default database are listed in Table 2. The defaults database was built using the DOE Reference Buildings (USDOE 2015), the Prototype Buildings (Thornton et al. 2013), and the Commercial Buildings Energy Consumption Survey (CBECS 2003) database. The building use type, number of floors and conditioned floor area determine the default values for construction types, including roof, walls, and floor constructions as well as window frame type and window to wall ratio. They also determine the default HVAC system type and heating fuel type. The year of construction, along with the building use type determines the glazing type (single pane versus double or triple pane) and lighting fixture type. For instance, all buildings constructed before 1980 (and not having undergone an envelope retrofit) are assumed to have single pane windows and buildings constructed after 2004 are assumed to have double pane windows. Preview doesn't require any user inputs for a building's geometry. These buildings are modeled with a rectangular footprint for simplicity and the default database specifies its aspect ratio and floor height based on its use type, conditioned floor area, and number of floors. Buildings with complex geometry ("U" or "E" shaped buildings, or buildings with varying heights) are not applicable for Preview analysis as the simplification through a rectangular shape might result in significantly varying surface-to-volume ratios. Table 3 provides examples of the defaults in the Preview database. A building's location doesn't affect any of the defaults used, however this value along with a building's use type and year of construction are used by the analytical engine to infer the thermal properties of all envelope components as well as the efficiencies of all HVAC equipment.

Table 1: Required Inputs for Creating a Preview Building

1	Building Name
2	Location
3	Year of Construction or Major Renovation
4	Conditioned Floor Area
5	Use Type
6	Orientation
7	Number of Floors

Table 2: Inputs Provided by Preview's Defaults Database

1	Roof Construction Type	7	Lighting Fixture Type
2	Floor Construction Type	8	Percentage Area served by Lighting Fixture
3	Wall Construction Type	9	HVAC System Type
4	Window Framing Type	10	Heating Fuel Type
5	Window Glass Type	11	Service Water Heating Fuel Type
6	Window to Wall Ratio		

Table 3. Default Database Values for Select Building Types

Building Use Type	Conditioned Floor Area	Number of Floors	Construction Properties			HVAC System	
			Roof Construction	Window Framing Type	WWR	HVAC System Type	
Education	Any	Any	Built-up w/ metal deck	Metal	0.35	VAV with Electric Reheat	
Office	Less than or equal to 5,000 sq.ft	Any	Built-up w/ wood deck	Wood/Vinyl/Fiber glass	0.25	Packaged Rooftop Heat Pump	
Office	Greater than 5,000 sq.ft but less than 50,000 sq.ft	Any	Built-up w/ metal deck	Metal	0.33	Packaged Rooftop VAV with Electric Reheat	
Office	Greater than 50,000 sq.ft	Any	Built-up w/ metal deck	Metal	0.36	VAV with Hot-Water Reheat	
Multi-family (4 floors or greater)	Any	4 floors or greater	Built-up w/ metal deck	Metal	0.15	Packaged Rooftop Air Conditioner	
Multi-family (fewer than 4 floors)	None	Less than 4 floors	Built-up w/ metal deck	Metal	0.23	Packaged Rooftop Air Conditioner	

Preview Workflow

Buildings for Preview analysis can be entered either one at a time or in bulk through batch analysis. For a single building input, the User Interface guides users through a series of inputs to collect the seven data points required to query the defaults database and create a complete Asset Score model. The User Interface displays the defaults to allow users to take one of the following three actions: (i) edit the default values if the known values are different from the defaults (ii) verify the default values or (iii) mark the defaults as unknown (Figure 2). The user action affects the uncertainty analysis as more "unknown" values indicate a larger sample size. The inputs that usually require additional data gathering efforts are not displayed for user's verification. These inputs include system controls (such as fan control, energy recovery ventilation, demand control ventilation, daylighting control, occupancy sensor control etc.) as well as geometry inputs (such as floor-to-floor height, building dimensions and aspect ratio). They are evaluated as uncertain variables for all scenarios. This allows Preview to be a quick analysis tool without burdensome data collection requirements. The Preview output is a score range and estimated improvement potential (Figure 3).
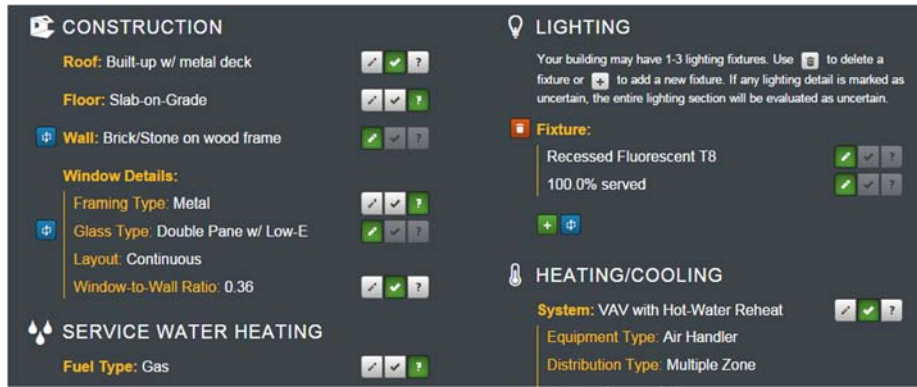
Figure 2. Screenshot of Preview Where Users Can Edit, Verify, or Mark the defaults

Figure 3. Preview Score Range Displayed on Asset Score Preview Report

The batch analysis capability allows for upload and analysis of a large number of buildings with a few simple steps. Results for these buildings can also be downloaded in a spreadsheet that provides information of the Preview score range for each building as well as the estimated minimum, maximum, and mean EUI. The Preview reports (in PDF format) can also be downloaded for all analyzed buildings through the batch analysis feature. Batch analysis offers a fast way to screen a large number of buildings with minimal data entry effort, allowing users to quickly identify buildings with the highest potential for energy savings.

Development of Uncertainty Analysis

The uncertainty analysis, built on regression models, estimates the EUI range (within the 95th percentile of the results) for a Preview building, based on the seven data points provided by the user as well as the defaults populated by the Preview database. The input values that are verified, edited or marked as unknown by the user affect the range of EUI calculated from the regression model. A larger number of uncertain inputs yield a wider EUI range. Some user inputs are more sensitive than others. For instance, the interior lighting power density for some use-types has a much higher degree of impact on a building's EUI than envelope parameters like roof or wall construction properties (Long et al. 2015). Hence the EUI range determined by the uncertainty analysis is not only affected by the number of uncertain parameters, but also the sensitivity of the parameters. The EUI range is post processed to give the corresponding score range, which is the main output in an Asset Score Preview report[1].

---

[1] The Preview score range is centered around the median EUI and reflects the skewedness of the EUI distribution. If the EUI distribution is positively skewed, the calculated score range reflects the same.

OpenStudio Analysis Framework

Development of the regression models required running hundreds of thousands of whole building simulations. The OpenStudio Analysis Framework (Long et al. 2014) was used to set up and manage the large quantity of simulation runs. The OpenStudio Analysis Framework allows for an efficient setup of analyses (e.g., sampling, optimization, etc.) by using an Excel-based spreadsheet to select the seed model, algorithm details, variables, and outputs. The seed model is the initial model from which all the model perturbations stem. Seed models were created for 22 buildings and were closely aligned to the PNNL Prototype Buildings (Thornton, et al., 2011). These seed models are in eXtensible Markup Language (XML) format with simplified geometry (all rectangular), uniformly distributed internal loads, and simplified thermal zone layout. The seed models contain default values for all the variables to be analyzed, such as the construction types, HVAC system types and their efficiency levels. The seed model defaults are of importance since the value of the input variables contained within the seed model (e.g. total floor area, U-factors, system efficiencies) are assumed to be the mean for that building type. A variation of the seed models was also created with different HVAC system types to include multiple system configurations. For example a medium office seed model was analyzed with a packaged variable air volume (VAV) with electric reheat system as well as a packaged rooftop air conditioner.

A total of 38 variables (Table 4), which form the inputs into the full Asset Score tool, were used to develop the regression models. Each variable was assigned a minimum, maximum, mean, standard deviation, and distribution (e.g. uniform, triangular) representative of typical efficiency ranges based on the vintage of the existing building stock as well as current technologies. Distributions that best represent the variable across the commercial building stock in the U.S. were then developed using these identified efficiency values (Wang et al. 2016). The OpenStudio Analysis Framework was used to simulate the seed models and their variations based on the pre-defined input distributions.

Table 4: Variables Analyzed for the Development of the Regression Models

No.	Variables	No.	Variables
1	Air Handler Fan Efficiency	20	Location
2	Aspect Ratio	21	Minimum Airflow Fraction
3	Chilled Water Reset	22	Number of Floors
4	Chiller Pump Control	23	Orientation
5	Condenser Pump Control	24	Perimeter Zone Depth
6	Cooling Efficiency	25	Roof Construction Type
7	Cooling Tower Control	26	Roof U-Value
8	Daylighting Control	27	Shading Height above Window
9	Demand Control Ventilations	28	Shading Projection Factor
10	Economizer	29	Supply Air Temperature Reset
11	Energy Recovery Ventilation	30	Total Floor Area
12	Fan Control	31	Wall Construction Type
13	Fan Static Pressure Reset	32	Wall U-Value
14	Floor Plate Area	33	Water Heater Efficiency
15	Floor R-Value	34	Window Solar Heat Gain Coefficient

16	Floor-To-Floor Height	35	Window Sill Height
17	Heating Efficiency	36	Window U-Value
18	Heating Fuel Type	37	Window Visible Light Transmittance
19	Interior Lighting Power Density	38	Window-To-Wall Ratio

Development of the Regression Model

Multiple regression techniques were investigated to determine which was best for use in meta-modeling. Initial testing revealed that non-linear regression models had significantly better R^2 and percentile error values. Therefore, a random forest algorithm was employed to develop the regression engine (Breiman, 2001, Breiman and Cutler, 2016) Several regression models were created for each seed model for a total of 43 regression models representing several HVAC system for each use-type.

To create each regression model 400 stratified samples of the parameter space, generated using Latin Hypercube Sampling (LHS), were run for each building type, system and climate zone combination, resulting in around 6,000 data-points for the 15 climate zones and each seed model analyzed. Each regression model was then randomly divided into two sets: a training set comprising 80% of the data points and a test set comprising the remaining 20% ,which was used to measure the accuracy of the regression model. The training set of around 4,800 data points was then passed to the random forest generating algorithm, contained in the R package random forest.[2] Figure 4 shows an example of the results of regression model testing and associated percentile error for a medium office seed model with a packaged VAV with electric reheat system. In this figure the red line is a 1-1 slope that represents the ideal relationship between simulated and predicted results, namely that they are the same. The majority of the test set falls very close to this line, and the R^2 value is approximately 0.95.
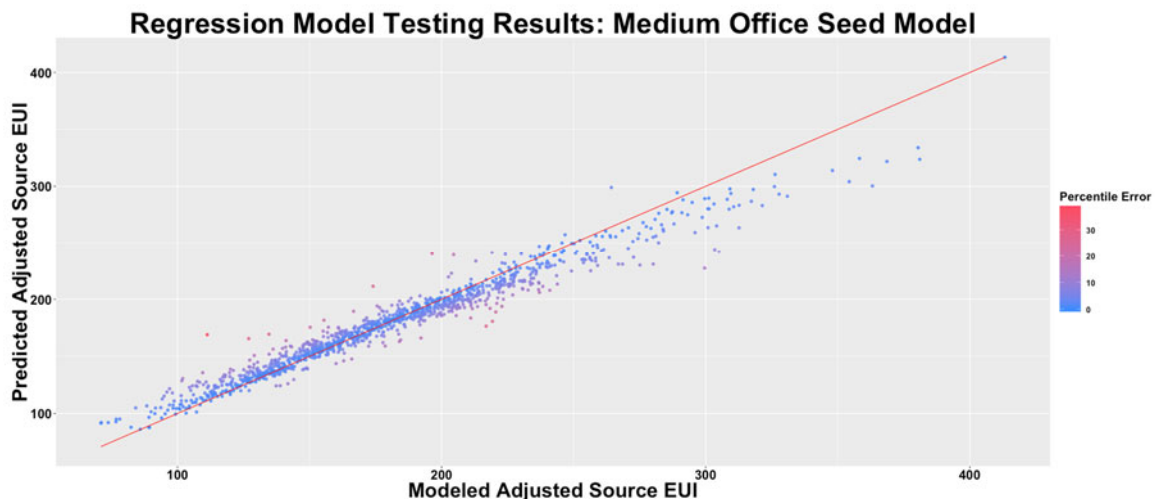

Figure 4. Results of Regression Model Testing For Medium Office Seed Model

In addition to R^2 values, the percentile error of each test datapoint was computed. This

---

[2] A random forest is a large aggregation of binary decision trees, each of which is independently trained. Although each random forest keeps robust internal accuracy statistics, the test set comprising 20% of the data points, or ~1200 data points is used to gauge the overall effectiveness of the regression.

percentile error represents the difference in percentile between the simulated and predicted results. This error measure is important as the Asset Score is roughly based off of a building's percentile among the estimated building stock, and thus this measure gives a better sense of the maximum score error returned by the regression model. The R^2 value of the predicted versus simulated EUIs ranged from 0.92 to 0.98, whereas the maximum percentile error could range as high as 45, particularly for chiller systems. There is therefore a significant possibility for score error should only a single data-point be used to calculate the score. This in large part drove the decision to use 10,000 predictions to estimate a score range as described in the next section, relying on the weak law of large numbers to ensure the robustness (Gallager, 2013).

EUI Prediction Methodology

The EUI range calculated from the regression models is determined by the information provided by the user on the certainty of different variable categories (HVAC, lighting, hot water, walls, roof, windows, and floor). The user input values and the default values are passed through the Asset Score XML and used to populate the regression engine. Variable categories specified as certain (i.e. either verified or edited by the user) are fixed for all data points and not sampled through the regression model. Variable categories specified as uncertain (i.e. marked as unknown by the user) are sampled over a triangular or uniform distribution depending on the type of variable category.

Figure 5 shows a flowchart of this process. Here the HVAC system variables are first designated as certain or uncertain depending on the system certainty. The remaining variables are subsequently parsed into certain and uncertain sets based on the user inputs. In the case of certain variables the values are fixed according to the XML file, while in the case of uncertain variables a triangular distribution centered at the default value determined by the Asset Score application and inferred by the analytical engine is used or a uniform distribution is assigned to the variable. Numerical variables such as lighting power density are sampled over triangular distributions and categorical variables such as wall construction types are sampled uniformly.
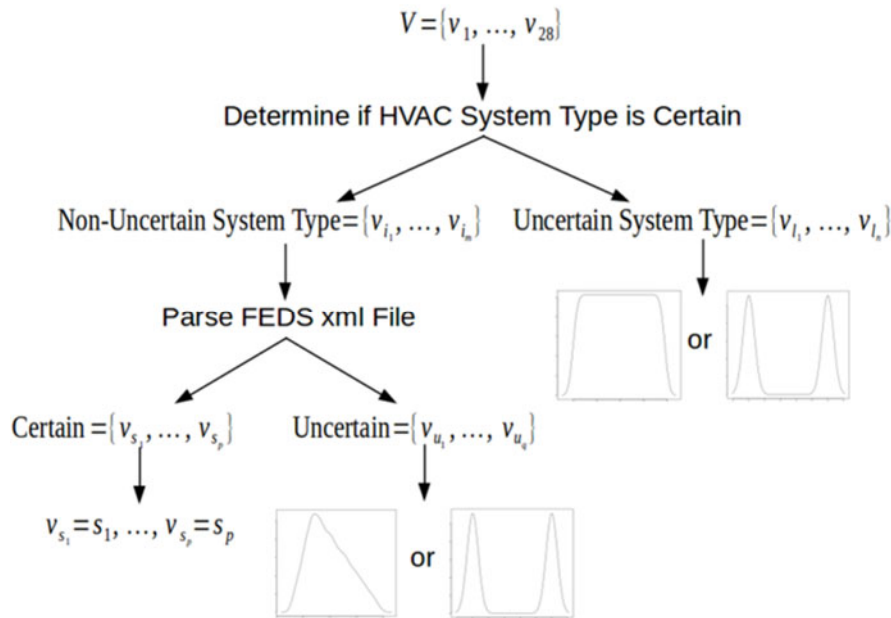

Figure 5. Workflow for Mapping Variables to Distributions

Testing the Uncertainty Analysis Model

The regression models were tested using two sets of models: (i) the seed models (i.e., simplified Prototype Buildings) developed for the uncertainty analysis, (ii) buildings entered by users into the Asset Scoring Tool during the pilot studies (pilot building models) conducted in 2012 and 2013 (Wang et. al. 2013). The intent of this testing process was to determine if the uncertainty analysis model could accurately reflect the efficiency of a building considering the limited data provided.

The seven basic building characteristics of the 22 seed models were entered into Preview to generate the score range for each seed model. The complete set of building characteristics was entered into the full Asset Score tool to generate the Asset Score using whole building energy simulation. The acceptance criteria employed in the testing process required the full Asset Score to be within the range returned by Preview. All buildings passed the test as the actual scores lay within the score range determined through Preview. Table 5 shows the example results. Building geometry and HVAC system controls, such as demand control ventilation, supply temperature reset etc., are always considered uncertain. This can have a significant impact on how a full building scores in reference to the Preview score range. The fact that some seed models score closer to the Preview high score and others closer to the low score depends on the sensitivity of the applicable HVAC system controls. For example, the medium office seed model is analyzed with a packaged VAV with electric reheat system with controls such as fan static pressure reset and minimum airflow fraction specified at 0.3. The Preview model assumes these HVAC controls to be uncertain and the EUI range determined includes the impact of a minimum airflow fraction specified at 0.4 and 0.3 as well as the VAV fans modeled with and without static pressure reset controls. A minimum airflow fraction input of 0.4 results in a much higher reheat energy use which causes the full seed model, modeled with a minimum airflow fraction at 0.3, to score closer to the Preview high score.

Table 5 Test Results for Select Seed Models

Seed Model	Preview Low Score	Preview High Score	Full Input Mode Score
Retail Stand Alone	6.5	9.5	6.5
Warehouse	5.5	8.5	8
Small Office	2.5	5.5	5
Medium Office	1.5	4.5	4
City Hall	3.5	6.5	4.5
Library	6	9	7
Small Hotel	5.5	8.5	8
Large Hotel	7	10	8.5

The pilot buildings were tested with two configurations, the minimal seven inputs with all Preview defaults marked unknown and the seven inputs with all defaults being edited or verified based on the actual building data. The first set, submitted with seven inputs and none of the default inputs verified, did not consistently display the actual scores within the uncertainty ranges. Over 160 pilot buildings were analyzed for this purpose. Of the tested buildings, 65% had scores within the range identified by Preview and the remaining 35% scored outside the

range identified by Preview. The alignment was dependent on how closely the preview defaults coincided with the actual values as well as the sensitivity of the parameters that did not coincide. Editing or verifying the default values to actual values significantly improved the alignment of the buildings analyzed.

A subset of 40 pilot buildings were randomly selected and entered in Preview with the seven minimal inputs and default values modified in accordance to the actual building inputs. In this analysis, the default values for lighting fixture types, HVAC system type, window glass type, and window-to-wall ratio were either verified or edited to reflect the actual systems present in the building. A higher degree of accuracy was observed where the actual score of the building fell within the score range calculated through the regression models for 95% of the test cases. Buildings falling outside the Preview score range were the highly efficient buildings with equipment performance ratios higher than typically found. Since equipment efficiencies, inferred by the Asset Score analytical engine, are based on CBECS data, these buildings were not well represented and the full score was much higher than the preview score. Figure 6 shows the example results of buildings tested in the second configuration. For most buildings tested, lighting power density was found to be the biggest factor affecting the correlation between the Preview score range and the actual building score. The HVAC system type and the window-to-wall ratio were also sensitive inputs that were found to affect the uncertainty analysis. Hence, with specifications of the lighting system, HVAC system type and window-to-wall ratio, the accuracy of the Preview score is significantly improved.
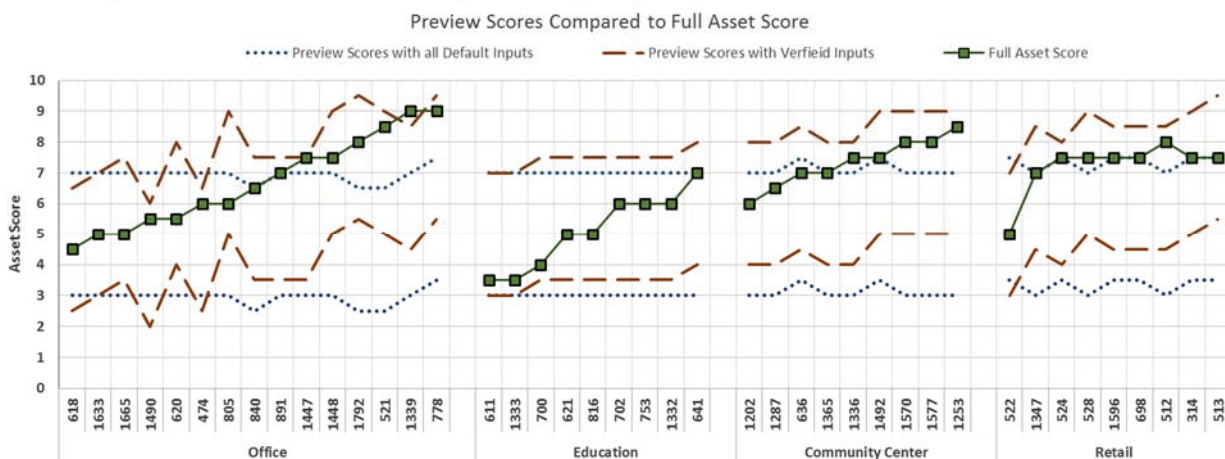

Figure 6. Testing Results for Pilot Buildings

Conclusion

The Asset Score Preview creates a viable framework for analyzing the efficiency of the building asset. With limited effort from users, it can provide useful information about a building's assets and identify buildings that have great potential for improvement or could benefit from a more in-depth analysis. The use of random forest models to ascertain the range of uncertainty in the building's EUI allows limited inputs from the user to generate meaningful results without the significant computation time and power required for more detailed whole building energy modeling. The validation of this framework shows that additional sensitive inputs verified by users can greatly improve the accuracy of the Preview results. Buildings with more complex geometry, mix-used type, or advanced HVAC systems may not be suitable for Preview because oversimplification likely results in low accuracy. The batch analysis capability

reduces the amount of manual work required to generate the Preview scores and can be used to pre-analyze large portfolios of buildings with relatively minimal initial effort. This offers a more friendly entry to the full capabilities of the Asset Score. The expected users include utility program administrators, cities energy program managers, and large real estate portfolio managers.

References

Breiman, Leo. 2001. Random Forests. Machine Learning 45 (1): 5–32.

Breiman, Leo, and Adele Cutler. 2016. "Random Forests for Scientific Discovery." http://www.math.usu.edu/adele/forests/ENAR _files/frame.html. (Accessed January 8.)

CBECS. 2003. Commercial Buildings Energy Consumption Survey 2003. Energy Information Administration, U.S. Department of Energy, Washington, D.C. Last accessed in July, 2009 at http://www.eia.doe.gov/emeu/cbecs/contents.html

Gallager, R. B. 2013. Stochastic Processes: Theory for Applications. Cambridge University Press, Cambridge, U.K.

Long, N., B. Ball, K. Fleming and D. Macumber. 2014. Scaling Building Energy Modeling Horizontally in the Cloud with OpenStudio. 2014 ASHRAE/IBPSA-USA Building Simulation Conference, Atlanta, GA.

Long, N., S. Goel and H. Horsey. 2015. U.S. Department of Energy's Asset Score Sensitivity and Scale Implementation. Building Simulation 2015

Thornton, B.A, M Rosenberg, EE Richman, W Wang, Y Xie, J Zhang, H Cho, VV Mendon, RA Athalye, and B Liu. 2011. Achieving the 30% Goal: Energy and Cost Savings Analysis of ASHRAE Standard 90.1-2010. PNNL-20405, Pacific Northwest National Laboratory, Richland, Washington.

USDOE. 2015b. Commercial Reference Buildings. http://energy.gov/eere/buildings/commercial-reference-buildings (accessed January 10, 2016)

PNNL. 2014. Facility Energy Decision System User's Guide, Release 7.0. PNNL-17848, Pacific Northwest National Laboratory, Richland, Washington. Available from http://www.pnnl.gov/feds/pdfs/FEDS_7-0_user_guide.pdf (December 2014).

Thornton, B.A, M Rosenberg, EE Richman, W Wang, Y Xie, J Zhang, H Cho, VV Mendon, RA Athalye, and B Liu. 2011. Achieving the 30% Goal: Energy and Cost Savings Analysis of ASHRAE Standard 90.1-2010. PNNL-20405, Pacific Northwest National Laboratory, Richland, Washington.

Wang, N., S. Goel, W. Gorrisen and A. Makhmalbaf. 2013. Understanding Building Infrastructure and Building Operation through DOE Asset Score Model: Lessons Learned from a Pilot Project. PNNL-SA-92883, ASHRAE Transactions, 119(pt. 2):Paper No. DE-13-C030

Wang, N., S. Goel, V. Srivastava and A. Makhmalbaf. 2015. Commercial Building Energy Asset Score System: Program Overview and Technical Protocol (Version 1.2). PNNL-22045 Rev 1.2, Pacific Northwest National Laboratory, Richland, Washington.

Wang, N., S. Goel, A. Makhmalbaf and N. Long. 2016. Development of Building Energy Asset Rating Using Stock Modeling in the USA. SN 1940-1493, Journal of Building Performance Simulation.