Can the Home Energy Score Serve as a Proxy for HERS? An Analysis of Energy Asset Ratings in San Diego County

November 2014

PREPARED FOR

San Diego Regional Energy Partnership

PREPARED BY

Center for Sustainable Energy





Headquarters 9325 Sky Park Court, Suite 100 San Diego, CA 92123 858.244.1177 www.energycenter.org Office Locations
San Diego
Los Angeles
Oakland

| This report was prepared for the San Diego Regional Energy Partnership by the Center for Sustainable Energy. The San Diego Regional Energy Partnership includes San Diego Gas & Electric, City of San Diego, County of San Diego, City of Chula Vista, San Diego Association of Governments and Port of San Diego. |
|--|
| |
| |
| |

Contents

| EXECUTIVE SUMMARY | 2 |
|---|----|
| INTRODUCTION TO ASSET RATINGS | |
| HERS | 2 |
| Home Energy Score | 7 |
| The value of asset ratings | |
| HOME ENERGY SCORE AND HERS ANALYSIS | 10 |
| Collecting the data | 10 |
| Producing Home Energy Scores | 12 |
| Using Home Energy Scores to predict HERS scores | 13 |
| Using Home Energy Score field measurements to predict HERS scores | 13 |
| Conclusions and caveats | 18 |



EXECUTIVE SUMMARY

An energy asset rating measures a building's energy efficiency performance based on its permanent equipment, design and construction. This type of rating allows for the comparison of energy efficiency across buildings while making standard assumptions about occupant behavior. An asset rating typically produces two outputs: a numerical score and a report that describes opportunities for energy efficiency upgrades. Asset ratings are useful to a variety of stakeholders, including homeowners and contractors engaged in energy efficiency upgrades; buyers, sellers, agents and lenders involved in real estate transactions; and utilities and local governments that wish to track energy performance improvements that result from energy efficiency programs.

In California, the primary asset rating system is the California Whole-House Home Energy Rating (HERS). Generating a HERS rating is a relatively complex endeavor, requiring diagnostic testing (i.e., blower door and duct blaster) and more than 50 inputs into the modeling software. In total, this requires four to six hours of work by a certified rater and costs approximately \$400-500.

An alternative asset rating is the Home Energy Score, which is predominantly used outside of California. The Home Energy Score tool is relatively basic, characterizing home efficiency on a 10-point scale. It typically requires a similar number of data inputs as a HERS rating, but does not require advanced diagnostic testing and can be completed by a qualified assessor in about one hour for a cost of approximately \$100-\$250.

Given the potential benefits of widespread asset ratings, the San Diego Regional Energy Partnership wanted to explore the viability of using the lower-cost, lighter-touch Home Energy Score and/or its inputs to predict a HERS score. The answer may also help inform the California Energy Commission as it considers the development of a "HERS Lite" rating system and seeks to better understand the technical basis of the Home Energy Score.

The Center for Sustainable Energy conducted this analysis by applying linear regression models and machine learning techniques to a sample of 255 homes in San Diego County. The results show that 10 key features derived from Home Energy Score field measurements can be used to predict a HERS score

within an accuracy range of \pm 27 points. Thus, while the Home Energy Score does not replace the HERS rating, it could serve as a cost-effective proxy for HERS scores in the San Diego region and help target limited resources where they are most needed. For instance, the Home Energy Score may be used to identify homes that warrant a more thorough evaluation with a HERS rating to plan energy efficiency improvements; i.e., homes that land on the high end of the proxy HERS scale may be worth the larger investment of an actual HERS rating.

It is important to note that this analysis did not examine the effectiveness of Home Energy Score or HERS in stimulating home energy upgrades. These conversion rates would likely be impacted by the full presentation of the Home Energy Score or HERS *report*, not just the numerical scores. Therefore, it would be useful to conduct additional research on the conversion rates from various asset ratings to energy efficiency upgrades. In addition, analysis using data from different climates and regions would indicate whether the conclusions presented here apply to areas beyond the San Diego region.



INTRODUCTION TO ASSET RATINGS

An energy asset rating measures a building's energy efficiency performance based on its permanent equipment, design and construction. This type of rating – more commonly applied to residential buildings than to nonresidential buildings – allows for the comparison of energy efficiency across buildings while making standard assumptions about occupant behavior. To produce a rating, a rater or assessor may collect information about the home's size, age, location, building characteristics, mechanical systems, appliances and lighting. The rater then enters this information into software to produce two outputs: a numerical score and a report that provides supplemental information and recommendations for energy efficiency upgrades.

In California, the primary asset rating system is the California Whole-House Home Energy Rating (HERS)¹, which is administered by the California Energy Commission.² Outside of California, an array of asset ratings has been developed by various state, regional and national entities.³ One of these systems is the Home Energy Score, developed by Lawrence Berkeley National Laboratory in collaboration with the U.S Department of Energy.

HERS

HERS scores homes on a scale of zero to 250, with a lower score indicating a more energy-efficient building. A home built to California's 2008 Building Energy Efficiency Standards (Title 24, Part 6) and no better than those standards would score a 100. A score of zero would indicate a net-zero-energy home.

¹ The whole-house rating was established in the 2009 update to the California Home Energy Rating System (HERS) regulations.

² Comprehensive Energy Efficiency Program for Existing Buildings, CEC, August 2012, p. xi (AB 758 Action Plan).

³ Rating Review Process, U.S. Department of Energy, June 2010.

A HERS rating requires more than 50 data inputs⁴ including:

- Climate zone
- Year of construction
- Number of stories
- Conditioned floor area
- Conditioned volume
- Number of bedrooms
- Ceiling and roof surface area
- Ceiling and roof orientation
- Ceiling and roof tilt
- Ceiling framing assembly
- · Roof pitch
- Roof deck type
- Roof deck solar reflectance and emittance
- Roof deck above deck insulation
- Roof deck above deck mass
- Roof deck framing members insulation
- Roof deck insulation below deck
- Roof deck radiant barrier
- Overhang parameters
- Side-fin parameters
- Attic height
- Attic free ventilation area
- Attic fraction located high in attic
- Inter-zone surface type
- Inter-zone surface area
- Leakage rate (CFM)
- · Air retarding wrap
- Special air system specs

- Mechanical ventilation (CFM)
- Ventilation height difference
- Exterior walls surface area
- Thermostat type
- Duct location specs
- Duct insulation specs
- Duct/air handler leakage
- · Reduced infiltratation due to duct sealing
- Inter-zone ventilation
- HVAC system type
- Heating system type
- Heating system efficiency
- Cooling system specs
- Free ventilation area
- Exterior wall orientation
- UIMC and areas if high mass
- Basement walls surface area
- Raised floors surface area
- Slab area and type
- Exposed slab area %
- Door area
- Door U-factor
- Door orientation
- Fenestration area
- Fenestration orientation
- Fenestration SHGC
- Fenestration U-factor
- Water heating specs

The most time-consuming of these data points include building leakage and duct leakage, which are measured using diagnostic tests with a blower door and duct blaster, respectively. To generate the HERS score and report, these data are entered into EnergyPro, an energy modeling software that requires a significant amount of training to use properly.

⁴ Home Energy Rating System Technical Manual, California Energy Commission, December 2008.

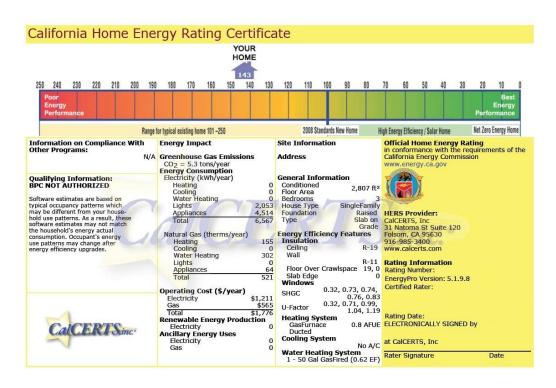


Figure 1: Page 2 of HERS Rating report; the HERS score is indicated at the top.

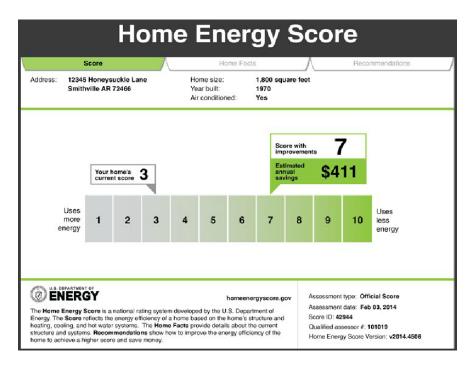


Figure 2: Page 1 of Home Energy Score report; the score is indicated on the left half of the page.

Home Energy Score

The Home Energy Score tool was designed to be a "low-cost, high-value assessment [that] can be provided as a stand-alone service or as an add-on to a home inspection or comprehensive energy assessment." The Home Energy Score rates homes on a scale of one to 10, with a higher score signifying higher energy efficiency. It does not use California building standards as a reference in the scoring system.

A Home Energy Score also requires more than 50 data inputs, although the blower door test is optional and one cannot enter duct blaster measurements at all.⁶ The data are entered into an online interface to generate the Home Energy Score and report. These inputs include:

- Address
- City
- State
- ZIP code
- House year built
- Townhouse or duplex
- Direction faced by front of house
- Conditioned floor area (sq. ft.)
- House # of bedrooms
- House # stories above grade
- House average ceiling height(ft.)
- Attic or ceiling type
- Attic floor insulation
- Foundation type
- Foundation insulation over basement or crawlspace
- Position of unit
- Blower door air leakage (CFM50)
- Has the house been professionally sealed
- Wall characteristics
- Wall construction
- Wall exterior finish
- Wall insulation
- Duct #1 location
- Duct #2 location
- Duct #3 location
- Heating system type
- Heating system AFUE or HSPF

- Heating system year installed
- Cooling system type
- Cooling system efficiency
- Cooling SEER or EER
- Cooling year installed
- Window panes
- Window area (front, back, right, left)
- Window frame type
- Window glazing type
- Window U-factor
- Window SHGC
- Skylights # panes
- Skylights total area (sq. ft.)
- Skylights frame type
- Skylights glazing type
- Skylights U-factor
- Skylights SHGC
- Roof construction
- Roof exterior finish
- Roof insulation level
- Roof color
- Hot water heating type
- Hot water energy factor
- Hot water system efficiency
- Hot water year installed
- Assessment type
- Assessment date

⁵ Home Energy Scoring Tool, Lawrence Berkeley National Laboratory.

⁶ Home Energy Scoring Tool Data Collection Sheet, U.S. Department of Energy.

Table 1: Comparing HERS Score and Home Energy Score

| | HERS | Home Energy Score | |
|---|---|---|--|
| Scoring system | 0-250 (lower is better) | 1-10 (higher is better) | |
| Diagnostic testing required | Blower door, duct blaster (if home is in bad condition or test cannot be conducted, default values may be used) | Blower door optional; cannot enter duct blaster measurement into Home Energy Score | |
| Number of data inputs | 50+ | 50+ | |
| Field work + data input time | 4-6 hours | 1.25 hours | |
| Estimated cost per rating | \$400-\$500 ⁷ | \$100-\$250 | |
| Required qualifications of rater/assessor | Must demonstrate competence in all areas of Section 1673(a)(1) of the HERS regulations (includes diagnostic testing and use of energy modeling software) to become a certified California Whole-House Home Energy Rater | As of Nov. 2014: Must be certified as a Building Performance Institute (BPI) Building Analyst or a Residential Energy Services Network (RESNET) HERS Rater to take Home Energy Score exam Expected in Jan. 2015: ASHI- and InterNACHI-certified home inspectors and NATE-certified HVAC technicians can take exam | |

The value of asset ratings

Asset ratings are an important tool in achieving energy efficiency goals on the federal, state, local and personal household levels. The Energy Commission states, "Ratings can be a powerful tool to communicate the energy assets of a property and can educate and motivate consumers to take action on an upgrade project." Common circumstances in which asset ratings are valuable include:

1. To inform homeowners and contractors engaged in home energy upgrades

Before a whole-house upgrade takes place, homeowners and contractors often use asset ratings to prioritize measures and plan the scope of work. For many projects that go through Energy Upgrade California[™] Home Upgrade, the modeling required to generate a HERS rating also is used to calculate rebate amounts – further assisting the selection of measures to include in the project. At the end of an upgrade, a homeowner or contractor may request another asset rating to be performed to officially verify the impact of the improvements.

⁷ The average pre-rebate cost of pre-upgrade HERS ratings conducted through the San Diego Region HERS Rating Rebate Program was \$492. The \$400 rebate may have inflated these costs.

⁸ Comprehensive Energy Efficiency Program for Existing Buildings Scoping Report, California Energy Commission, August 2012.

2. To inform buyers, sellers, agents and lenders involved in real estate transactions

The real estate community is increasingly valuing energy-efficient homes and demanding third-party verified information about the efficiency of homes. Home sellers want to market their homes as energy efficient or "green" to earn a premium for the energy upgrades they've invested in. Home buyers want to know they will end up in homes with reasonable utility bills. Some multiple listing services (MLS) now provide the option to enter asset ratings, either as a searchable data field or simply by uploading the report as an attachment. Lenders can use asset ratings to qualify homes for energy-efficient mortgages. A few local jurisdictions are exploring requirements for conducting asset ratings at time of sale.

3. To inform energy savings for utility and local government programs

Asset ratings can be used to track improvements in residential energy performance that result from direct install programs, financing or incentive programs, or time-of-sale requirements. (Although it should be noted that this is not the same as tracking improvements in actual energy consumption, which is impacted by occupant behavior and weather.)



HOME ENERGY SCORE AND HERS ANALYSIS

Given the potential benefits of widespread asset ratings and the fact that HERS is currently the only recognized asset rating system in California, the San Diego Regional Energy Partnership (SDREP) wanted to explore the viability of using the lower-cost, lighter-touch Home Energy Score and/or its inputs to predict a HERS score. The Center for Sustainable Energy (CSE) conducted this analysis using data from homes that received HERS ratings through a rebate program conducted by CSE for the SDREP. Because the data required to generate a Home Energy Score is a subset of the data collected for a HERS rating, we were able to generate Home Energy Scores for each of these homes. We then used linear regression models and machine learning techniques to answer the following questions.

- Can the Home Energy Score by itself predict a HERS score?
- Can the Home Energy Score field measurements, or values derived from the Home Energy Score field measurements, predict a HERS score?

The answer to these questions also may help inform the Energy Commission as it considers the development of a "HERS Lite" rating system and seeks to "better understand the technical basis of the Home Energy Score, and the extent to which the differences in HEScore system deviate from outcomes consistent with California energy policy goals for calculations and ratings."9

Collecting the data

The data used in this analysis were drawn from HERS ratings conducted on single-family homes in San Diego County from April 2013 to May 2014 (Figure 3). In all, 257 viable projects were included in the study, with 163 households in climate zone 7, 92 in climate zone 10 and two in climate zone 14. The two climate zone 14 households were excluded from the HERS prediction model due to insufficient sample size. Individual

⁹ Comprehensive Energy Efficiency Program for Existing Buildings Scoping Report, California Energy Commission, August 2012.

household records were recorded in the .bld file format for use in the building energy analysis software EnergyPro by EnergySoft. The total sample size of .bld files used in the model was 255 and represented the coastal and inland regions of San Diego. The median actual HERS score of the distribution was 155. The range of the HERS sample ranged from 92 to 313.

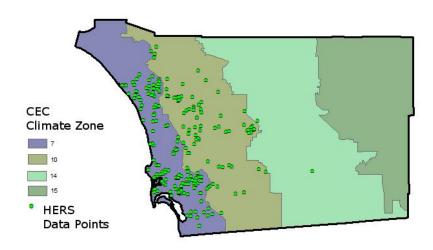


Figure 3: Spatial distribution of residential data points used for Home Energy Score/HERS analysis.

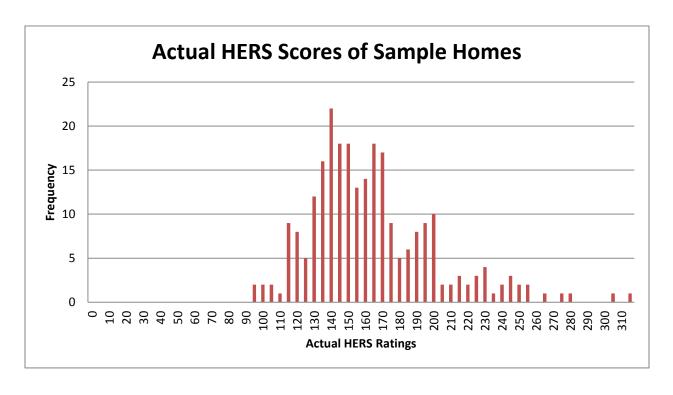


Figure 4: The distribution of actual HERS scores from 255 San Diego County homes.

Producing Home Energy Scores

EnergyPro was used to produce both the HERS scores and Home Energy Scores (Figure 5). We created a script in the programming language Python to batch process .bld files into the Home Energy Score API to produce Home Energy Scores. Many of the HERS ratings were also stored in PDF certificates. We built a PDF scraper in Python to extract this information. We then calculated Home Energy Scores twice: using actual blower door measurements of building leakage and mimicking the absence of a blower door test (i.e., the standard Home Energy Score process). This was done to test the added value of the optional blower door test in the Home Energy Score process.

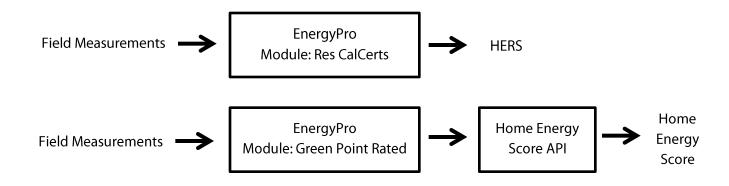


Figure 5: The HERS and Home Energy Score were produced using the Res CalCerts and Green Point Rated modules within EnergyPro, respectively.

For Home Energy Score assessments that do not include a blower door test, the Home Energy Score API calculates a default normalized leakage based upon building characteristics as seen in Equation 1.¹⁰ The normalized leakage (NL) can then be converted to CFM_{50} or specific leakage area (SLA) using ACEEE standards.

Equation 1:
$$NL = \exp(Floor\ Area\ x\ C_{floor\ area} + Building\ Height\ x\ C_{height} + (sealed = yes)\ x\ C_{sealed} + C_{vintage} + C_{IECC} + C_{foundation} + C_{duct})$$

Substituting actual blower door tests for calculated default values did not change the overall distribution of Home Energy Scores. The median of the actual Home Energy Scores is 8 for both inclusion and exclusion samples (Figure 6). For the blower door exclusion sample, 71 of the 256 samples scored a 10. The HERS curve seen in Figure 4 and the Home Energy Score curve in Figure 6 differ in shape with HERS ratings showing a median near the center of the 0-250 scale and Home Energy Scores in Figure 6 showing a median towards the right of the 1-10 scale at 8.

¹⁰ Home Energy Saver: Engineering Documentation, Lawrence Berkeley National Laboratory.

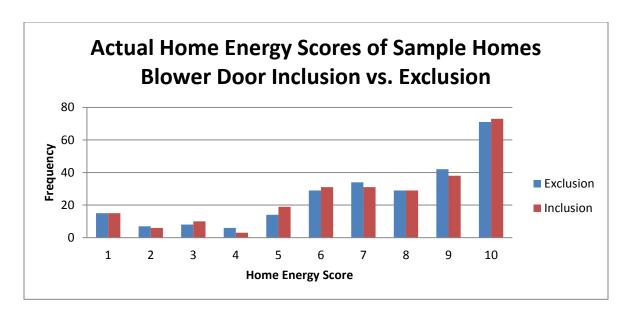


Figure 6: The distribution of actual Home Energy Scores from 255 San Diego County homes.

Using Home Energy Scores to predict HERS scores

The first model analyzed the explanatory power of just the Home Energy Score, without the Home Energy Score field measurements, in predicting the HERS score. This was done by creating a simple linear regression in the statistical modeling software R. This simple model is expressed in Equation 2.

Equation 2:
$$HERS = f(Home\ Energy\ Score)$$

Our findings show an adjusted R-squared value of 0.36, which indicates that a Home Energy Score can explain 36% of the variance in the HERS index using linear regression.

Using Home Energy Score field measurements to predict HERS scores

The next step was to run a model using various Home Energy Score field measurements (e.g., air conditioner SEER, window size, etc.) to predict HERS scores.

Making the raw field measurements more useful

Some Home Energy Score field measurements did not offer enough explanatory power on their own and needed to be processed before modeling. Examples include:

• The addresses of the homes were first geocoded and their corresponding Energy Commission climate zone was used as a categorical variable for the model.

- The location of HVAC ventilation, which can be in the crawlspace, attic, outdoors, etc., was categorized either conditioned space or unconditioned space.
- The roof, windows, slab and walls of every home often have multiple U-values and areas within each given category. For example one may have eight U-values and eight areas corresponding to eight walls in a home. To reduce multiple parameters, as well as the degrees of freedom in the model, we used the area weighted average calculation as seen in Equation 3.

Equation 3: Area Weighted Spec =
$$\frac{\sum_{i=1}^{n} Area_i \times Spec_i}{\sum_{i=1}^{n} Area_i}$$

In this equation, Area; is the area of the window, wall or roof panel of interest and Spec; corresponds to either the U-value or the solar heat gain coefficient (SHGC).

One challenge in dealing with window parameters is that a statistical model cannot derive full explanatory value from a list of window areas, solar heat gain coefficients (SHGCs) and orientation. Despite this complexity, including windows is critical to the modeling effort as they represent a key point of heat exchange in the building envelope, impacting overall efficiency through the interaction of orientation, size (area) and specification (U-factor, transmittance, solar heat gain and leakage). To capture this effect, the location, orientation and tilt values were run through the National Renewable Energy Lab's PVWatts API to calculate incident solar insolation on the window surface. Next, the solar heat gain value was applied. This process was automated using Python to quickly generate a proxy for the annual heat gain in BTU from every window in a given home.

Selecting the most relevant features

A key factor in the robustness of a model is how the variables, also known as features, are selected. It is best practice to select only features that have a statistically significant correlation with the dependent variable in order to reduce multicollinearity. There are different methods of rewarding model simplicity and punishing model complexity. This analysis used linear regression and three standard feature selection tools, Akaike information criterion (AIC), Bayesian information criterion (BIC) and Mallow's Cp (C_p), to identify the features with the highest explanatory power. The list of potential features from which we selected is in Table 2. Many of these features were composite features as explained in the previous section.

After running the linear regression, 10 out of 27 features were chosen for our model. When predicting a HERS score, the five most statistically significant features (noted in bold in Table 2) are the Energy Commission climate zone, the weighted roof area U-factor, the weighted west wall U-factor, conditioned floor area and the annual solar window heat gain. Of the top five significant features, four are composite features that took raw Home Energy Score data inputs and used additional formulas to add more explanatory value to the feature.

Table 2: Home Energy Score Features that Predict HERS Score

| List of all Home Energy Score features considered Some of these are raw Home Energy Score field measurements; some are composite features that applied formulas to raw Home Energy Score inputs to add more explanatory value | Identified as statistically significant in predicting HERS score | Predictive value Lower indicates more statistically significant |
|---|--|--|
| Energy Commission climate zone | ✓ | < 2 x 10 ⁻¹⁶ |
| Year built | | |
| Number of stories | ✓ | 1.39 x 10 ⁻² |
| Conditioned floor area | ✓ | 5.18 x 10 ⁻⁷ |
| Ceiling height | ✓ | 1.51 x x10 ⁻¹ |
| Number of bedrooms | | |
| Building leakage | | |
| Weighted foundation area U-factor | ✓ | 1.10 x 10 ⁻¹ |
| Weighted roof area U-factor | ✓ | 1.44 x 10 ⁻⁸ |
| North wall U-factor weighted area | | |
| East wall U-factor weighted area | | |
| South wall U-factor weighted area | | |
| West wall U-factor weighted area | ✓ | 6.57 x 10 ⁻⁸ |
| Annual solar window BTU (calculated with PVWatts API) | ✓ | 1.51 x 10 ⁻⁵ |
| North window weighted area | ✓ | 4.24 x 10 ⁻³ |
| East window weighted area | | |
| South window weighted area | | |
| West window weighted area | | |
| HVAC duct R-value | | |
| Air conditioner SEER | ✓ | 5.89 x 10 ⁻⁴ |
| Heating efficiency | | |
| Vent location | | |
| Space heating fuel source | | |
| Water heater energy factor (EF) | | |
| Water heater rated input | | |
| Water heater tank size | | |

Testing the accuracy of our predictions

We used both linear regressions and machine learning techniques to predict HERS scores. Linear regressions train a model to make future predictions based on past data – a strategy that performs well when trying to predict past outcomes and when seeking quick insight into the predictive value of Home Energy Scores vs. HERS scores. This methodology requires access to a large data set that is representative of the diversity found in the underlying population. When predicting values based on new data that is un- or under-represented in the base data set, the model may not perform as well.

The advantage of machine learning is that it allows one to gauge model performance based on new data that has not been trained on before, resembling real-world conditions. Using the "hold-out" method, one must hold out data from the training set and designate this data as test data. The model then uses sophisticated pattern recognition that teaches itself using this test data. Machine learning offers more accurate error estimations that cannot be identified with linear regressions.

Linear Regression

We used a multivariate linear regression model represented by the following equation.

Equation 4: HERS = f(10 statistically significant Home Energy Score features in Table 2)

The results, seen in Figure 7, produced an adjusted R-squared of 0.60. This indicates that by including the Home Energy Score features into the linear regression, we were able to account for 60% of the HERS variance – or 25% more than using the Home Energy Score alone. One key check in validating linear regressions is to observe if any variables correlate with one another that can lead to invalid results; this is called multicollinearity. After feature selection, no significant multicollinearity is observed in this analysis; the results are valid by this check.¹¹

Machine Learning

We built a predictive model based on 190 samples and then tested the model on the remaining "hold out" data of 65 homes. We used the Random Forest algorithm to test the limits of predicting the HERS value and identify real-world error. The algorithm was run 300 times to approximate the root mean square error (RMSE) of the actual HERS scores and the predicted HERS scores (Equation 5). The RMSE is a standard metric used to compare the accuracy of different models; it punishes large prediction errors greatly due to the squared term.

Equation 5:
$$RMSE_{HERS} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(HERS_{actual} - HERS_{predicted})^2}$$

The results, seen in Figure 8, show an average RMSE of \pm 27.06. This means that when predicting HERS scores using San Diego-region residential building characteristic data that is routinely collected during a Home Energy Score assessment, one can expect accuracy within \pm 27.06 points of the actual HERS score on a scale of 0-250. Interestingly, by taking the additional step of completing a blower door test, HERS prediction accuracy showed no improvement with an average RMSE of \pm 27.08 over 300 runs.

Page 16

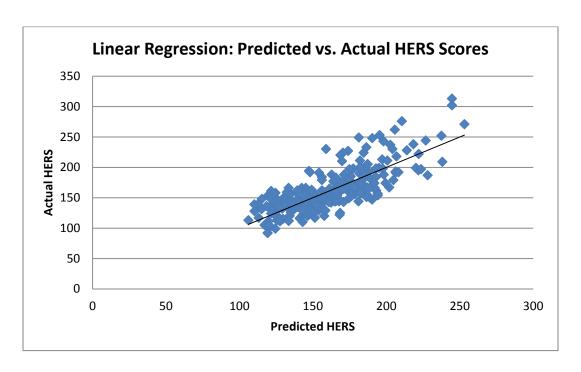


Figure 7: This figure shows the multivariate linear regression with an R-squared of 0.62 and adjusted R-squared of 0.60.

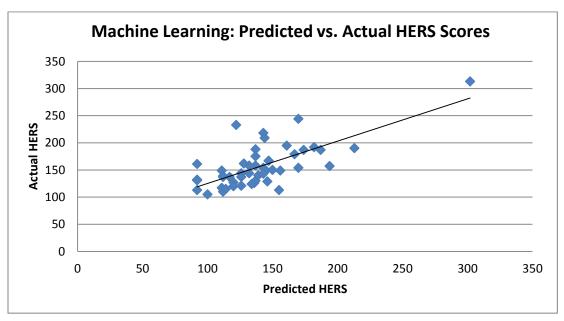


Figure 8: Using the Random Forest algorithm and a test size of 300 this chart shows a single machine learning run with an RMSE of 23.18. These algorithm runs were repeated 300 times and averaged to a representative RMSE of \pm 27.08.

Conclusions and caveats

This research suggests that by using machine learning techniques, 10 key features derived from Home Energy Score field measurements can be used to predict a HERS score within an accuracy range of \pm 27 points. Thus, while the Home Energy Score does not replace the HERS rating, it could serve as a cost-effective proxy for HERS scores in the San Diego region and help target limited resources where they are most needed. For instance, the Home Energy Score may be used to identify homes that warrant a more thorough evaluation with a HERS rating to plan energy efficiency improvements; i.e., homes that land on the high end of the proxy HERS scale may be worth the larger investment of an actual HERS rating.

It is important to note three caveats to this analysis and areas where further research is needed.

First, this model was tuned to the San Diego region, specifically the coastal and inland climate zones for which there was adequate sample size. As discussed previously, statistical regression and machine learning techniques are limited by the availability and representativeness of test data. This model as currently specified may or may not perform as well in a different region or climate. Further analysis using HERS data from different climates and regions is required before substantiated interregional HERS estimates can be made in different climates.

Second, 255 samples were used in this analysis; as the sample size increases, the features selected to be most relevant in prediction may change as more data is available. A larger sample size should improve prediction accuracy and R-squared values in the future.

Finally, we did not examine the relative effectiveness of Home Energy Score and HERS in moving consumers from asset rating to energy efficiency upgrades. Both HERS and Home Energy Score provide a report, which include supplemental information and recommendations for energy upgrades, in addition to the numerical score. While the numerical score by itself is highly valuable for some purposes (for example, for tracking relative energy performance of buildings or prompting homebuyers to value energy-efficient homes), the larger report may be useful for policies or programs designed to stimulate homeowners to undertake energy upgrades. Thus, it would be useful to compare conversion rates to analyze the impact of the full HERS and Home Energy Score experience on consumer decision-making in regard to home energy upgrades.

