Assessment of the U.S. Department of Energy's Home Energy Scoring Tool: 2014 Update

Noel Merket and Mike Heaney

NREL is a national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy
Operated by the Alliance for Sustainable Energy, LLC

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Technical Report
NREL/TP-5500-62263
August 2014

Contract No. DE-AC36-08GO28308

Assessment of the U.S. Department of Energy's Home Energy Scoring Tool: 2014 Update

Noel Merket and Mike Heaney

Prepared under Task No. BE4R.0601

NREL is a national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy
Operated by the Alliance for Sustainable Energy, LLC

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

National Renewable Energy Laboratory
15013 Denver West Parkway
Golden, CO 80401
303-275-3000 • www.nrel.gov

Technical Report
NREL/TP-5500-62263
August 2014

Contract No. DE-AC36-08GO28308

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Available electronically at http://www.osti.gov/scitech

Available for a processing fee to U.S. Department of Energy and its contractors, in paper, from:

U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831-0062
phone: 865.576.8401
fax: 865.576.5728
email: mailto:reports@adonis.osti.gov

Available for sale to the public, in paper, from:

U.S. Department of Commerce
National Technical Information Service
5285 Port Royal Road
Springfield, VA 22161
phone: 800.553.6847
fax: 703.605.6900
email: orders@ntis.fedworld.gov
online ordering: http://www.ntis.gov/help/ordermethods.aspx

Cover Photos: (left to right) photo by Pat Corkery, NREL 16416, photo from SunEdison, NREL 17423, photo by Pat Corkery, NREL 16560, photo by Dennis Schroeder, NREL 17613, photo by Dean Armstrong, NREL 17436, photo by Pat Corkery, NREL 17721.

NREL prints on paper that contains recycled content.

Acknowledgments

This work was funded by the U.S. Department of Energy (DOE) Building Technologies Program. The authors wish to thank David Lee (DOE Team Leader, Residential Buildings) for his continued support. We would also like to thank Norm Bourassa and Leo Rainer of Lawrence Berkley National Laboratory (LBNL), and Joan Glickman and Glenn Dickey of DOE.[1]

[1] In acknowledging individuals and organizations we do not mean to imply their endorsement of the research results. Our intention is simply to acknowledge their contributions and thank them.

Executive Summary

In 2012 the National Renewable Energy Laboratory (NREL) conducted a series of assessments of the DOE proposed Home Energy Scoring Tool (HEScore). The primary objectives were to: (1) assess the accuracy of HEScore as it was being developed; and (2) provide information that was useful to DOE program managers and the HEScore development team at LBNL. The work was documented by Roberts et al. (2012).

This report is an update to the 2012 work, assessing the 2014 HEScore release. The analyses here follow the same methods as those in the previous report to facilitate comparison. However, the analysis in this updated report focuses on energy use only. A comparison of the software-assigned score to the score a building would receive based on the actual energy use is omitted. The scoring methodology for the 2014 version was updated to include end uses only (space conditioning and hot water) that were affected by the inputs to the model. Other end uses (lighting and miscellaneous electric loads) are occupant driven and are not used in the asset score calculation. Because end uses cannot be separated from gross electricity and natural gas bills, a score could not be calculated for the measured energy use of the houses in this sample for comparison.

Comparison of Predicted and Measured Energy Uses

Predictions of electricity and natural gas consumption were compared with weather-normalized utility billing data for a mixture of newer and older homes located in Oregon, Wisconsin, Minnesota, North Carolina and Texas (see Table 1 and Table 2).[2] The 877 electricity use comparisons and 548 natural gas use comparisons yielded the following:

- HEScore 2014 overpredicted electricity use by a median of 2%.
- HEScore 2014 underpredicted gas use by a median of 10%.
- There was no significant change in the variability from HEScore 2012 to HEScore 2014 (as measured by the standard deviation of the differences) for electricity use.
- For natural gas use there was a 9% reduction in variability from HEScore 2012 to HEScore 2014.

[2] A limitation of this approach is that the HEScore assesses the performance of the energy-related assets of a home under typical operating conditions; utility billing data reflect the performance of the energy-related assets of a home under actual operating conditions. The uncertainty associated with this limitation is addressed in later sections of the report.

Table 1. Statistical Summary of Differences Between Predicted and Measured Electric Energy Use (Predicted kWh – Measured kWh)

	HEScore v2012	HEScore v2014
Number of Observations	877	877
Mean Measured[3]	10659	10769
Mean Predicted	10340	10526
Mean Difference	–319	–243
Median Difference	30	177
Mean Percent Difference	13%	13%
Median Percent Difference	0%	2%
Median Absolute Difference	2382	2253
Median Absolute Percent Difference	24%	22%
Percent of Homes < ± 25% Different	54%	55%
Percent of Homes < ± 50% Different	80%	81%

Table 2. Statistical Summary of Differences Between Predicted and Measured Natural Gas Use (Predicted Therms – Measured Therms)

	HEScore v2012	HEScore v2014
Number of Observations	548	548
Mean Measured[4]	863	837
Mean Predicted	828	754
Mean Difference	–35	–82
Median Difference	–41	–71
Mean Percent Difference	1%	–6%
Median Percent Difference	–6%	–10%
Median Absolute Difference	185	181
Median Absolute Percent Difference	24%	25%
Percent of Homes < ± 25% Different	52%	50%
Percent of Homes < ± 50% Different	84%	84%

Statistical Modeling

We used multiple linear regression analysis to develop empirical models, using energy use differences as the dependent variable. This enabled us to identify potential issues driving differences between HEScore-predicted energy uses and measured energy uses. The number of bedrooms and the number of stories above grade contributed significantly to the difference between predicted and actual electric energy consumption. This may be due in part to assumptions about occupancy, base loads, and lighting in HEScore. Contributors to the difference between predicted and measured natural gas use include the number of heating degree-days, window area, and heating system efficiency. The statistical model indicates that HEScore is over- or underresponsive to these features to some degree. It is important to note that

[3] The measured energy use differs between versions because the bills were normalized to Typical Meteorological Year 3 (TMY3) for HEScore 2014 and TMY2 for HEScore 2012 to match the simulated weather.
[4] The measured energy use differs between versions because the bills were normalized to TMY3 for HEScore 2014 and TMY2 for HEScore 2012 to match the simulated weather.

the statistical model applies only to the current dataset. Electric furnaces and electric baseboards were significant contributors to delta electricity for HEScore 2012, but not for HEScore 2014. The modeling has apparently improved for this heating type.

Whole-House Leakage Sensitivity Analysis

HEScore accepts either a quantitative measurement of whole-house leakage using a blower door or a qualitative assessment of whether the home has been air sealed. When the qualitative air sealing question is answered instead of providing a blower door measurement, an algorithm using data from LBNL's Residential Diagnostics Database[5] is used to estimate the whole-house leakage using other known inputs about the house as well as the air sealing question. In HEScore 2014 this infiltration estimation model saw significant updates. This analysis shows that updated model and determines whether it can reliably replace a blower door measurement when used to calculate a Home Energy Score.

In the data acquired from the HEScore 2012 national launch, blower door measurements were performed for 1,489 homes. Additionally the Building America Field Data Repository has blower door data for 1,075 homes. NREL reran these homes through HEScore 2014 three times using different inputs for whole-house air leakage:

- The blower door data (quantitative input)
- The qualitative assessment of *Sealed*
- The qualitative assessment of *Unsealed*.

For the HEScore data, when compared to the predictions stemming from quantitative input, the average source energy use decreased by 2 MMBtu/yr (1.1%) when the sealed qualitative input was used and increased by 5 MMBtu/yr (2.4%) when the unsealed qualitative input was used. For the Building America Field Data Repository data, the average source energy decreased 1 MMBtu/yr (0.7%) when the sealed qualitative input was used and increased by 2 MMBtu/yr (1.2%) when the unsealed qualitative input was used. These estimates reflect the average predicted energy for the entire dataset. From this we concluded that the infiltration measurement, when replaced by a qualitative assessment, has a relatively small effect on the overall average predicted energy. In other words, the qualitative assessment is doing a reasonably good job of estimating leakage. Infiltration (envelope leakage) does have a significant effect on individual home energy use based on the statistical analysis in Section 3 of this report. On average the infiltration assumption for sealed is more airtight than a blower door measurement and the assumption for unsealed is less airtight than a blower door measurement, which is logical because the blower door test would have been done on houses that were both sealed and unsealed.

When evaluated in terms of the score, using the qualitative infiltration assessment produced scores within ±1 bin of the score generated using a blower door measurement for 95% of the homes investigated, whether or not the house was represented as air sealed. This indicates that the infiltration assumptions from the qualitative assessment are adequate, especially when considering the resolution of the score.

[5] http://resdb.lbl.gov/

Nomenclature

ACH50	Air Changes per Hour at 50 Pascals of Pressure Differential
AFUE	Annual Fuel Utilization Efficiency
API	Application Programming Interface
CDD	Cooling Degree Day
CFM25	Cubic Feet per Minute at 25 Pascals of Pressure Differential
CFM50	Cubic Feet per Minute at 50 Pascals of Pressure Differential
DF	Degrees of Freedom
DOE	U.S. Department of Energy
BAFDR	Building America Field Data Repository
HDD	Heating Degree Day
HEScore	Home Energy Scoring Tool
LBNL	Lawrence Berkeley National Laboratory
MLR	Multiple Linear Regression
NREL	National Renewable Energy Laboratory
SHGC	Solar Heat Gain Coefficient
TMY	Typical Meteorological Year

Table of Contents

List of Figures

List of Tables

1 Introduction

In 2012 the National Renewable Energy Laboratory (NREL) conducted a series of assessments of the U.S. Department of Energy's (DOE) proposed Home Energy Scoring Tool (HEScore). The primary objectives were to: (1) assess the accuracy of HEScore as it was being developed, and (2) provide information that was useful to DOE program managers and the HEScore development team at Lawrence Berkeley National Laboratory (LBNL). The work was documented by Roberts et al. (2012).

This report is an update to the 2012 work, assessing the 2014 HEScore release. This report provides updated results for following 2012 analysis activities:

- Comparison of predicted energy uses to measured energy uses
- Statistical modeling
- Whole-house leakage sensitivity analysis.

This report is an assessment of the 2014[6] HEScore release. Comparisons to the 2012 version are from the data used in the previous report. Only the houses in both datasets are used in the comparison.

1.1 Home Energy Scoring Tool Updates

HEScore 2014 is a re-architecture of the HEScore software backend and application programming interface (API) with the goal of modernizing a code base, parts of which date back to the late 1990s. The goal of the re-architecture was to replicate the energy modeling of the legacy software and provide a modular code base that will facilitate modeling improvements in the future. Nevertheless, changes were made that modestly affect the energy predictions from the model. Among them are:

- Building simulations now use Typical Meteorological Year 3 (TMY3) weather data instead of TMY2.
- The site-to-source energy multipliers for each fuel type were updated with data from ENERGY STAR® Portfolio Manager (ENERGY STAR 2013).[7]
- The air leakage model now uses the 2012 version of the Residential Diagnostics Database.
- Duct efficiency is calculated with an hourly DOE-2 function and updated to account for regain correctly.
- The hot water draw assumptions were revised to use current Building America House Simulation Protocols.
- Appliance efficiency defaults were updated. Upgrade efficiencies were updated to the latest ENERGY STAR criteria.

[6] The version of HEScore used was 2014.4693. The simulations were performed on February 20, 2014.

[7] HEScore 2014 electricity and natural gas source multipliers are 3.14 and 1.05, respectively; HEScore 2012 multipliers were 3.365 and 1.092.

- Miscellaneous electric load assumptions were updated to use the latest Building America House Simulation Protocols.
- Boiler electrical energy use assumptions and the combined boiler/water heater model were corrected.
- Models were improved for electric baseboards, electric furnaces, and room air conditioners.
- Basement regain modeling was improved.
- Fuel price and emissions data were updated.

A detailed release history is available on the Home Energy Scoring Tool website.[8]

1.2 Home Energy Score National Launch

In early 2011 DOE conducted a pilot of the Home Energy Score with 10 agencies, spread throughout the United States, that volunteered to test the concept. The agencies conducted dozens to hundreds of home assessments, entered data collected into the pilot version of the Home Energy Scoring Tool Web interface, and generated Home Energy Score labels. The pilot participants provided feedback to DOE on several aspects of the proposed program (homeowner interest and acceptance of the score, likelihood of stimulating retrofit activities, ease of use, accuracy, etc.). In 2012, incorporating the feedback from the pilots, HEScore was launched nationally and many partners have been using it in their programs across the country.

More than 4,000 homes were scored during the first several months of the version 2012 national launch (see Figure 1 for geographic locations). The data collected and results generated are stored in a database that is accessible by the HEScore development team at LBNL. These "session" data were provided to NREL for use in this assessment. Of particular interest was the subset of the pilot homes for which blower door tests were conducted to assess whole-house leakage. These data were used to conduct the sensitivity analysis discussed in Section 4.

[8] https://sites.google.com/a/lbl.gov/hes-public/home-energy-scoring-tool/release-history

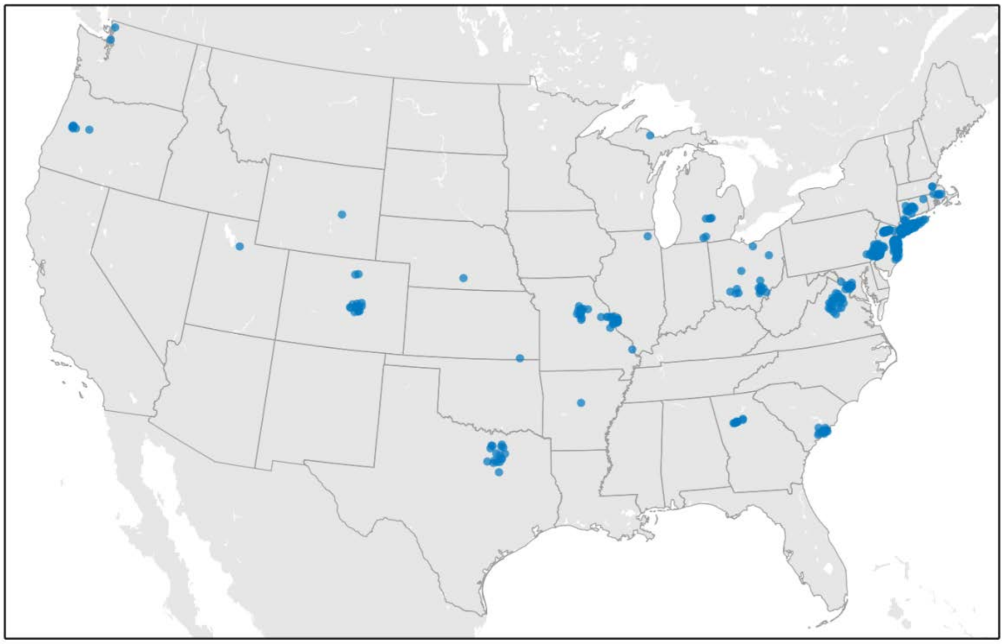

Figure 1. Home Energy Score pilot test locations

1.3 Building America Field Data Repository

The Building America Field Data Repository (BAFDR) contains historical energy audit data combined with utility bills. It was used to compare HEScore predicted energy use to actual energy use. The subset of the BAFDR used in this report is the same that was used by Roberts et al. (2012). Figure 2 shows a map of the data locations and relative number of houses used.

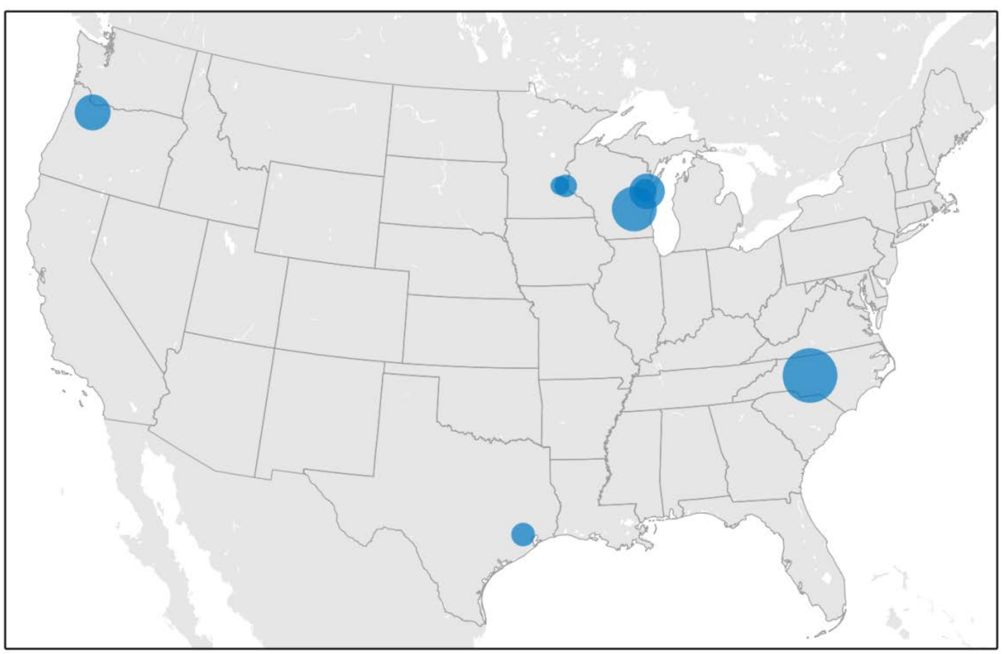

Figure 2. BAFDR dataset locations

The BAFDR, data sources, and data translation are discussed in more detail in Appendix A.

1.4 Overview of Approach

Predicted energy uses from the HEScore were compared to measured energy uses (i.e., weather-normalized utility billing data). Roberts et al. (2012) compared predicted energy uses from two other residential energy analysis tools to measured energy uses. In this assessment, the current version of HEScore is compared to the previous version. NREL's BAFDR and supporting translation algorithms were used to conduct these comparative analyses. The BAFDR is discussed in more detail in Appendix A. The results of the comparative analyses are presented and discussed in Section 2.

Multivariate linear models of measured energy use and of the residuals between predicted and measured use were developed to examine the impacts of HEScore inputs. These models inform potential changes to the software that may improve agreement between predictions and measurements. Results of the statistical modeling are presented in Section 3.

The HEScore input structure allows either a qualitative assessment or a quantitative measurement of whole-house air leakage. A question that is important to DOE is whether to require blower-door measurement as part of the Home Energy Score assessment process (currently an optional input to the tool). NREL leveraged data collected during the first few months of the Home Energy Score national launch to examine the sensitivity to using quantitative versus qualitative input in HEScore. This analysis is described in Section 4.

2 Comparisons to Previous Version and to Measured Data

Data from the BAFDR were programmatically mapped to HEScore. Some homes in the BAFDR were not included in these comparative analyses. Homes were programmatically excluded for a variety of reasons: missing utility billing data, poor data quality, or known asset features that cannot be modeled in HEScore.

Data from the BAFDR were mapped to the 2014 release of LBNL's HEScore. These simulation data were batched and programmatically submitted to the API used by the HEScore user interface. The results returned by the API were collected into a database. The process of mapping BAFDR data to HEScore inputs is detailed in Appendix B.

To facilitate comparison between the 2012 and 2014 versions of HEScore, only homes that were simulated in the 2012 version were simulated in the 2014 version. This provides a more direct comparison that is not affected by additional datasets that confound the results.

2.1 Current Predictions Versus Previous Version

Comparing the energy predictions of the new version of the HEScore software against the old version provides a method to verify that the software modifications are performing as expected and to quantify the change. It does not speak to the accuracy of the software prediction, which is considered in Section 2.2.

Figure 3 shows the comparison of predicted electricity use between versions. Figure 4 shows the comparison of predicted natural gas use between versions.

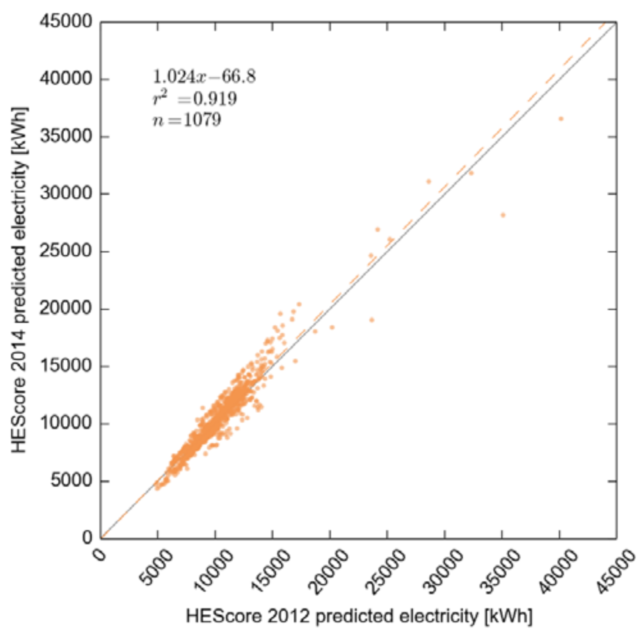

Figure 3. Site electricity use as predicted by HEScore 2014 versus HEScore 2012

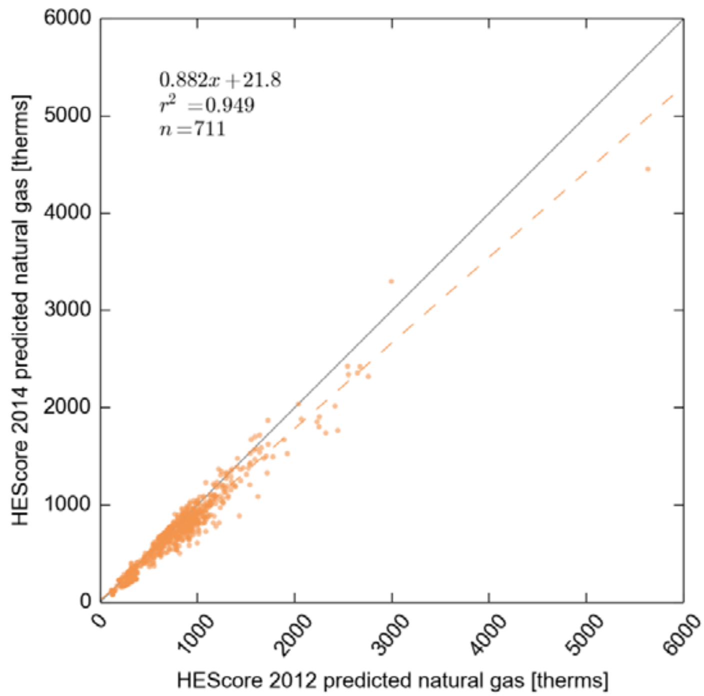

Figure 4. Site natural gas use as predicted by HEScore 2014 versus HEScore 2012

Average predicted site electricity use remains mostly unchanged between versions, although there is a very slight positive shift in higher predicted electricity-using homes in HEScore 2014. The mean electric energy use accordingly increases slightly. Site natural gas use has a noticeable negative shift, meaning that in HEScore 2014 less natural gas is predicted to be used for the same houses compared to HEScore 2012. A summary of the regression statistics is shown in Table 3.

Table 3. Statistical Summary of Predicted Electricity and Natural Gas Use Between HEScore 2012 and HEScore 2014

	Electricity (kWh)	Natural Gas (therm)
Number of Observations	1079	711
Mean HEScore 2012 Prediction	10320	761
Mean HEScore 2014 Prediction	10496	693
Slope of Regression	1.024	0.882
Intercept of Regression	–66.8	21.8
R^2 of Regression	0.919	0.949

Given that the purpose of the HEScore 2014 software update was an internal architecture overhaul with few intentional changes to energy modeling, replicating the energy results of the previous version indicates that the updates did not compromise the modeling assumptions or methods. With that understanding, the very comparable modeling results are logical and encouraging.

2.2 Predicted Versus Measured Energy Uses

Figure 5 shows HEScore-predicted site versus weather-normalized measured site electricity use. HEScore 2012 tends to underpredict homes with high measured electricity use and overpredict electricity use in homes with low measured use. HEScore 2014 improves the prediction. HEScore 2014 still underpredicts homes with high measured electricity use, but the pattern is less distinct than in the previous version. HEScore uses standard occupancy assumptions as defined in the Building America House Simulation Protocols; thus, it would not be expected to respond to unusually low or high energy use. The points to the far right of the graph, well below the line of perfect agreement, are likely homes with electrical loads that are not considered in the asset assessment: swimming pools, hot tubs, aquariums, waterbeds, second refrigerators, etc. Information about these end uses is not readily available in the BAFDR.

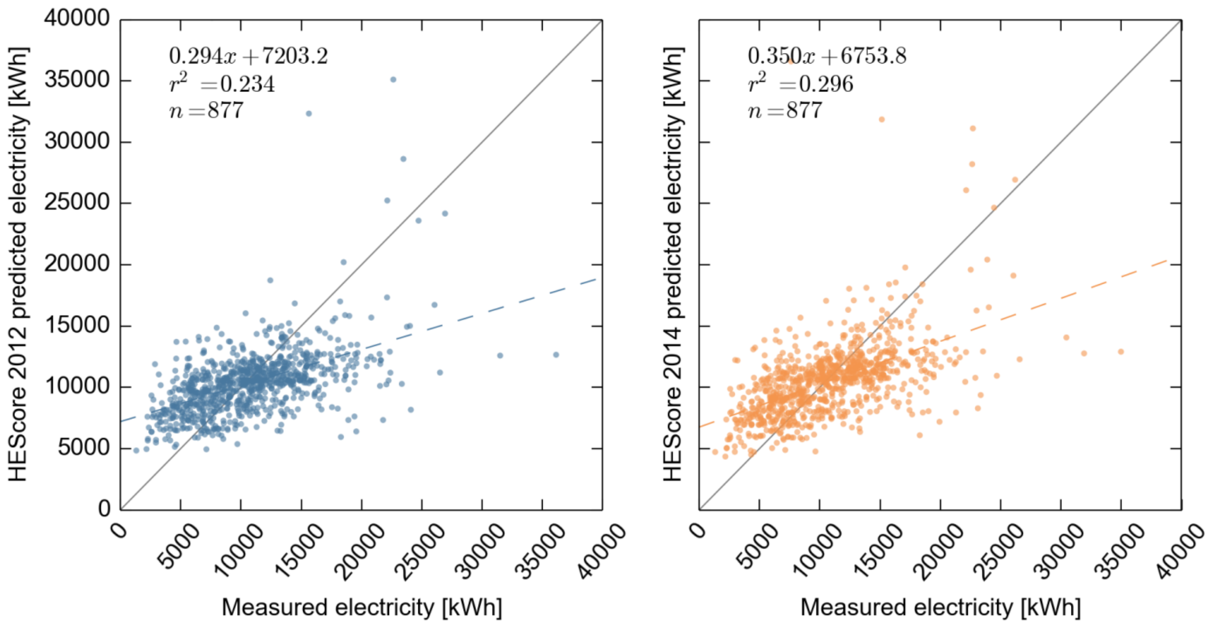

Figure 5. HEScore-predicted site versus weather-normalized measured site electricity use

Figure 6 shows the distribution of differences between HEScore-predicted site weather-normalized measured site electricity use. The distribution is asymmetrical, with a slight negative bias. Again, this is expected because HEScore does not account for extraordinary electricity end uses (e.g., swimming pool pumps) and the fact that energy use distributions are not normal: they are bounded by zero, but there is not a bound at the upper limit.

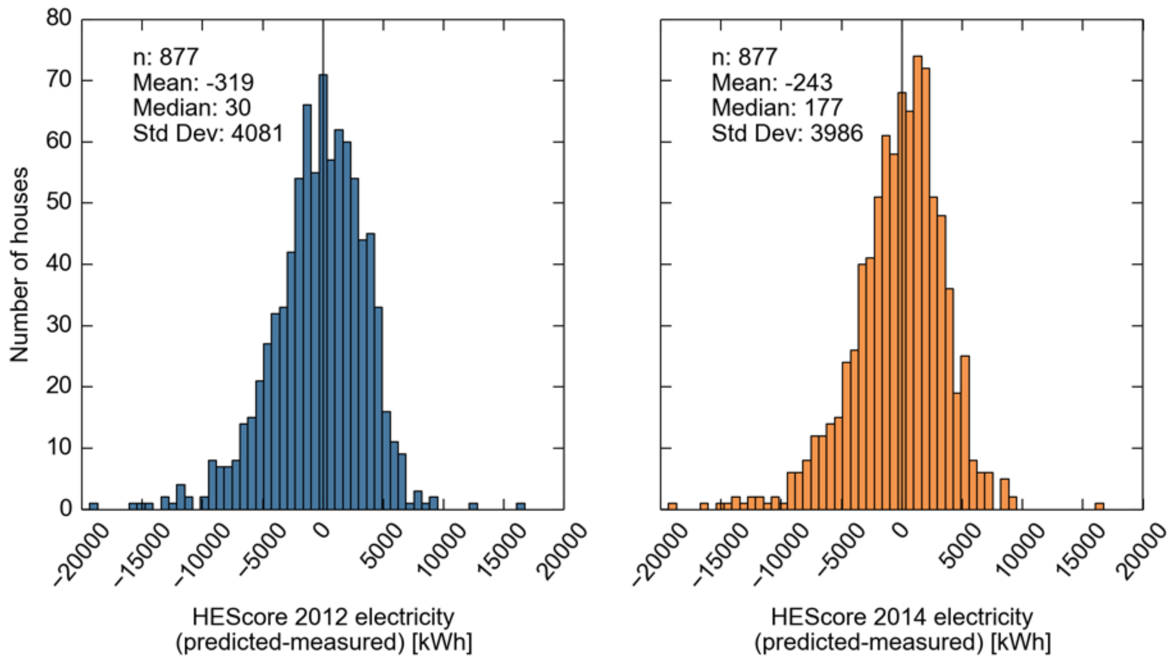

Figure 6. Distribution of differences between HEScore-predicted and measured site electricity use

Table 4 summarizes the differences between predicted and weather-normalized measured electricity uses. The distributions are not exactly normal, but generally resemble a normal distribution. An F-test using the standard deviations indicates that the variance in the HEScore 2014 electricity use (predicted – measured) is not significantly different from HEScore 2012.

Table 4. Statistical Summary of Differences Between Predicted and Weather-Normalized Measured Electricity Use (Predicted kWh – Measured kWh)[9]

	HEScore v2012	HEScore v2014
Number of Observations	877	877
Mean Measured	10659	10769
Median Measured	10348	10379
Mean Predicted	10340	10526
Median Predicted	10368	10584
Mean Difference	–319	–243
Median Difference	30	177
Standard Deviation of Difference	4081	3986
Mean Percent Difference	13%	13%
Median Percent Difference	0%	2%
Mean Absolute Difference	3008	2916
Median Absolute Difference	2382	2253
Mean Absolute Percent Difference	35%	33%
Median Absolute Percent Difference	24%	22%

[9] Equations for the statistics presented can be found in Appendix D.

	HEScore v2012	HEScore v2014
Percent Root Mean Square Error	38%	37%
Percent of Homes < ± 25% Different	54%	55%
Percent of Homes < ± 50% Different	80%	81%
R^2 of Regression	0.234	0.296
Slope of Regression	0.294	0.350
Intercept of Regression	7203	6754

Figure 7 shows HEScore-predicted site versus weather-normalized measured site natural gas use. Figure 8 shows the distribution of differences between the HEScore-predicted site and weather-normalized measured site natural gas use. The distribution is nearly symmetrical around zero, with only a slight negative bias.

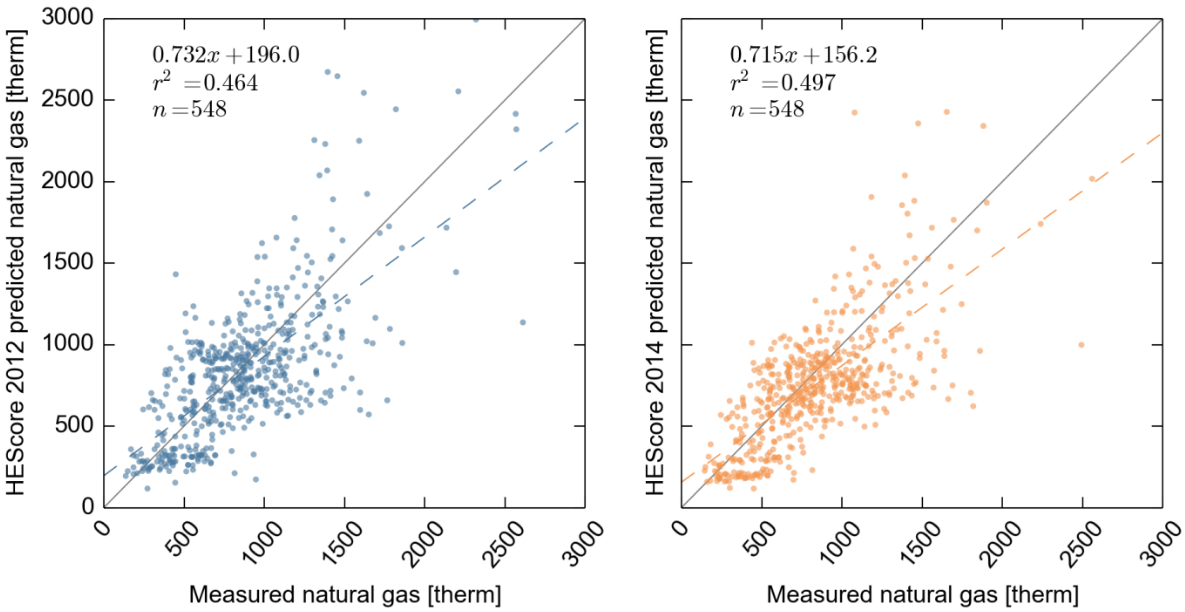

Figure 7. HEScore-predicted site versus weather-normalized measured site natural gas use

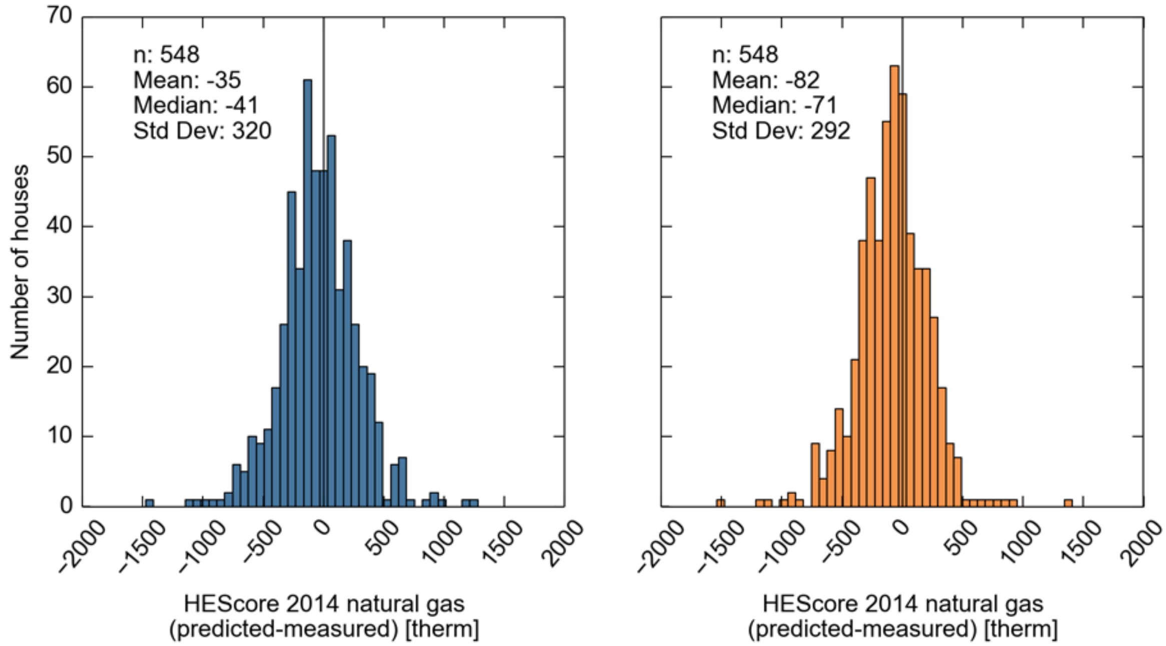

Figure 8. Distribution of differences between HEScore-predicted and measured site natural gas use

Table 5 summarizes the differences between predicted and weather-normalized measured natural gas use. The distributions are not exactly normal, but are reasonably close. An F-test using the standard deviations indicates that the variance in the HEScore 2014 natural gas use (predicted – measured) is significantly different from HEScore 2012. HEScore 2014 has about a 9% lower standard deviation for natural gas use compared to 2012.

Table 5. Statistical Summary of Differences Between Predicted and Weather-Normalized Measured Natural Gas Use (Predicted Therms – Measured Therms)

	HEScore v2012	HEScore v2014
Number of Observations	548	548
Mean Measured	863	837
Median Measured	841	804
Mean Predicted	828	754
Median Predicted	801	736
Mean Difference	-35	-82
Median Difference	-41	-71
Standard Deviation of Difference	320	292
Mean Percent Difference	1%	-6%
Median Percent Difference	-6%	-10%
Mean Absolute Difference	243	227
Median Absolute Difference	185	181
Mean Absolute Percent Difference	30%	28%

	HEScore v2012	HEScore v2014
Median Absolute Percent Difference	24%	25%
Percent Root Mean Square Error	37%	36%
Percent of Homes < ± 25% Different	52%	50%
Percent of Homes < ± 50% Different	84%	84%
R^2 of Regression	0.464	0.497
Slope of Regression	0.732	0.715
Intercept of Regression	196	156

Figure 9 shows a cumulative distribution of percent differences between predicted and weather-normalized measured electric energy use. At the 50% point on the x-axis, the lines cross the median percent difference value on the y-axis. The cumulative distribution changes very little between versions. HEScore crosses the 50% point at 0% overprediction for HEScore 2012 and at 2% overprediction for HEScore 2014.

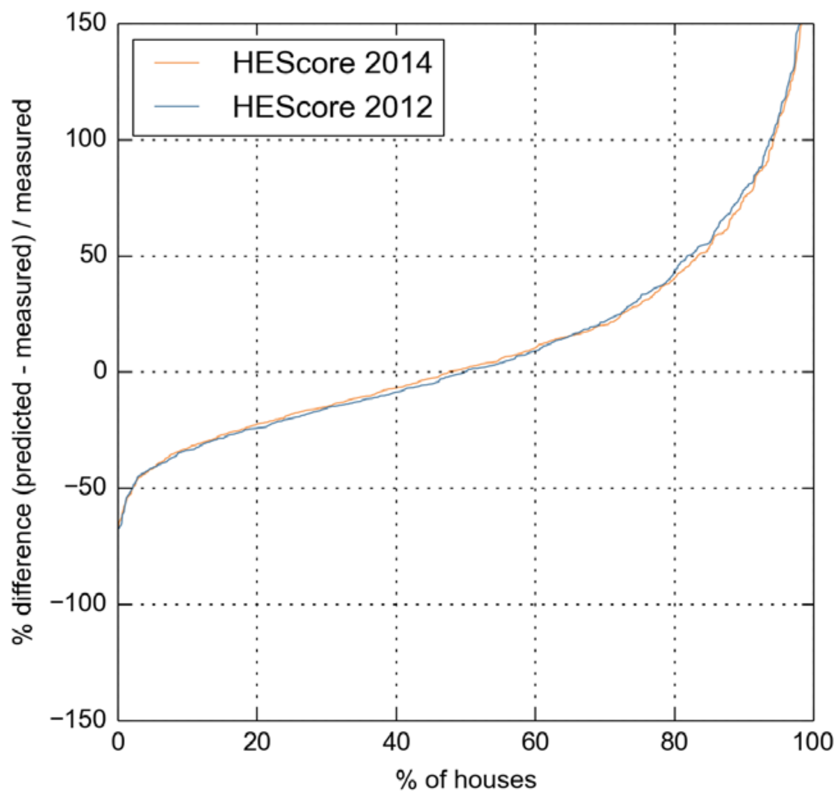

Figure 9. Cumulative distribution plot of percent differences between predicted and weather-normalized measured site electricity use for the three tools evaluated[10]

[10] Data points above 150% difference are not shown on the graph.

Figure 10 shows a cumulative distribution of the percent difference between predicted and weather-normalized measured natural gas use. On the whole, HEScore 2014 predicts lower natural gas use and underpredicts natural gas use more often. HEScore 2014 underpredicts natural gas use 63% of the time; HEScore 2012 underpredicted 56% of the time. The percent root mean square error (RMSE) decreased slightly from 37% to 36% (see Table 5), indicating a slight but not necessarily significant improvement in overall prediction. However, there is a slight bias of the natural gas use error toward underprediction, which increased between the 2012 and 2014 versions of HEScore from a median percent difference of –6% and –10%, respectively.

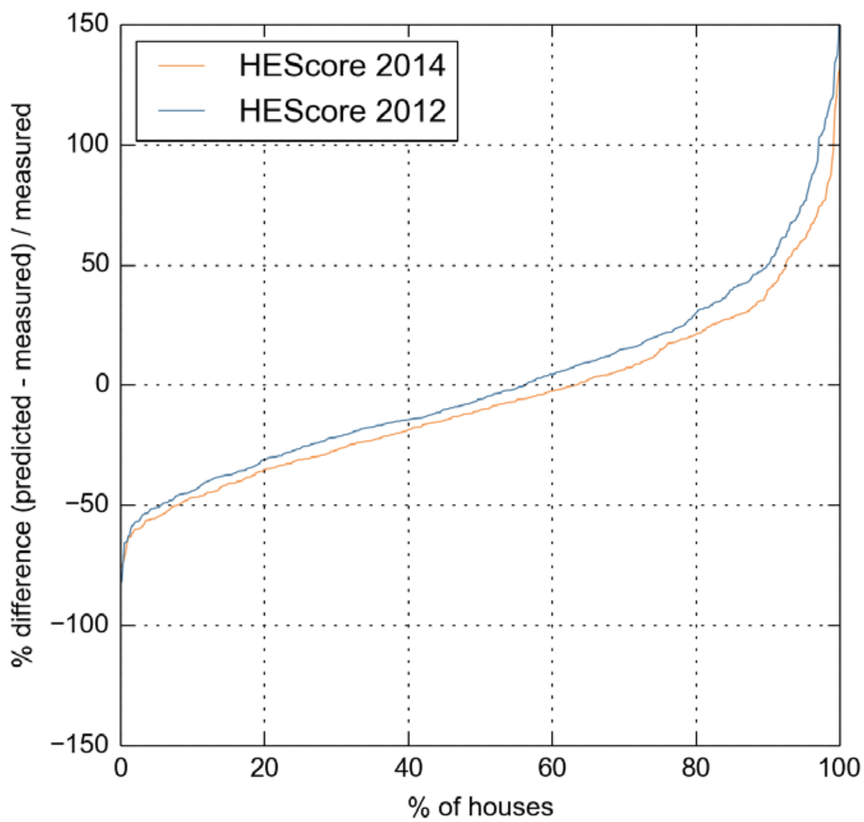

Figure 10. Cumulative distribution plot of percent differences between predicted and weather-normalized measured site natural gas use for the three tools evaluated[11]

[11] Data points above 150% difference are not shown on the graph.

3 Statistical Models

To estimate which inputs contribute the most to differences between HEScore predictions and measured energy uses, we applied a statistical analysis approach to the BAFDR records. More specifically, multiple linear regression (MLR) was used to develop empirical models from HEScore inputs and utility billing data. This section covers the approach taken, the resulting models, and what can be concluded from these models.

3.1 Approach

In MLR, a least-squares-fit algorithm is applied to a dataset that contains multiple records; each record contains one y-value and its associated x-values. The general model equation for MLR follows:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n + \varepsilon$$

where,

y is the dependent variable

β_0 is the intercept

β_1 through β_n are the coefficients

x_1 through x_n are the independent variables (inputs)

ε is the remaining error.

Most statistical software programs calculate the coefficients and probability values that allow one to estimate which coefficients are significant. Polynomial terms (i.e., x_n^2) and interaction terms (i.e., $x_1 x_2$) are sometimes included in the model if they improve the overall fit and have minimal correlation with the other independent variables. Although one starts out with a model containing practically all possible independent variables, common practice is to eliminate insignificant variables until a "reduced" model containing only significant variables is achieved.

The output from most MLR programs is a table that contains the following statistics for each variable used in the model:

- Coefficient
- Standard error
- t value
- Probability value.

The coefficient is determined from the least-squares fit, the standard error is essentially the standard deviation calculated for each coefficient, the t value is the coefficient divided by the standard error, and the probability value is determined from the *Student's t distribution* (Princeton University 2007). The probability value (Pr(>|t|)) is used to determine whether a variable is significant. The general approach is to keep variables in the model if the probability value is 0.05 or less. For this study, only variables that met the 0.05 criterion were kept in the reduced model, unless noted. The R language was used for statistical analysis of data for this study (R Core Team 2013). The output from the R program uses the term *estimate* for coefficients (Rodríguez 2013). The absolute t value reasonably indicates the importance of a variable. In addition to the table, the regression output includes the estimates for R-squared and adjusted R-squared. R-squared is also called the *coefficient of determination* and indicates how well data points fit a line or curve regardless of whether the terms in the model are significant. Adjusted R-squared takes into consideration whether the model significantly improves as terms are added (Montgomery 1997).

3.2 HEScore Dependent Variables and Inputs

To evaluate HEScore 2014 results, we used the following dependent variables in four separate empirical models:

1. Measured site electricity use (weather-normalized)
2. Measured site natural gas use (weather-normalized)
3. Difference site electricity = (predicted site electricity use) – (measured site electricity use)
4. Difference site natural gas = (predicted site natural gas use) – (measured site natural gas use).

Separate models for *measured site electricity* and *measured site natural gas* were created to estimate which inputs correlate with measured use at a significant level and to evaluate how much variability in the measured results can be explained by these inputs. The next step was to model differences between HEScore predictions and measured energy uses. Again, separate models were created for site electricity and site natural gas use. The coefficients from these difference models can be examined to evaluate which HEScore inputs contribute to overpredictions and which contribute to underpredictions.

There are approximately 40 HEScore inputs. Some, such as conditioned floor area, are numeric, but many use construction codes to describe discrete types of building construction components (Mills 2008). Wall types, roof types, foundation types, and many other components have separate codes that make up a building. For statistical analysis, the frequency of each specific code was examined and then a binary variable was defined for each. More details on variable coding for statistical analysis are given in Section 3.4. An example of a ceiling construction code is "ecwf30," which is defined as 3.5-in. wood ceiling joists @ 24 in. o.c., 10.5-in. (R-30) fiberglass fill ceiling insulation, and 0.5-in. gypsum wallboard. These construction codes were used in the variable names to allow lookup in the tables for further details. Because there were many construction codes, we decided that in cases with at least seven observations, a binary variable would be created. Insulation R-values were also extracted from these construction codes. Because insulation R-values can be treated as numeric variables, they often are more understandable in the MLR model. Numeric R-values were kept instead of binary construction code variables if the resulting model was comparable. For the variable Roof_R_Value, the R-values were extracted from both "roof" construction codes and "ceiling" construction codes, because often a building had insulation listed for one but not for the other.

3.3 Dataset Limitations and Bias

The current BAFDR contains both measured utility data and HEScore building asset characteristics for 1085 homes. The data are limited primarily to five states (Minnesota, North Carolina, Oregon, Texas, and Wisconsin).[12] Only houses that used electric or natural gas space heating were included in this analysis. To be included, a house needed to have positive measured electricity use. In addition, houses heated with natural gas needed to have positive measured natural gas use. This reduced the dataset to 845 homes.

Table 6 lists how many houses from each historical dataset and state could be considered in the statistical analysis after filtering out homes heated by fuel oil or liquid petroleum gas. Roberts et al. (2012) found no significant difference that could be attributed to the different datasets, but Minnesota appeared to be a significant factor. This was most likely the case because the datasets were highly correlated with the states. For this study each state was given its own binary variable, but not the individual datasets.

Table 6. Summary of Homes Considered for Statistical Analysis

Dataset Name[13]	Dataset Source (Delivered to NREL)	Location of Homes	Number of Housing Units	New/ Existing
Oregon EPS Pilot Study	Earth Advantage	OR	168	Existing
Wisconsin Energy Star Study	EPA	WI	57	New
Wisconsin Housing Characterization Study	Energy Center of Wisconsin	WI	172	Existing
Advanced Energy SystemVision Program	Advanced Energy	NC (2 in TN)	254	New
Houston ENERGY STAR Homes Program	Blasnik and Associates	TX	81*	New
Wisconsin Building America	Wisconsin Energy Conservation Center	WI	68	Existing
Minnesota Building America	Center for Energy and Environment	MN	45	Existing

*This dataset represents a larger number of homes due to the sampling approach.

Often an input or explanatory variable can be correlated with a particular state. This can result in distorted estimates of variable coefficients using MLR and cause some variables to appear significant when they are not. MLR models give indications of the most likely inputs that correlate with the dependent variable, but they do not provide absolute certainty. In cases where an input variable correlated strongly with a state, the input variable was kept in the model and the state binary variable was excluded.

In addition to HEScore inputs, climate differences are believed to be important. To capture actual climatic differences, two additional independent variables, *heating degree days* (HDDs) (base 65°F) and *cooling degree days* (CDDs) (base 65°F), were joined to the dataset and treated as

[12] Two houses from North Carolina Advanced Energy are located in Tennessee. From the 5-digit zip code, these houses appear to be located in eastern Tennessee, very near other houses in North Carolina; hence, they are kept in the analysis.
[13] Further description of these datasets can be found in Appendix B.

numeric variables. Values for these variables were taken from TMY weather files at weather stations near home locations (based on zip code values).

3.4 Variable Coding

Coding for the HEScore inputs was done almost exactly as before (Roberts et al. 2012) with only a few exceptions. For continuous (numeric) or pseudo-continuous variables, the coding involved subtracting the mean and then dividing this difference by the standard deviation. This is called *univariate coding*. House orientation was not treated as a numeric variable, but rather categorized by the different directions. Instead of using window areas for each side of the house as separate variables, window area for the house was summed up to a total window area. U-values and solar heat gain coefficient (SHGC) were calculated using a weighted average (by window area) for each house. A similar approach was used for skylight area, skylight U-value, and skylight SHGC. Water heater energy factor was not included in the analysis because it depends on fuel type and because some natural gas-heated homes had electric water heaters and vice versa. Heating efficiency was included in the analysis, but only for homes using natural gas for heating.

Categorical variables were checked for frequency of occurrence and coded in a similar fashion as before. Table 7 demonstrates the categorical coding for heating type with *record counts* based on the most current dataset. The resulting number of usable predictor variables was 79. A complete list of HEScore variables with descriptions is included in Appendix B.

Table 7. Example of Binary Coding for Heating Type Category HEScore Input

Heating Fuel and Type Description	Record Count	C_HT_EBB	C_HT_EFN	C_HT_EHP	C_HT_GBL	C_HT_GWF
Natural Gas Furnace (Control)	577	0	0	0	0	0
Electric Baseboard	9	1	0	0	0	0
Electric Furnace	3	0	1	0	0	0
Electric Heat Pump	221	0	0	1	0	0
Natural Gas Boiler	28	0	0	0	1	0
Natural Gas Wall Furnace	6	0	0	0	0	1
Natural Gas None (Did Not Code)	1	0	0	0	0	0

3.5 Models of Measured Energy Use

For this assessment, the observations were not separated into a model set and a test set, because Roberts et al. (2012) already confirmed that inputs found to be significant did in fact predict reasonably well. For comparison to the previous study, site electricity use reported in kWh was converted to site MMBtu and site natural gas use reported in therms was also converted to site MMBtu.

Table 8 shows the resulting MLR model with *measured site electricity* as the dependent variable. All variables listed are significant at a confidence level of 95% or greater. The variables are sorted by the absolute t value from the largest to the smallest; hence, the variables that have the greatest influence are listed at the top of the table and those with the least influence are listed at the bottom. The intercept is not a variable coefficient and is listed at the top of the table for convenience. If a variable has a positive estimate (coefficient), the MLR model predicts that electric energy use increases as the variable value increases, or for binary variables, if the variable value is 1 ("True"). Conversely, if the estimate is negative, the MLR model predicts a decrease in electric energy usage as the variable value increases.

An adjusted R-squared value of 0.437 resulted; this implies that the model can explain approximately 44% of the observed variability in measured electricity use. This is very close to the adjusted R-squared value of 0.433 determined in the previous study. Three of the most important variables are the same as those found in the previous model. In fact, a fourth variable, C_hwFuel_elec is essentially the same as C_hwFuel with just a different coding. The estimates are also comparable to the previous model estimates. In addition, variables indicating the presence of cooling equipment significantly increased electricity use. The degrees of freedom (DF) essentially indicate how many values can be used for error determination after the intercept and coefficients are determined.

Table 8. Measured Site Electricity Model:
R-Squared = 0.443 , Adjusted R-Squared = 0.437, DF = 830

Variable	Variable Description	Estimate (MMBtu)	Std. Error	Abs (t value)	Pr(>|t|)
(Intercept)		28.09	0.95	29.60	< 2e-16
C_hwFuel_elec	Electricity used for domestic hot water	12.35	1.35	9.13	< 2e-16
C_number_bedrooms	Number of bedrooms	3.81	0.47	8.04	0.0000
C_CDD65	CDDs (base 65°F)	3.71	0.49	7.55	0.0000
C_conditioned_floor_area	Conditioned floor area (ft^2)	4.69	0.63	7.40	0.0000
C_HT_EBB	Heating type is electric baseboard	29.90	4.28	6.98	0.0000
C_CT_ehp	Cooling type is electric heat pump	10.44	1.74	6.02	0.0000
C_envelope_leakage	Envelope leakage	2.24	0.50	4.53	0.0000
C_CT_cac	Cooling type is central air conditioner	3.97	1.12	3.53	0.0004
C_CC_ecwf49_60	Ceiling construction code ecwf49 or 60 (3.5-in. wood joists @ 24 in. o.c., R-49 or R-60 fiberglass fill insulation, 0.5-in gypsum wallboard)	–3.87	1.39	2.79	0.0055
C_RC_rfwf11co	Roofing construction code rf.wf11co (5.5-in. wood roof rafters @ 24 in. o.c., 2-in. air space + R-11 mineral fiber batt insulation, 0.5-in. gypsum wallboard, 0.25-in. asphalt shingles, 0.125-in. felt membrane, 0.625-in. plywood sheathing)	–7.90	3.19	2.47	0.0136

Figure 11 shows a graph of the measured site electricity use versus the MLR prediction. Although there is extensive variability, a definite trend can be observed, confirming that the correlations of model inputs to measured site electricity use are valid.

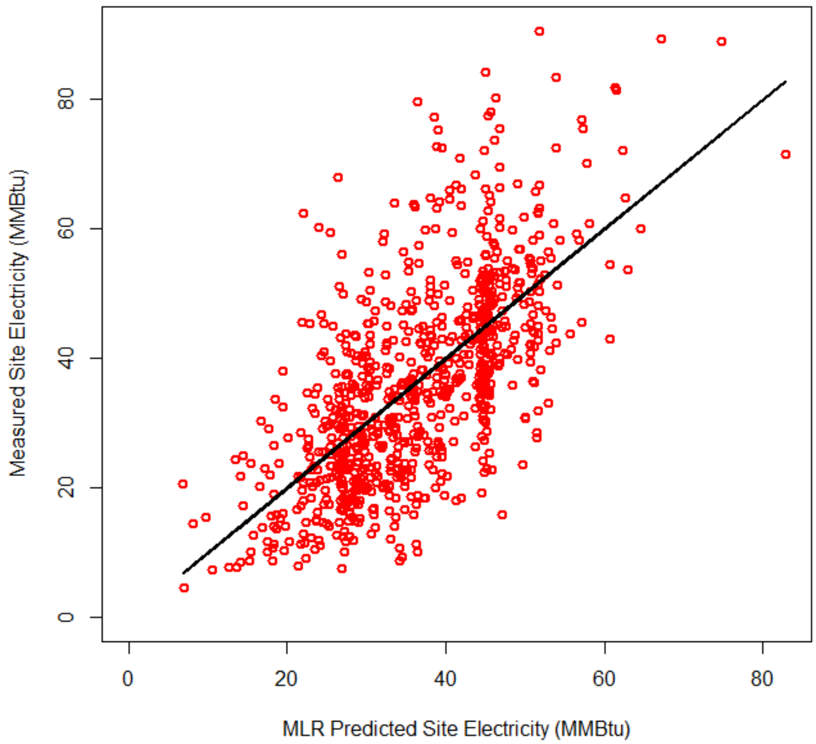

Figure 11. Measured versus MLR-predicted site electricity use

For modeling *measured site natural gas*, only buildings where natural gas is used for space heating were included. This reduced the number of observations to 543. Table 9 shows the resulting MLR model with *measured site natural gas* as the dependent variable. The resulting adjusted R-squared for this model is 0.670, which is slightly greater than the previous model adjusted R-squared of 0.650. This indicates that the model explains 67% of the variability. The graph in Figure 12 shows measured versus MLR-predicted site natural gas use. Most of the model variable coefficients appear to agree with how we might expect the input to influence natural gas use. For example, houses in locations with more HDDs would be expected to use more natural gas for space heating. Increased envelope leakage and increased floor area both contribute to greater natural gas use. A gas furnace with higher efficiency reduces natural gas use. Six of the significant variables are essentially the same as those in the previous model. Wall R-value appears to replace the two floor construction codes found in the previous model. Total window area replaces the two window areas that included orientation. The *age in years* indicates a reduction in natural gas use for older buildings. Although a few additional homes were added to the current dataset, the correlation of age in years with the Oregon dataset still exists as it did in Roberts et al. (2012). As older homes in other states are added to the BAFDR, a better test should result for the *age in years* variable.

Table 9. Measured Site Natural Gas Model:
R-Squared = 0.670, Adjusted R-Squared = 0.664, DF = 524

Variable	Variable Description	Estimate (MMBtu)	Std. Error	Abs (t value)	Pr(>|t|)
(Intercept)		62.70	1.43	43.88	< 2e-16
C_HDD65	HDDs (base 65°F)	22.26	1.15	19.33	< 2e-16
C_total_window_area	Total window area (ft^2)	11.12	1.23	9.06	< 2e-16
C_envelope_leakage	Envelope leakage	8.83	1.22	7.25	0.00000
C_conditioned_floor_area	Conditioned floor area (ft^2)	9.12	1.57	5.81	0.00000
C_NG_HeatingEff	Natural gas heating efficiency	-5.77	1.02	5.66	0.00000
C_HW_EST	Water heater is electric storage tank	-16.24	3.07	5.29	0.00000
C_wtmean_window_shgc	Weighted average of window SHGC	4.96	1.14	4.35	0.00002
C_HT_GBL	Heating type is natural gas boiler	15.23	4.67	3.26	0.00120
C_age_years	Age of house (2011 - year built)	–4.81	1.55	3.11	0.00200
C_Wall_R_Value	Wall R-value	–3.17	1.07	2.95	0.00330

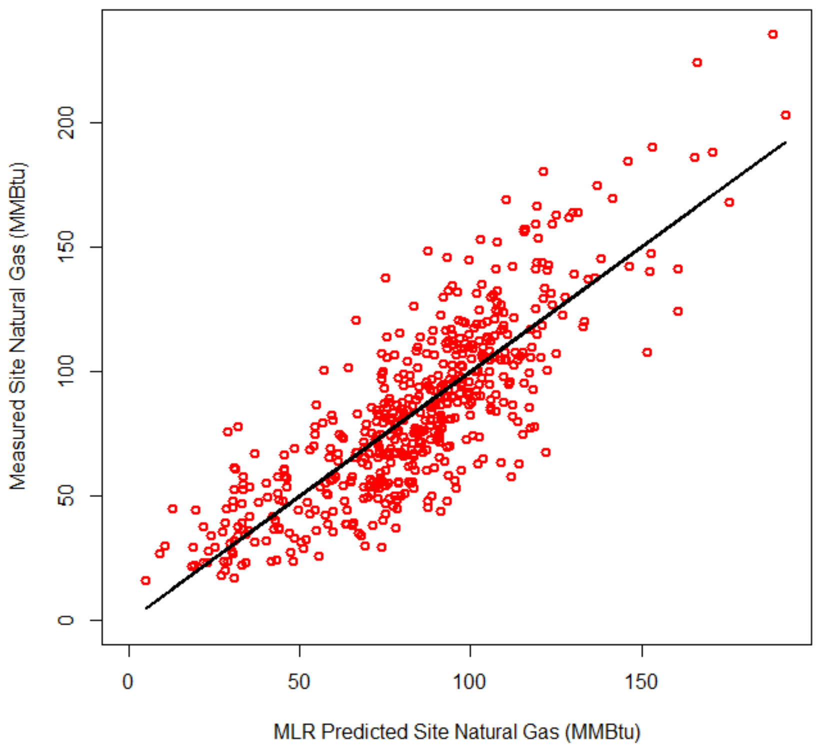

Figure 12. Measured versus MLR-predicted site natural gas use

3.6 Models of Differences Between Predicted and Measured Energy Uses

Table 10 shows the resulting MLR model with difference or delta site electricity use as the dependent variable. Again, the delta is *predicted site electricity* minus *measured site electricity*. The resulting adjusted R-squared value for this model is 0.155, which is less than 0.199 of the previous model. Only the number of bedrooms was in the previous model, although it is the most important in both models and has similar coefficients. This model does contain at least one wall construction code (wwwf21wo), indicating that a potentially higher wall R-value contributes to decrease in delta electricity. This is consistent with the wall R-value coefficient found in the previous model. Also, C_HT_EHP is highly correlated with C_CT_ehp, both indicating an electric heat pump is used, so essentially this is consistent.

The sign of the coefficients indicates the direction in which increasing a variable causes the delta variable to change. For example, as the number of bedrooms increases, the delta electricity decreases. The variable by itself does not necessarily result in over- or underprediction, but rather the prediction for any given house can be greater or less than actual energy use based on the combination of all variables. If a particular one-bedroom house had predicted energy use greater than measured, modeling this house with two bedrooms (holding all other variables constant) would result in a lower delta (hence, potentially better agreement). Modeling this house with three bedrooms (holding all other variables constant) would result in an even lower (or possibly in a negative) delta, indicating that the model is now underpredicting electricity use.

Table 10. Delta Electricity Model: R-Squared = 0.164, Adjusted R-Squared = 0.155, DF = 831

Variable	Variable Description	Estimate (MMBtu)	Std. Error	Abs (tvalue)	Pr(>|t|)
(Intercept)		1.68	0.66	2.55	0.0111
C_number_bedrooms	Number of bedrooms	–2.69	0.42	6.34	0.0000
C_HT_EHP	Heating type is electric heat pump	–6.13	1.16	5.30	0.0000
C_CC_ecwf49_60	Ceiling construction code ecwf49 or 60 (3.5-in. wood joists @ 24 in. o.c., R-49 or R-60 fiberglass fill insulation, 0.5-in. gypsum wallboard)	3.48	1.21	2.87	0.0043
C_roof_absorptance	Roof absorptance	1.14	0.40	2.84	0.0047
C_wtmean_window_shgc	Weighted average of window SHGC	–1.16	0.46	2.50	0.0126
C_num_floor_above_grade	Number of floors above grade	1.09	0.46	2.38	0.0175
C_orientation_west	House orientation, front facing west	–2.30	0.97	2.37	0.0180
C_WC_ewwf21wo	Wall construction code ecwf21wo (5.5-in. wood studs @ 16 in. o.c., R-21 mineral fiber batt insulation, 0.5-in. lapped wood siding, 0.5-in. fiberboard sheathing, 0.5-in. gypsum wallboard)	–4.19	1.80	2.34	0.0197
C_CC_ecwf30	Ceiling construction code ecwf30 (3.5-in. wood joists @ 24 in. o.c., R-30 fiberglass fill insulation, 0.5-in. gypsum wallboard)	–2.17	0.94	2.32	0.0206

The very low R-squared values indicate that only a fraction of the difference between HEScore-predicted and measured values can be explained by the inputs. Nevertheless, a few variables might still be worth investigating. At least two—and possibly three—(C_HT_EHP correlated with C_CT_ehp) of the significant variables in the delta site electricity model occur in the measured site electricity model. The negative coefficients for number of bedrooms, electric heat pump, wall code ewwf21wo, and a few other variables potentially correlate to a decrease in the delta electricity. The number of bedrooms may still not reflect total occupant electricity use in HEScore. The negative coefficient for the electric heat pump might suggest that estimates of heat pump performance are too optimistic. The positive coefficients for variables such as ceiling construction ecwf49_60, roof absorptance, number of floors above grade potentially correlate to an increase in the delta electricity. Very high levels of ceiling insulation may provide an additional unanticipated benefit, such as covering duct systems that might be exposed with less insulation.

Table 11 shows the resulting MLR model with *difference or delta site natural gas* as the dependent variable. The resulting adjusted R-squared value for this model is 0.384 compared to 0.456 in the previous model. A simple analysis of variance test was done by pooling the results from delta electricity model with the delta natural gas model. This test did not disprove the null hypothesis (no difference between the previous and current HEScore results); hence, the reduction in adjusted R-squared cannot be considered significant. Even if found to be significant, the difference could be attributed to some of the changes made in modeling the data, a few more observations available, or other factors. In fact, total window area was significant with a negative coefficient, whereas north-facing window area was significant in the previous model, also with a negative coefficient. The negative coefficient for window area indicates that HEScore underestimates the heat losses of windows. Increasing total window area correlates with increasing natural gas use. The delta model indicates increasing the total window area lowers the difference value, and may indicate underprediction.

Other variables in the previous model that at least have the same sign on the coefficient are HDDs, roof R-value, and natural gas heating efficiency. HDDs were significant for both the *measured site natural gas* and the delta model. Both had positive coefficients. The indication is that HEScore overpredicts the impact of more HDDs, perhaps for a variety of reasons: the assumed heating set point may be too high, variation in indoor air temperature, imperfect modeling of empty wall cavities, etc. Basically the same or similar variables still contribute to the difference between predicted and measured natural gas use.

Heating efficiency was significant for both the *measured site natural gas* and the delta model, but had negative coefficients. The *measured site natural gas* model shows a reasonable trend, declining natural gas use as the heating system efficiency increases. The negative estimate for the delta model indicates that HEScore may still not completely capture the impact of increasing system efficiency. Duct location binary variables were tested and the duct location in an unconditioned basement was significant. Other factors such as furnace location, which are not currently modeled in HEScore could also affect the heating efficiency relationships.

Table 11. Delta Natural Gas Model: R-Squared = 0.399, Adjusted R-Squared = 0.384, DF = 520

Variable	Variable Description	Estimate (MMBtu)	Std. Error	Abs (t value)	Pr(>|t|)
(Intercept)		–9.97	2.35	4.24	0.0000
C_total_window_area	Total window area (ft^2)	–8.07	1.06	7.59	0.0000
C_HDD65	HDDs (base 65°F)	9.01	1.23	7.31	0.0000
C_Roof_R_Value	Roof R-value determined from roof construction and ceiling construction inputs	–5.65	1.05	5.37	0.0000
C_FT_slab	Foundation type is slab	12.36	2.59	4.78	0.0000
C_Wall_R_Value^2	Wall R-value squared, determined from wall construction codes	3.14	0.77	4.07	0.0001
C_NG_HeatingEff	Natural gas heating efficiency	–3.16	1.03	3.08	0.0021
C_wtmean_window_shgc	Weighted average of window SHGC	–3.29	1.13	2.91	0.0038
C_CT_cac	Cooling type is central air conditioner	–6.21	2.16	2.88	0.0042
C_HW_EST	Water heater is electric storage tank	–8.14	3.20	2.55	0.0111
C_DL_uncond_basement	Duct location in unconditioned basement	6.70	3.05	2.20	0.0285
C_orientation_north	House orientation, front facing north	–4.76	2.20	2.17	0.0306
C_num_floor_above_grade	Number of floors above grade	–2.13	1.01	2.11	0.0350
C_Wall_R_Value	Wall R-value determined from wall construction codes (kept in model because wall R-value squared term significant)	–1.55	1.39	1.12	0.2653

The roof R-value was not significant in the *measured site natural gas* model, but is significant in the delta model. One would expect that as roof R-value increases, the natural gas use should decrease. However, actual natural gas use may not decrease as much as HEScore predicts. Hence the predicted use might be lower than the measured use.

Slab foundation was not significant in other models, but was significant in the delta natural gas model. The delta model indicates a possible over-prediction by HEScore for homes with this type of foundation. This might be an artifact of limited data, although 25% of the natural gas-heated homes had slab foundations. Also, slab foundations are somewhat correlated (R = 0.50) with the Texas binary variable.

A central air conditioner appears to decrease the delta natural gas use. This might indicate greater leakage in homes with air conditioners.

For the previous HEScore dataset, wall R-value was not significant. Instead, a couple of wall construction codes were significant; a low insulation wall had a positive coefficient and a higher insulation wall had a negative coefficient. As mentioned in Section 3.1, polynomial terms are sometimes added to the statistical models when significant and an improvement in the model is observed (greater adjusted R-squared value). For the current HEScore dataset, wall R-value squared was significant, indicating the relationship to delta natural gas use might not be a simple straight line. Wall R-value has a negative coefficient indicating an initial decrease in delta natural gas use for low wall R-values. Although the probability value for wall R-value is greater than 0.05, this term is kept in the model based on the hierarchy principle (Pennsylvania State University 2014).

3.7 Summary

MLR models indicate significant correlations between site measured energy use and several HEScore inputs. These methods also indicate significant correlations between differences (HEScore predicted minus measured energy use) and several HEScore inputs. How these inputs are collected and used in the HEScore prediction models can be investigated to identify causes for differences from measured energy usage and potential improvements to software inputs and models. Even with the improvements in HEScore predictions, there still appears to be substantial differences between predicted and measured energy use that correlate strongly with model inputs.

When comparing the results from HEScore 2012 to HEScore 2014, one must keep in mind that more observations were available for 2014. Also, the MLR models for 2012 used only 75% of the available data; the remaining 25% were used to validate the models. The predictive ability of the 2012 models was reasonably good (Roberts et al. 2012), so it was decided there was no need to perform this check again. Instead, all available data were used to create the MLR models. Therefore, some of the observed differences in models are likely due to using more data.

Table 12 summarizes the similarities and differences between the delta electricity model for 2012 compared to 2014. At least four significant variables are either the same or similar interpretations. Other variables could be due the differences in data. Two of the heating types (electric furnaces and electric baseboards) were significant in 2012 but not in 2014. This could be an improvement in that heating type (other than heat pumps) no longer contributes to the delta electricity in HEScore 2014.

Table 13 summarizes the similarities and differences between the delta natural gas model for 2012 compared to 2014. At least eight significant variables are either the same or similar interpretations. Other variables could be due the differences in data. The interpretation of a specific variable is not always understood. A variable can occasionally be deemed significant when it actuality is not because of the data. The most significant variables found in both models are likely truly significant.

Table 12. Similarities and Differences Between Delta Electricity Model for 2012 Compared to 2014

	Variable From 2012	Variable From 2014	Summary
Similarities	C_numberBedrooms	C_number_bedrooms	Most significant in both models, coefficients have same sign and magnitude.
	C_CT_ehp	C_HT_EHP	Indicates heat pump, highly correlated with each other (hence essentially the same variable), coefficients have same sign and magnitude.
	C_WallRValue	C_WC_ewwf21wo	Indicates that increase in wall R-value decreases delta electricity, coefficients have same sign, magnitude not directly comparable because of different coding.
	C_floorArea	C_num_floor_above_grade	Not strongly correlated in data, but both variables indicate house size to some extent. Coefficients have same sign.
Differences	C_HT_EFN		Heating type electric furnace. The fact that this is not significant in 2014 indicates HEScore 2014 likely models electric furnaces reasonably well.
	C_ST_dseab		Skylight category not significant in 2014.
	State_MN		Minnesota not significant in 2014.
	C_WC_ewps19wo		Wall construction category with positive coefficient in 2012.
	C_HT_EBB		Heating type electric baseboard. The fact that this is not significant in 2014 indicates HEScore 2014 likely models electric baseboard heaters reasonably well.
	C_CC_ecwf21		Ceiling construction category with positive coefficient in 2012.
		C_CC_ecwf49_60	Ceiling construction category with positive coefficient in 2014.
		C_roof_aborptance	Roof aborptance with positive coefficient in 2014.
		C_wtmean_window_shgc	Weighted average of window SHGC with negative coefficient in 2014.
		C_orientation_west	House orientation, front facing west with negative coefficient in 2014.
		C_CC_ecwf30	Ceiling construction category with negative coefficient in 2014.

Table 13. Similarities and Differences Between Delta Natural Gas Model for 2012 Compared to 2014

	Variable From 2012	Variable From 2014	Summary
Similarities	C_HDD_65F	C_HDD65	HDDs (referenced to 65°F), coefficients have same sign and magnitude.
	C_E_WASG	C_wtmean_window_shgc	For 2012, east window area × SHGC was significant. SHGC weighted by window area on each side of house was used in 2014 analysis. Coefficients have same sign.
	C_heatingEfficiency	C_NG_HeatingEff	Heating system efficiency coefficients had same sign, but different values (possibly because of differences in data).
	C_RoofRValue	C_Roof_R_Value	Roof R-value coefficients have same sign and magnitude.
	C_N_WindowArea	C_total_window_area	North-facing window used in 2012 analysis. Only total window area was used in 2014 analysis, but coefficient sign agrees.
	C_houseOrientation	C_orientation_north	Coding for 2014 was different than 2012, hence not directly comparable, but interesting that orientation significant in both models.
	C_WC_ewwf00wo C_WC_ewwf19wo	C_Wall_R_Value C_Wall_R_Value^2	Coding for 2014 was different than 2012, but overall trend of decreasing delta natural gas use with increasing R-value occurs.
Differences	C_airLeakage50ip		Air leakage (cfm) found significant in 2012, but not significant in 2014.
	C_age_years		House age in years significant in 2012, but not significant in 2014.
	C_HT_GBL		Heating type of gas boiler increased delta natural gas use in 2012, but not significant in 2014.
		C_FT_slab	Foundation type slab had positive coefficient in 2014.
		C_CT_cac	Cooling type of central air conditioner had negative coefficient in 2014.
		C_HW_EST	Water heater of electric storage tank had negative coefficient in 2014.
		C_DL_uncond_basement	Duct location in unconditioned basement had positive coefficient in 2014.
		C_num_floor_above_grade	Number of floors above grade had negative coefficient in 2014.

It is important to note that the statistical models discussed in this section apply only to the current BAFDR data used to develop the models. As more data are collected, the coefficients will probably change, new inputs will be significant, and current significant inputs might prove to be not significant.

4 Sensitivity to Assessment of Whole-House Leakage

HEScore accepts either a quantitative measurement of whole-house leakage using a blower door or a qualitative assessment of whether the home has been air sealed.[14] When a user enters the quantitative results of a blower door test in CFM50, the software uses this datum to calculate the leakage area of the home, a direct input into the underlying DOE-2 infiltration model. When a user enters the qualitative assessment, HEScore estimates from historical data the leakage area of the home based on this input and other house characteristics.[15] The data and algorithm for estimation of leakage area were updated between HEScore 2012 and HEScore 2014.

4.1 Approach

In the data from the HEScore 2012 launch that NREL received, blower door measurements were collected for 1,489 homes. Also, of the homes in the BAFDR used for this HEScore analysis, blower door data were collected for 1,075 homes. NREL reran these homes through the Home Energy Saver 2014 API. Each home was run three times using different inputs for whole-house air leakage:

- Using the blower door data (CFM50)
- Using the qualitative assessment of *Sealed*
- Using the qualitative assessment of *Unsealed*.

4.2 Results of Simulating HEScore National Launch Data

Results of these runs are presented in Figure 13, which compares the total predicted source energy from HEScore using the airsealed/not airsealed input to the total predicted source energy from HEScore using the blower door measurement.

[14] The quantitative input is entered as cubic feet per minute at 50 Pascals of pressure (CFM50). The qualitative input is choosing *Yes* or *No* to the question: *Has the house been professionally air sealed?* The HEScore website help tip for the qualitative input reads as follows: *Answer "No" unless specific efforts have been made to seal the majority of the air leaks (thermal bypasses) in the home.*

[15] Details on the Home Energy Saver infiltration model can be found in the Home Energy Saver engineering documentation, available online at https://sites.google.com/a/lbl.gov/hes-public/.

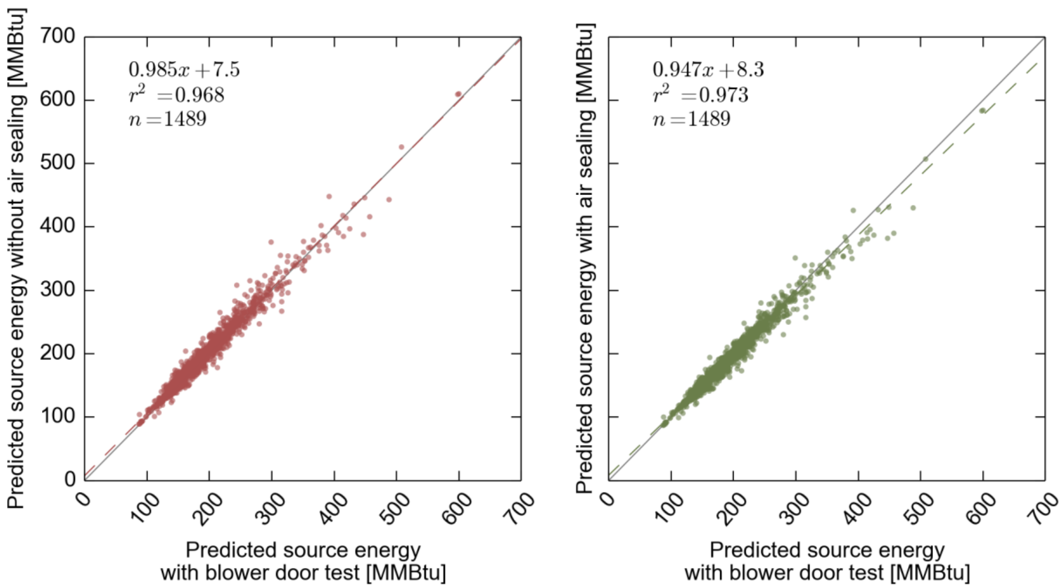

Figure 13. Predicted source energy use from HEScore using "Unsealed" input for whole-house air leakage (left) and "Sealed" input (right) versus measured whole-house leakage for HEScore homes

Figure 14 shows the frequency distribution of the predicted annual source energy use using the three variations in whole-house leakage input.

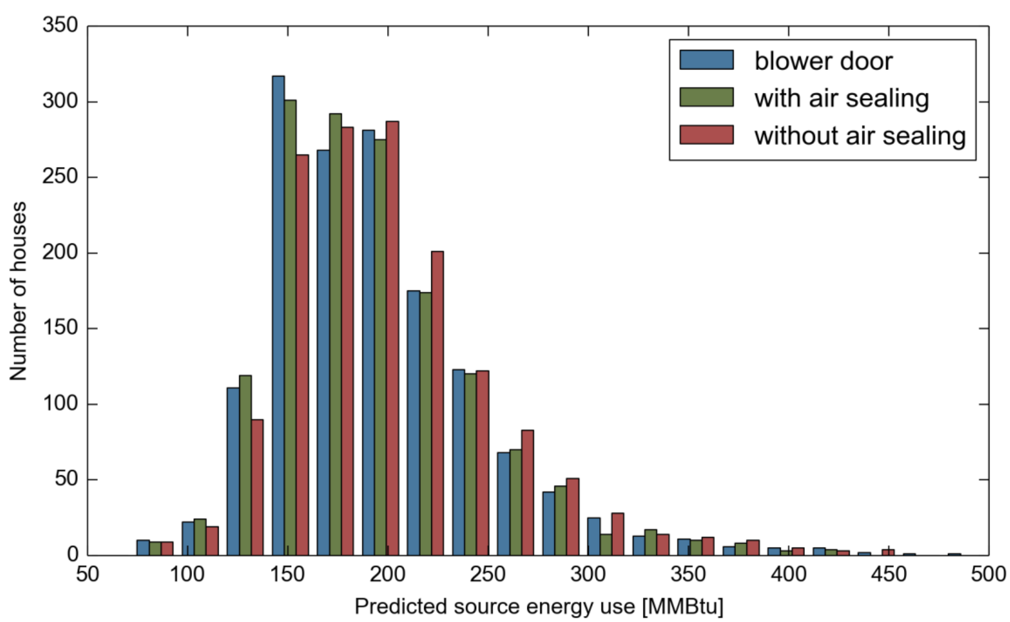

Figure 14. Distribution of HEScore-predicted source energy use for HEScore homes generated using three input scenarios for whole-house leakage

Table 14 shows the mean value of the source energy for each of the three whole-house infiltration input scenarios. As expected, an air-sealed house uses less energy than an unsealed house. Also, the houses with blower door measurements predict energy that is on average between the sealed and unsealed energy use.

Table 14. Average HEScore-Predicted Source Energy Use for Each Infiltration Assumption for HEScore Homes

	CFM50	Sealed	Unsealed
Mean Predicted Source Energy (MMBtu/yr)	195	193	200

Figure 15 shows the difference in predicted source energy use generated by HEScore using the quantitative and qualitative inputs for whole-house air infiltration. On average, when compared to the predictions from quantitative input, the source energy use decreases 2 MMBtu/yr when the sealed qualitative input is used and increases 5 MMBtu/yr when the unsealed qualitative input is used.

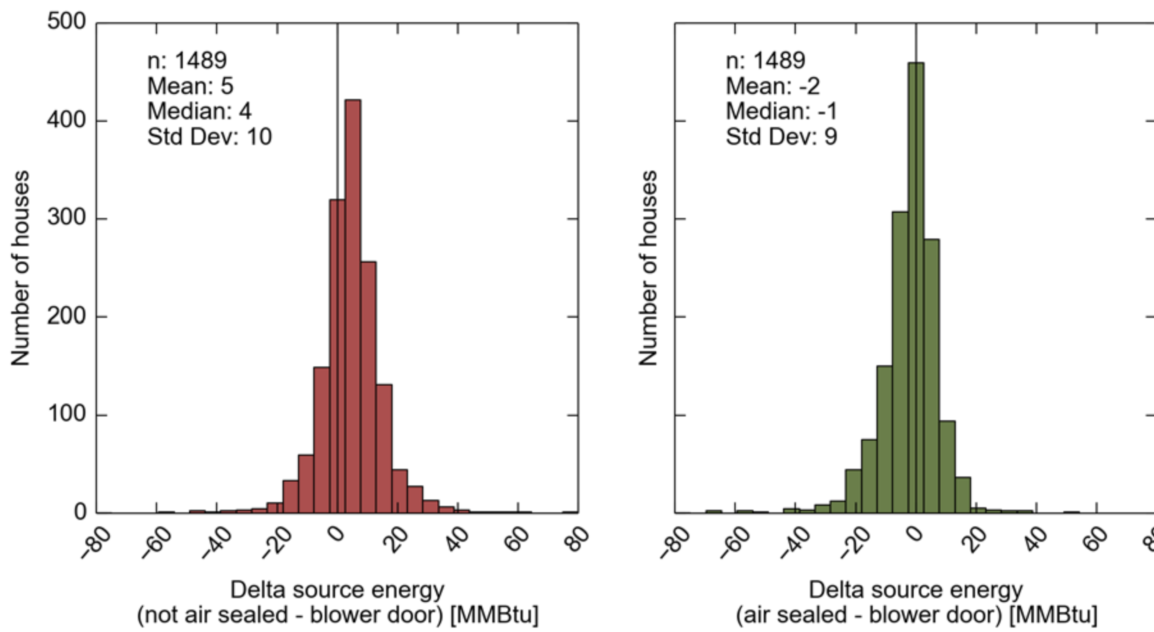

Figure 15. Distribution of differences in HESscore-predicted source energy use using qualitative and quantitative input for whole-house leakage for HEScore homes

An important measure of the functional accuracy of the infiltration assumptions is its effect on the end product for the consumer—the score. Figure 16 shows distribution of the differences in score when comparing qualitative inputs and blower door input for whole-house infiltration. When the "not air sealed" qualitative input is selected instead of the blower door test, the score remains within ±1 bin for 95% of the homes. When the “air sealed” quantitative input is selected, the score remains within ±1 bin for 96% of the homes.

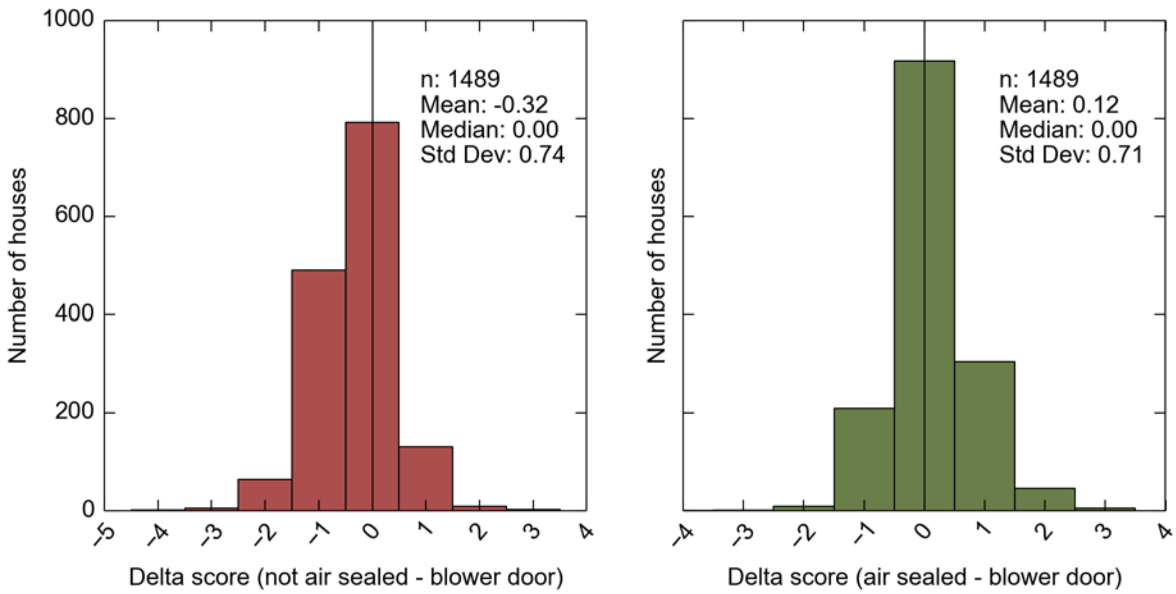

Figure 16. Distribution of differences in score using qualitative and quantitative input for whole-house leakage for HEScore homes

4.3 Results of BAFDR Data

The same approach that was used on the HEScore data was used to simulate the homes from the BAFDR through HEScore for each case. Results of these runs are presented in Figure 17. The figure compares the total predicted source energy from HEScore using the airsealed/not airsealed input to the total predicted source energy from HEScore using the blower door measurement.

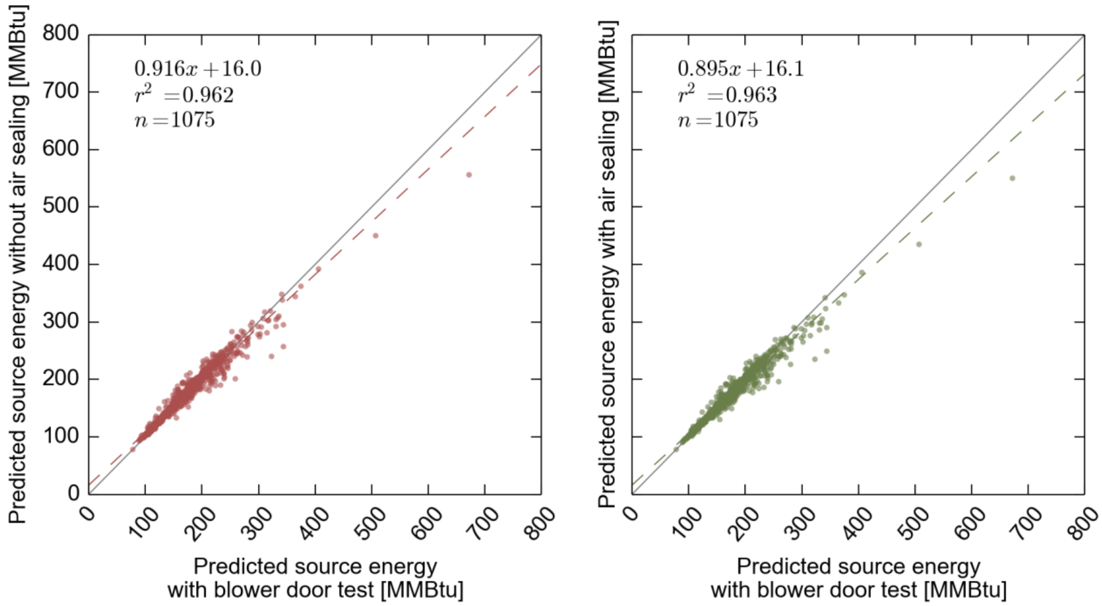

Figure 17. Predicted source energy use from HEScore using "Unsealed" input for whole-house air leakage (left) and "Sealed" input (right) versus measured whole-house leakage for BAFDR homes

Figure 18 shows the frequency distribution of the predicted annual source energy use using the three variations in whole-house leakage input.

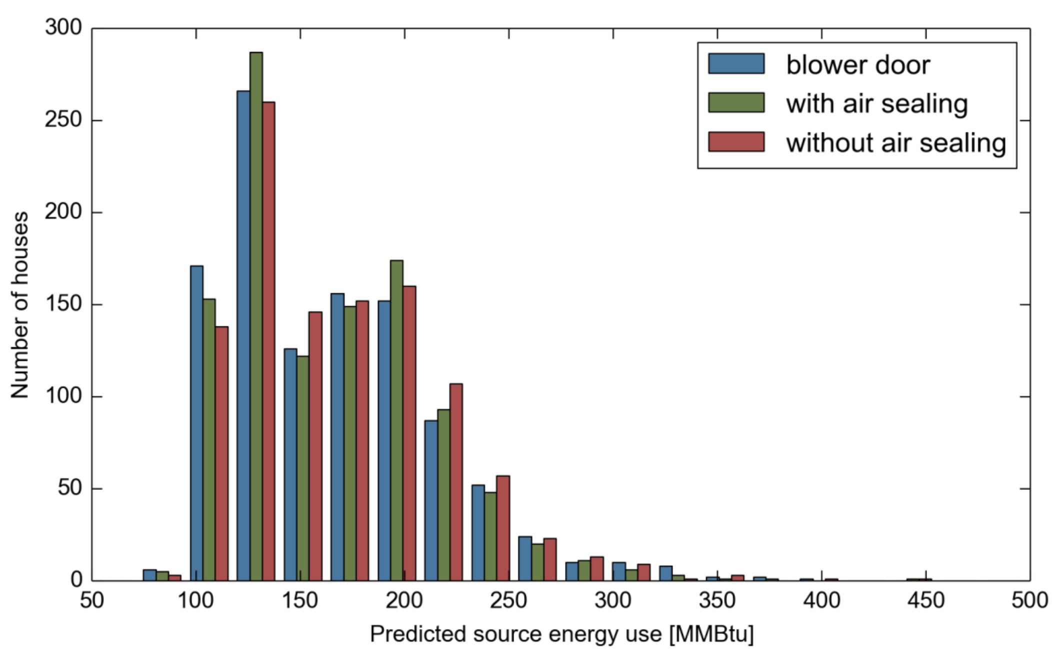

Figure 18. Distribution of HEScore-predicted source energy use for BAFDR homes generated using three input scenarios for whole-house leakage

Table 15 shows the mean value of the source energy for each of the three whole-house infiltration input scenarios. The averages follow the same trend as with the HEScore homes, except that overall predicted energy use is lower in the BAFDR homes. The average change in predicted energy use when changing from "air sealed" to "not air sealed" is on the order of 2%–3%.

Table 15. Average HEScore-Predicted Source Energy Use for Each Infiltration Assumption for BAFDR

	CFM50	Sealed	Unsealed
Mean Source Energy (MMBtu/yr)	166	165	168

Figure 19 shows the difference in predicted source energy use generated by HEScore using the quantitative and qualitative inputs for whole-house air infiltration. On average, when compared to the predictions stemming from quantitative input, the source energy use decreases 1 MMBtu/yr when the sealed qualitative input is used and increases 2 MMBtu/yr when the unsealed qualitative input is used. Also, there is a standard deviation of 11 MMBtu/yr, meaning that although the average difference seen from selecting a quantitative versus a qualitative input is small on an individual house, it can be—and often is—much larger.

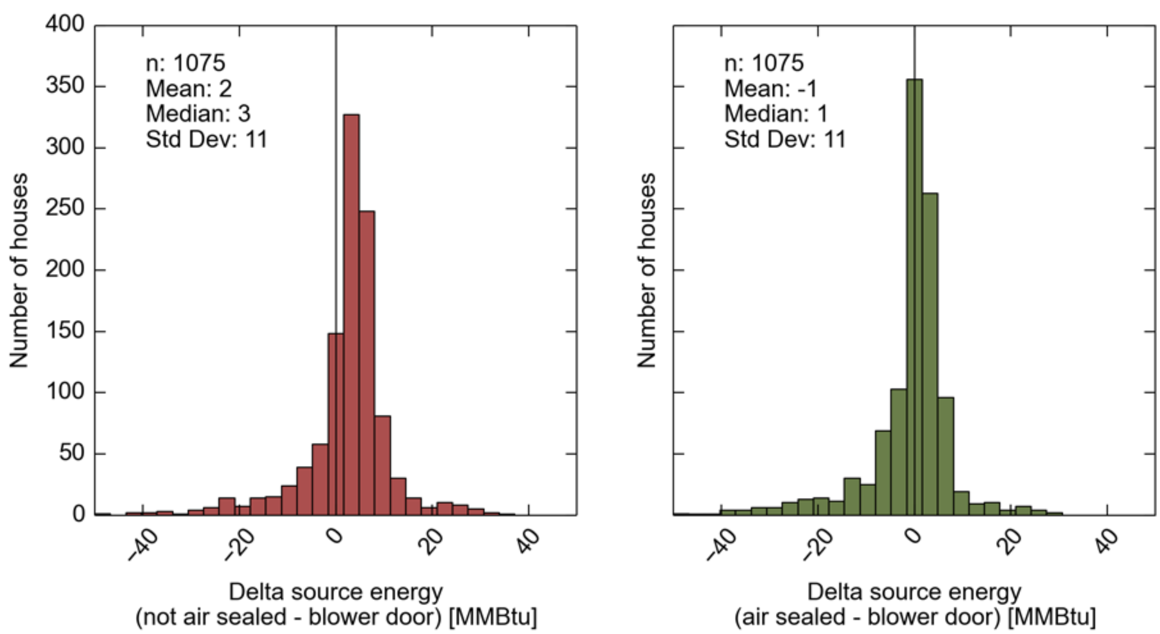

Figure 19. Distribution of differences in HEScore-predicted source energy use using qualitative and quantitative input for whole-house leakage for BAFDR homes

Once again, it is important to review these same differences in terms of the score. Figure 20 illustrates the effect on the score of selecting the "not air sealed" and "air sealed" qualitative inputs, respectively, compared to using the blower door measurement. In both cases, 95% of the homes score within ±1 bin.

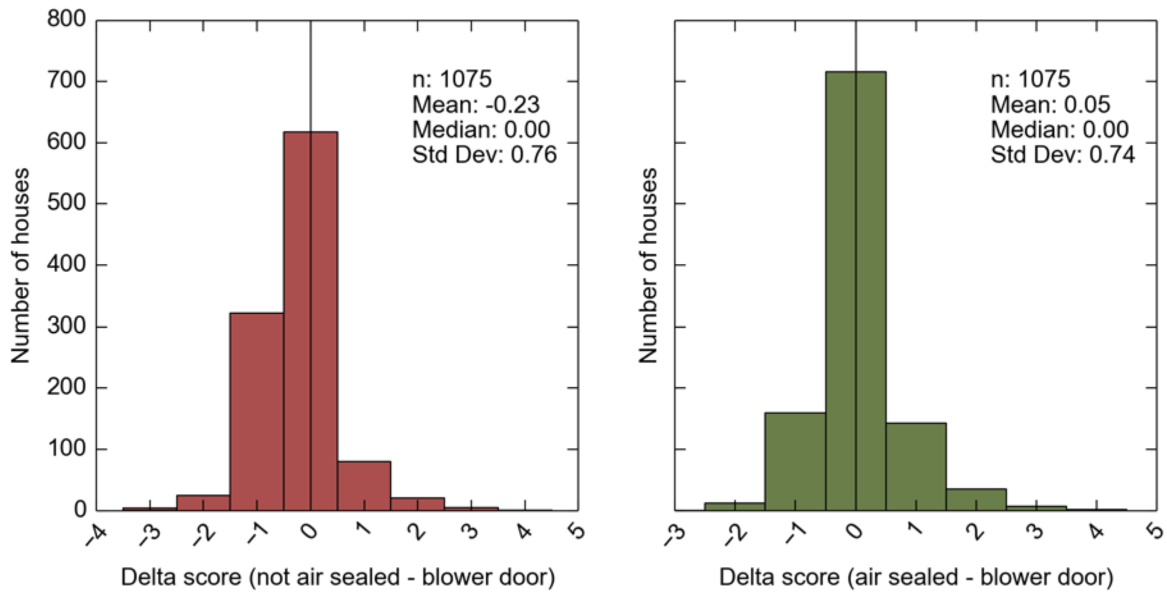

Figure 20. Distribution of differences in score using qualitative and quantitative input for whole-house leakage for BAFDR homes

4.4 Infiltration Study Conclusions

From these data and the data from the HEScore 2012 simulations, the effect of the whole-house leakage on overall average energy use appears to be small, whether it is measured with a blower door or inferred from data on populations of houses based on the characteristics of a particular house. However, infiltration plays a significant role in whole-house energy use. The role of the "air sealed" choice is only one factor in the infiltration estimate algorithm used in HEScore. This analysis measures the impact of that choice only, versus using a blower door. Other factors used in the estimate (floor area, age of home, etc.) remain constant. This indicates that the algorithm, including *all* the factors it considers, is doing a relatively good job of estimating whole-house leakage compared to measurements.

This analysis also indicates that the sensitivity to the air sealing input compared to the blower door measurements is slight, which means that getting it "wrong" in the audit has a minimal effect on the predicted source energy (and therefore score). For both datasets, the score remained within ±1 bin 95% of the time when either qualitative input was selected. From this we can see that the accuracy of the infiltration assumptions is adequate for the purposes of the score when a blower door test is not performed.

5 Conclusions

When comparing HEScore 2012 to HEScore 2014 predicted energy use in Section 2.1, the version to version differences were small. Because the goal of HEScore 2014 was primarily to update the architecture—and not necessarily to significantly change the modeling—this is not a surprising result.

The same trends were observed in HEScore 2012 and HEScore 2014 when comparing model-predicted energy use to energy use from weather-normalized utility billing data. Both electricity and natural gas use saw slightly better r^2 values for the linear regression. However, when we applied an analysis of variance to the regressions, the improved r^2 was revealed to be statistically insignificant. There does appear to be better agreement with natural gas use for HEScore 2014, as indicated by the significantly lower standard deviation in the difference (refer to Figure 8). This is presumably due to the modeling changes that were made.

In the evaluation of the defaults for whole-house infiltration and their effect on predicted energy use, we found that on average the total energy effect of selecting a qualitative input versus performing a blower door test was on the order of 2%–3%. This indicates that the specific qualitative input for whether air sealing was present had a slight effect, and that the overall estimate of infiltration when a blower door measurement isn't specified is accurate.

As previously mentioned, the scores could not be calculated from the utility bill data because we could not extract energy use for heating, cooling and domestic hot water only. Nevertheless, there was a reduction in delta natural gas use for HEScore 2014 and two electric heating types (furnaces and baseboards) were no longer significant contributors to the delta electricity models. These potential improvements in energy prediction should also reduce the probability of assigning completely different scores to similar houses.

References

ASHRAE (2002). *ASHRAE Guideline 14-2002: Measurement of Energy and Demand Savings*. Atlanta, GA: American Society of Heating, Refrigerating and Air-Conditioning Engineers.

BPI (2011). *BPI-2400-S-2011 Standardized Qualification of Whole House Energy Savings Estimates.* Malta, NY: Building Performance Institute, Inc. Accessed March 27, 2014: http://www.bpi.org/tools_downloads.aspx?selectedTypeID=1&selectedID=115

Deru, M.; Torcellini, P. (2007). *Source Energy and Emission Factors for Energy Use in Buildings.* Golden, CO: National Renewable Energy Laboratory. NREL/TP-550-38617.

Earth Advantage Institute, Conservation Services Group (EAI/CSG). (2009). *Energy Performance Score (EPS) 2008 Pilot Findings & Recommendations Report*. Energy Trust of Oregon.

ENERGY STAR. (2013). *Portfolio Manager Technical Reference: Source Energy.* Accessed March 27, 2014: http://www.energystar.gov/buildings/tools-and-resources/portfolio-manager-technical-reference-source-energy

Hassel, S.; Blasnik, M.; Hannas, B. (2009). *Houston Home Energy Efficiency Study.* Raleigh, NC: Advanced Energy.

Lynch, S. (2003). “Outliers and Influential Cases (Soc 504)” Princeton University. Accessed January 30, 2014: http://www.princeton.edu/~slynch/soc504/outliers.pdf

Mills, E. (2008). “The Home Energy Saver” Lawrence Berkeley National Laboratory. Accessed March 30, 2014: http://evanmills.lbl.gov/pubs/pdf/home-energy-saver.pdf

Montgomery, D.C. (1997). *Design and Analysis of Experiments*. 4th edition. New York: John Wiley & Sons, pp. 556-557.

Nettleton, G.; Edwards, J., (2012), *Data Collection-Data Characterization Summary*, NorthernSTAR Building America Partnership, Building Technologies Program, USDOE.

Pennsylvania State University (2014). “17.1 – Polynomial Regression.” Accessed June 6, 2014: https://onlinecourses.science.psu.edu/stat501/print/book/export/html/235

Pigg, S.; Nevius, M., (2000). *Energy and Housing in Wisconsin: A Study of Single-Family Owner-Occupied Homes.* Energy Center of Wisconsin, 199-1.

Princeton University (2007). “Interpreting Regression Output.” Accessed January 13, 2014: http://dss.princeton.edu/online_help/analysis/interpreting_regression.htm

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/

Roberts, D.; Merket, N.; Polly, B.; Heaney, M.; Casey, S.; Robertson, J. (2012). *Assessment of the U.S. Department of Energy’s Home Energy Scoring Tool.* Golden, CO: National Renewable Energy Laboratory, NREL/TP-5500- 54074.

Rodríguez, G. (2013). “Introducing R.” Princeton University, Accessed January 13, 2014: http://data.princeton.edu/R/linearModels.html

Appendix A: Use of BAFDR in Scoring Tool Assessment

This appendix describes the state and use of the BAFDR in the HEScore assessment project. To maintain consistency with the previous HEScore report, only data available and used in that report were again simulated in the newer version to allow an unbiased comparison.

A.1 BAFDR Data Collection

NREL has been working to obtain historical datasets containing robust, research-grade characteristics data coupled with utility billing data. It has proven to be a challenging process; there are a number of obstacles, including a paucity of data and legal issues related to customer privacy. Nevertheless, NREL managed to accumulate useful data.

To start with, the BAFDR team focused on datasets that were available in the form of REM/Rate software input files. REM/Rate has been the most widely used home energy rating software for more than a decade, and thus REM/Rate input files are a relatively common format for existing datasets. It is used for home energy ratings (primarily supporting ENERGY STAR Qualified Homes Program), state and utility efficiency programs, and income-qualified weatherization programs. The software is used to evaluate new construction and retrofits, and its input files contain fairly detailed building characteristics. The software also has a feature that allows batch exporting of input files into a relational database. This REM/Rate database format became the starting point for the BAFDR.

Initial datasets that were collected and aggregated into the BAFDR included:

- ***Oregon Energy Performance Pilot Study (EAI/CSG) data.***[16] A mix of 190 newer and older homes audited in 2008 and located in Portland and Bend.
- ***Wisconsin Housing Characterization Study (Pigg and Nevius 2000) data.***[17] Collected in 1991, a mix of 299 new and existing homes located throughout Wisconsin.
- ***Houston utility study.*** A subsample of 82 homes statistically derived from a utility program evaluation conducted by Hassel et al. (2009) involving a sample of 226,000 homes built from 2002 through 2007.
- ***Advance Energy's SystemVision homes.***[18] Four hundred ninety high-efficiency homes in North Carolina and Tennessee receiving full home energy ratings.
- ***Building America Energy Audit Assessment Housing Characterization (Nettleton and Edwards 2012) data.*** One hundred twenty-five older retrofit-candidate homes in Minnesota and Wisconsin receiving full home energy ratings.
- ***EPA ENERGY STAR Qualified Homes evaluation study.***[19] Seventy-five ENERGY STAR-qualified homes in Minnesota and Wisconsin.

[16] This dataset was generously provided by Earth Advantage Institute. In particular, the authors would like to thank David Heslam of Earth Advantage and Diane Ferrington of The Energy Trust of Oregon.
[17] The authors would like to thank Scott Pigg of the Energy Center of Wisconsin for providing this dataset.
[18] The authors would like to thank Jonathan Coulter of Advanced Energy for providing this dataset.
[19] The authors would like to thank Mat Gates of Residential Science Resources, LLC for providing this dataset.

These datasets resulted in a total of 1,183 homes in the BAFDR after they were processed as described in the next section.

A.2 BAFDR Data Processing

Housing characteristics data across these initial datasets were consistent and required no processing. The data were aggregated using REM/Rate batch database capabilities.

Utility billing data had to be processed for a subset of the datasets. To compare predicted energy use to measured energy use, the utility billing data had to be normalized for differences between the weather during the utility billing period and the climate data used in the energy simulation. HEScore 2012 used TMY2 climate data – TMY data for the 30-year period from1961–1990. HEScore 2014 uses TMY3 climate data – an update to TMY2 which provides more weather sites and covers the years 1991–2005. Comparisons to measured use for HEScore 2012 were normalized to TMY2. Comparisons to measured use for HEScore 2014 were normalized to TMY3.

The utility billing data were weather normalized following the procedures outlined in *BPI-2400-S-2011 Standardized Qualification of Whole House Energy Savings Estimates* (BPI 2011) and *ASHRAE Guideline 14-2002, Annex D: Regression Techniques* (ASHRAE 2002). Historical daily average temperatures were obtained from the nearest weather stations with sufficient data managed by the National Climatic Data Center.[20]

Three-point heating models were created with natural gas billing data. Three-point heating and cooling models and a five point heating and cooling model were created with electricity billing data. For each model, goodness-of-fit criterion were established as an adjusted coefficient of variation-RMSE ≥ 20%. For cases with insufficient goodness-of-fit, the model coefficients were not used and an annual average energy use was calculated from the utility billing data. For electricity modeling when more than one model met the goodness-of-fit criterion, the one with the highest R^2 was selected. For regression models with sufficient goodness-of-fit, the model coefficients were applied to the calculated degree-days from TMY2 or TMY3 climate data, calculating the normalized annual consumption.

Utility billing data could not be obtained for the Oregon dataset, but climate-normalized annual energy uses for electricity and natural gas were provided by Earth Advantage Institute (2009).

A.3 Translation of BAFDR Data to Software Inputs

For this project, “interpreters” were written in the Python programming language to translate BAFDR data to software inputs for HEScore. Detailed descriptions of these data translations are presented in Appendix B.

A.4 Processing BAFDR Results

HEScore source energy was calculated using national average site-to-source multipliers from ENERGY STAR Portfolio Manager for HEScore 2014 and Deru and Torcellini (2007) for HEScore 2012. Only homes for which successful results were returned from both versions of HEScore and sufficient utility bills could be obtained were included in the analysis, yielding a total of 1,085 homes.

[20] Web url: http://www.ncdc.noaa.gov/oa/climate/isd/

Appendix B: Translation of BAFDR Data to HEScore Inputs

The original translation of BAFDR data to HEScore was for the 2012 version of the API. In HEScore 2014 the API was modified. To run the BAFDR homes through HEScore 2014, the HEScore 2012 inputs obtained as described below were mapped into the newer version of the API using documentation from the HEScore development team.[21]

In translating inputs from BAFDR to HEScore, the goal was to provide an "out-of-the-box" set of inputs to HEScore; i.e., to use HEScore as closely as possible to the way an assessor entering data would use it. The following is an explanation of each input to HEScore and how it was derived from BAFDR data. The HEScore input variables are identified by *italicsAndCamelCase*.

B.1 General

- *zipcode* − The ZIP code from the house address in REM/Rate was used. For homes without a ZIP code recorded, the ZIP code was looked up for the city and state using postal code data from the Geonames Project.[22]
- *year* − The year the house was built.

B.2 House Shape and Size

- *houseOrientation* − In the BAFDR data the orientation of the windows is known. Overall house orientation is not known. House orientation was estimated by taking the side of the house with the greatest window area and assuming that it is the back of the house.
- *storiesAboveGround* − Number of stories above ground.
- *floorArea* − Total conditioned floor area.
- *ceilingHeight* − Average ceiling height was calculated by dividing the total conditioned volume by the floor area. The result was rounded to the nearest foot.

B.3 Number of Bedrooms

- *numberBedrooms* − Number of bedrooms. This number was collected as part of the Home Energy Score inputs.

B.4 Airtightness

- *airLeakage50ip* − The blower door measurement, measured in CFM50. If the infiltration was measured in air changes per hour at 50 Pascals (ACH50) it was converted to CFM50 using

  $$CFM50 = ACH50 \cdot Volume/60$$

  If the infiltration was measured in CFM25, it was converted to CFM50 using

  $$CFM50 = CFM25 \cdot (50/25)^{0.65}$$

[21] The Scoring Tool API comparison table used in this mapping is available at http://goo.gl/1EoSVN
[22] http://www.geonames.org

- *airSealingPresent* – If the infiltration units were not measured in CFM50, CFM25, or ACH50, the infiltration measurement was omitted in HEScore and the house was marked as "not air sealed."

B.5 Foundation and Floor

- *foundationType* – The mapping in Table 16 was used to convert foundation type from BAFDR to HEScore compatible foundation types:

Table 16. Mapping Used To Convert Foundation Types

BAFDR Foundation Type	HEScore Foundation Type
Slab	Slab-on-grade Foundation
Open Crawlspace	Vented Crawlspace
Enclosed Crawlspace	'Unvented Crawlspace' if the crawlspace type in REM/Rate is 'Unvented'. 'Vented Crawlspace' if the crawlspace type in REM/Rate is 'Vented' or 'Operable Vents'
Conditioned Basement	Conditioned Basement
Unconditioned Basement	Unconditioned Basement
Conditioned Crawlspace	Unvented Crawlspace

  If the foundation type was "more than one foundation type" the foundation wall or slab with the greatest perimeter to ambient/ground was specified as the foundation type in HEScore.

- *foundationSideInsulationRValue* – If the foundation type was a slab the R-value of the slab insulation with the greatest area was returned and rounded to the nearest of R-0 or R-5 (the only options for slab insulation in HEScore). If the foundation type was a basement or crawlspace, the sum of the exterior, cavity, and interior rigid insulation R-values of the foundation wall with the greatest area was returned and rounded to the nearest of R-0, R-11, and R-19 (the only options for basement and crawlspace foundation insulation in HEScore).

- *floorConstruction* – Insulation level of the floor above the basement or crawlspace. This was calculated by (1) identifying the largest frame floor between conditioned space and the open crawlspace, enclosed crawlspace, conditioned basement, or unconditioned basement depending on the foundation type, and then (2) adding the cavity and continuous insulation R-values and rounding to the closest of R-0, R-11, R-13, R-15, R-19, R-21, R-25, R-30, and R-38 (the only available options in HEScore). Other foundation types were assumed to have zero floor insulation.

B.6 Walls

- *wallsSameAllSides* – Indicates if different wall types are described on each side of the house or if one wall type is used to describe all of the exterior walls. This input was always set to *true* to specify one wall type for the whole house.

- *wallConstructionFront* – Represents the construction of all the walls in this case because the *wallsSameAllSides* input above was set to *true*. Wood stud walls were input with the cavity insulation rounded to the closest of the available R-values in HEScore (R-0, R-3, R-7, R-11, R-13, R-15, R-19, R-21). Any continuous insulation on a wood stud wall was assumed to be 1-in. expanded polyethylene sheathing because that was the only available option in the HEScore interface. All siding on wood stud walls was assumed to be wood siding. Structural brick walls and concrete block walls, where applicable, were also translated accordingly. The R-values of the continuous insulation were rounded to the nearest values available in the HEScore input (R-0, R-5, R-10 and R-0, R-3, R-6, respectively).

B.7 Doors and Windows

- *windowArea(Front|Back|Left|Right)* – Window area was totaled for each side and returned. For windows facing a direction between two sides (e.g., facing northeast), the window area was divided between the two sides (1/2 area facing north, 1/2 area facing east).
- *windowUValue(Front|Back|Left|Right)* – An area-weighted average U-value was calculated for each window direction.
- *windowSolarGain(Front|Back|Left|Right)* – An area-weighted average SHGC was calculated for each window direction.
- *windowShade(Front|Back|Left|Right)* – An area-weighted average interior shading factor was calculated for each window direction. A qualitative input was then selected that most closely matched the values in Table 9, p. 89, of the HEScore documentation.[23]

B.8 Skylights

- *skylightsPresent* – For homes with any skylight area, this was set to "true." For homes with no skylight area, this was set to "false" and no other skylight inputs were specified.
- *skylightType* – A skylight from the HEScore library was selected that most closely matched the area weighted average U-value and SHGC of the skylights.
- *skylightArea* – Total skylight area.

B.9 Attic and Roof

Only one ceiling can be specified in HEScore. The ceiling with the greatest area for the house in the BAFDR was selected. All others were ignored.

- *atticType* – 'Vaulted' ceiling in BAFDR was translated to a 'Cathedral Ceiling' in HEScore. 'Attic' in BAFDR was translated to 'Unconditioned Attic' in HEScore.
- *roofConstruction* – This input is a code that represents the roofing material, roof insulation (not attic floor insulation), and the presence of a radiant barrier. All roofs were assumed to have composition shingles. For roofs with vaulted ceilings, the insulation indicated in the BAFDR was assumed to be in the roof cavity and the nearest R-value (R-0, 11, 13, 15) available for roof insulation in HEScore was selected. If a roof had a radiant barrier and no roof insulation, a radiant barrier was selected in HEScore. For roofs above an unfinished attic, no insulation was specified in the *roofContruction*, but was instead specified on the attic floor in *ceilingConstruction*.

[23] http://evanmills.lbl.gov/pubs/pdf/home-energy-saver.pdf

- *ceilingConstruction* – This indicates the R-value of insulation on the attic floor. For roofs above an unfinished attic, the R-value from the BAFDR was assumed to be on the attic floor and the nearest option for attic floor insulation in HEScore was selected (R-0, 3, 9, 11, 19, 21, 25, 30, 38, 49, 60). For roofs above a vaulted ceiling, no insulation was specified in the *ceilingConstruction*.
- *roofAbsorptivityValue* – The roof absorptivity was translated from qualitative to a quantitative value using the values in the HEScore documentation (Table 17):

Table 17. Translation of Roof Absorptivity From Qualitative to Quantitative

REM/Rate Roof Color	Absorptance
Light	0.60
Medium	0.75
Dark	0.90
Reflective	0.20

B.10 Ducts and Pipes

- *ductLocation* – Duct location was translated from BAFDR to HEScore inputs according to the following mapping (Table 18):

Table 18. Translation of Duct Location From BAFDR to HEScore

REM/Rate Duct Location	HEScore Duct Location
Open Crawlspace	Vented crawlspace
Enclosed Crawlspace	Unconditioned basement or unvented crawlspace
Conditioned Crawlspace	Conditioned space
Unconditioned Basement	Unconditioned basement or unvented crawlspace
Conditioned Basement	Conditioned space
Attic, Under Insulation	Conditioned space
Attic, Exposed	Unconditioned attic
Conditioned Space	Conditioned space
Wall With No Top Plate	Unknown/not applicable
Garage	Unknown/not applicable
Exterior Wall	Unknown/not applicable
Floor Cavity Over Garage	Unknown/not applicable
Under Slab Floor	Conditioned space

"Unknown/not applicable" is not available as an option in HEScore 2014, so houses with that duct location were not simulated and were therefore not included in the updated analysis.

- *ductInsulationPresent* – For homes in the BAFDR where the R-value of the ducts in the primary duct system was greater than R-1, the value for this input was set to "true."
- *ductSealingPresent* – The air handler flow rate in cubic feet per minute (cfm) was estimated by assuming air conditioners and heat pumps in cooling mode operate at 400 cfm/ton and furnaces and heat pumps in heating mode operate at 275 cfm/ton. If necessary, the measured duct leakage was converted from CFM50 to CFM25. If the duct leakage was not measured in CFM50 or CFM25, the HEScore default of "unsealed" was assumed. The measured duct leakage in CFM25 was divided by the estimated total CFM

to obtain a percent leakage. For homes with CFM25 duct leakage ≤22.5% of air handler flow, the ducts were assumed to be "sealed."

- *hwFromBoiler* – For homes in the BAFDR where the water heating equipment that handles the largest percentage of the water heating load handles some portion of the space heating load and is a gas or oil boiler, HEScore specified that the boiler provides hot water. Otherwise, it was input into HEScore as having separate hot water and space heating equipment. If it was determined that the boiler provides hot water and the boiler's hot water tank volume in the BAFDR is greater than zero, HEScore specified that the boiler has an indirect tank providing hot water; if not, the boiler was specified as having a tankless coil to provide hot water.

B.11 Heating Equipment

For each house in the BAFDR, the heating system that handles the greatest percentage of the heating load was selected. All other heating equipment was ignored.

- *heatingType* – The type of primary heating equipment was translated from BAFDR.
- *heatingEfficiency* – The heating efficiency of the heating equipment was input from the BAFDR into heating seasonal performance factor for heat pumps and AFUE for anything else. In cases where the efficiency of a heat pump was specified in coefficient of performance, it was converted to heating seasonal performance factor by dividing by 0.293.
- *heatingCapacity* – The heating capacity was converted from kBtu/h to Btu/h and input into HEScore.

B.12 Cooling Equipment

For each home in the BAFDR, the mechanical equipment that handles the greatest percentage of the cooling load was selected. Any other cooling equipment was ignored.

- *coolingType* – The type of cooling system. All homes in the BAFDR have either central air conditioning, electric heat pumps, or no air conditioning.
- *coolingEfficiency* – Seasonal energy efficiency ratio was entered for central air conditioners and heat pumps.

B.13 Water Heating

For each house in the BAFDR, the mechanical equipment that handles the greatest percentage of the hot water load was selected. All other hot water equipment was ignored.

- *hwFuel* – The fuel type of the primary water heater was translated from BAFDR.
- *hwEnergyFactor* – The energy factor of the water heater was translated from BAFDR.

Appendix C: Additional Statistical Model Information

This appendix includes additional information and background materials for the statistical models developed in Section 3. Section C.1 covers outlier identification and removal. A complete list of coded variables and descriptions is included in Section C.2.

C.1 Outlier Identification and Removal

Often datasets contain observations that, even after modeling, vary considerably from most observations. These outliers may bias the MLR models. There are only guidelines for detecting outliers rather than established rules (Lynch 2003). The approach for this study was to first use all observations to create an intermediate reduced model. The intermediate reduced model is done using a stepwise regression method (stepAIC, available in the R language) after creating a model with all available variables. The residuals (observations minus MLR predicted) are examined and any observation outside plus or minus three standard deviations of the residuals is considered as an outlier. For most of the models, only two to four observations were found to be outliers. These observations were removed before further reduction in the MLR models.

C.2 HEScore Variables Tested in Statistical Analysis

This section lists the HEScore variables and coding used for the statistical analysis. In addition to inputs listed in Appendix B, base 65°F HDDs, base 65°F CDDs, and a few combined variables were included in the analysis. Table 19 gives a complete list of HEScore variables and descriptions.

Table 19. Variables Used for Statistical Analysis and Descriptions

Number	Coded Variable	Original Variables and Inputs	Original Variable Description
1	blower_door_test	blower_door_test	Indication (Yes/No) of blower door test
2	ducts_insulated	ducts_insulated	Indication (Yes/No) of ducts insulated
3	ducts_sealed	ducts_sealed	Indication (Yes/No) of ducts sealed
4	cooling_present	cooling_present	Indication (Yes/No) of cooling system
5	C_orientation_north	orientation	Orientation = north
6	C_orientation_north_east		Orientation = north_east
7	C_orientation_north_west		Orientation = north_west
8	C_orientation_south		Orientation = south
9	C_orientation_south_east		Orientation = south_east
10	C_orientation_south_west		Orientation = south_west
11	C_orientation_west		Orientation = west
12	C_RC_rfwf11co	roof_assembly_code	Roof assembly code = rfwf11co
13	C_RC_rfwf15co		Roof assembly code = rfwf15co
14	C_RC_rfrb00co		Roof assembly code = rfrb00co
15	C_CC_ecwf00	ceiling_assembly_code	Ceiling assembly code = ecwf00
16	C_CC_ecwf03_09		Ceiling assembly code = ecwf03 through 09
17	C_CC_ecwf11		Ceiling assembly code = ecwf11
18	C_CC_ecwf19		Ceiling assembly code = ecwf19
19	C_CC_ecwf21		Ceiling assembly code = ecwf21

Number	Coded Variable	Original Variables and Inputs	Original Variable Description
20	C_CC_ecwf25		Ceiling assembly code = ecwf25
21	C_CC_ecwf30		Ceiling assembly code = ecwf30
22	C_CC_ecwf49_60		Ceiling assembly code = ecwf49 through 60
23	C_AT_cath_ceiling	roof_type	Attic type = cathedral ceiling
24	C_FT_slab	foundation_type	Foundation type = slab
25	C_FT_uncond_base		Foundation type = unconditioned basement
26	C_FT_unvent_crawl		Foundation type = unvented crawlspace
27	C_FT_vent_crawl		Foundation type = vented crawlspace
28	C_FC_efwf11ca_15ca	floor_assembly_code	Floor assembly code = efwf11ca through 15ca
29	C_FC_efwf19ca_21ca		Floor assembly code = efwf19ca through 21ca
30	C_FC_efwf25ca		Floor assembly code = efwf25ca
31	C_FC_efwf30ca_38ca		Floor assembly code = efwf30ca through 38ca
32	C_FC_none		Floor assembly code = none
33	C_HT_EBB	heating_type and heating_fuel_primary	Heating type/fuel = electric baseboard
34	C_HT_GBL		Heating type/fuel = electric baseboard
35	C_HT_EFN		Heating type/fuel = electric central furnace
36	C_HT_EHP		Heating type/fuel = electric heat pump
37	C_HT_GWF		Heating type/fuel = gas wall furnace
38	C_HF_elec	heating_fuel_primary	Heating fuel = electric
39	C_CT_ehp	cooling_type	Cooling type = electric heat pump
40	C_CT_cac		Cooling type = central air conditioner
41	C_HW_EST	dhw_type and dhw_fuel_primary	Hot water/fuel = electric storage tank
42	C_HW_GTL		Hot water/fuel = gas tankless
43	C_hwFuel_elec	dhw_fuel_primary	Hot water fuel = electric
44	C_WC_ewps15wo	wall_assembly_code	Wall assembly code = ewps15wo
45	C_WC_ewps19wo_21wo		Wall assembly code = ewps19wo through 21wo
46	C_WC_ewwf00wo_00br		Wall assembly code = ewwf00wo through 00br
47	C_WC_ewwf03wo_07wo		Wall assembly code = ewwf03wo through 07wo
48	C_WC_ewwf11wo		Wall assembly code = ewwf11wo
49	C_WC_ewwf13wo		Wall assembly code = ewwf13wo
50	C_WC_ewwf15wo		Wall assembly code = ewwf15wo
51	C_WC_ewwf21wo		Wall assembly code = ewwf21wo
52	C_WC_ewps00wo		Wall assembly code = ewps00wo
53	C_WC_ewps11wo		Wall assembly code = ewps11wo
54	C_WC_ewps13wo		Wall assembly code = ewps13wo

Number	Coded Variable	Original Variables and Inputs	Original Variable Description
55	C_DL_none	duct_location	Duct location = none
56	C_DL_uncond_attic		Duct location = unconditioned attic
57	C_DL_uncond_basement		Duct location = unconditioned basement
58	C_state_OR	state	State = Oregon
59	C_state_MN		State = Minnesota
60	C_state_NC		State = North Carolina
61	C_state_TX		State = Texas
62	C_number_bedrooms	number_bedrooms	Number of bedrooms
63	C_num_floor_above_grade	num_floor_above_grade	Stories above ground level
64	C_floor_to_ceiling_height	floor_to_ceiling_height	Floor to ceiling height (feet)
65	C_conditioned_floor_area	conditioned_floor_area	Conditioned floor area (ft^2)
66	C_envelope_leakage	envelope_leakage	Envelope leakage
67	C_roof_absorptance	roof_absorptance	Roof absorptance
68	C_skylight_area	skylight_area	Sum or skylight area
69	C_total_window_area	total_window_area	Sum of total area from each side
70	C_wtmean_window_u_value	wtmean_window_u_value	Mean window U value weighted by window area on each side
71	C_wtmean_window_shgc	wtmean_window_shgc	Mean window SHGC value weighted by window area on each side
72	C_HDD65	HDD65	Heating degree days (base 65°F)
73	C_CDD65	CDD65	Cooling degree days (base 65°F)
74	C_age_years	year_built	Age years = 2011 - year_built
75	C_Fnd_Flr_R_Value	Fnd_Flr_R_Value	Foundation or floor R value ($ft^2·°F·h/Btu$)
76	C_Roof_R_Value	Roof_R_Value	Roof or ceiling R value ($ft^2·°F·h/Btu$)
77	C_Wall_R_Value	Wall_R_Value	Wall R value ($ft^2·°F·h/Btu$)
78	C_NG_HeatingEff	heating_efficiency	Natural gas heating efficiency
79	C_CoolingEff	cooling_efficiency	Cooling efficiency

Table 20 gives details on the coding used. As mentioned in Section 4.2, some inputs such as floor area are numeric, but many inputs use codes to describe various types of building construction components (Mills 2008). Wall types, roof types, foundation types, and many other components that make up a building have separate codes. The "Natural Gas" column indicates that the variable was tested in the natural gas models. The "Electricity" column indicates that the variable was tested in electricity models.

Table 20. Coding Details for Variables Used for Statistical Analysis

Number	Coded Variable	Variable Type	Coding	Control for Binary	Natural Gas	Elec-tricity
1	blower_door_test	Binary	Yes = 1, No = 0	N/A	x	x
2	ducts_insulated	Binary	Yes = 1, No = 0	N/A	x	x
3	ducts_sealed	Binary	Yes = 1, No = 0	N/A	x	x
4	cooling_present	Binary	Yes = 1, No = 0	N/A	x	x
5	C_orientation_north	Binary	Yes = 1, No = 0	East	x	x
6	C_orientation_north_east	Binary	Yes = 1, No = 0	East	x	x
7	C_orientation_north_west	Binary	Yes = 1, No = 0	East	x	x

Number	Coded Variable	Variable Type	Coding	Control for Binary	Natural Gas	Elec-tricity
8	C_orientation_south	Binary	Yes = 1, No = 0	East	x	x
9	C_orientation_south_east	Binary	Yes = 1, No = 0	East	x	x
10	C_orientation_south_west	Binary	Yes = 1, No = 0	East	x	x
11	C_orientation_west	Binary	Yes = 1, No = 0	East	x	x
12	C_RC_rfwf11co	Binary	Yes = 1, No = 0	rfwf00co	x	x
13	C_RC_rfwf15co	Binary	Yes = 1, No = 0	rfwf00co	x	x
14	C_RC_rfrb00co	Binary	Yes = 1, No = 0	rfwf00co	x	x
15	C_CC_ecwf00	Binary	Yes = 1, No = 0	ecwf38	x	x
16	C_CC_ecwf03_09	Binary	Yes = 1, No = 0	ecwf38	x	x
17	C_CC_ecwf11	Binary	Yes = 1, No = 0	ecwf38	x	x
18	C_CC_ecwf19	Binary	Yes = 1, No = 0	ecwf38	x	x
19	C_CC_ecwf21	Binary	Yes = 1, No = 0	ecwf38	x	x
20	C_CC_ecwf25	Binary	Yes = 1, No = 0	ecwf38	x	x
21	C_CC_ecwf30	Binary	Yes = 1, No = 0	ecwf38	x	x
22	C_CC_ecwf49_60	Binary	Yes = 1, No = 0	ecwf38	x	x
23	C_AT_cath_ceiling	Binary	Yes = 1, No = 0	vented attic	x	x
24	C_FT_slab	Binary	Yes = 1, No = 0	conditioned basement	x	x
25	C_FT_uncond_base	Binary	Yes = 1, No = 0	conditioned basement	x	x
26	C_FT_unvent_crawl	Binary	Yes = 1, No = 0	conditioned basement	x	x
27	C_FT_vent_crawl	Binary	Yes = 1, No = 0	conditioned basement	x	x
28	C_FC_efwf11ca_15ca	Binary	Yes = 1, No = 0	efwf00ca	x	x
29	C_FC_efwf19ca_21ca	Binary	Yes = 1, No = 0	efwf00ca	x	x
30	C_FC_efwf25ca	Binary	Yes = 1, No = 0	efwf00ca	x	x
31	C_FC_efwf30ca_38ca	Binary	Yes = 1, No = 0	efwf00ca	x	x
32	C_FC_none	Binary	Yes = 1, No = 0	efwf00ca	x	x
33	C_HT_EBB	Binary	Yes = 1, No = 0	natural gas central furnace	x	x
34	C_HT_GBL	Binary	Yes = 1, No = 0	natural gas central furnace	x	x
35	C_HT_EFN	Binary	Yes = 1, No = 0	natural gas central furnace	x	x
36	C_HT_EHP	Binary	Yes = 1, No = 0	natural gas central furnace	x	x
37	C_HT_GWF	Binary	Yes = 1, No = 0	natural gas central furnace	x	x
38	C_HF_elec	Binary	Yes = 1, No = 0	natural gas	x	x
39	C_CT_ehp	Binary	Yes = 1, No = 0	none	x	x
40	C_CT_cac	Binary	Yes = 1, No = 0	none	x	x
41	C_HW_EST	Binary	Yes = 1, No = 0	natural gas storage	x	x
42	C_HW_GTL	Binary	Yes = 1, No = 0	natural gas storage	x	x
43	C_hwFuel_elec	Binary	Yes = 1, No = 0	natural gas	x	x

Number	Coded Variable	Variable Type	Coding	Control for Binary	Natural Gas	Elec-tricity
44	C_WC_ewps15wo	Binary	Yes = 1, No = 0	ewwf19wo	x	x
45	C_WC_ewps19wo_21wo	Binary	Yes = 1, No = 0	ewwf19wo	x	x
46	C_WC_ewwf00wo_00br	Binary	Yes = 1, No = 0	ewwf19wo	x	x
47	C_WC_ewwf03wo_07wo	Binary	Yes = 1, No = 0	ewwf19wo	x	x
48	C_WC_ewwf11wo	Binary	Yes = 1, No = 0	ewwf19wo	x	x
49	C_WC_ewwf13wo	Binary	Yes = 1, No = 0	ewwf19wo	x	x
50	C_WC_ewwf15wo	Binary	Yes = 1, No = 0	ewwf19wo	x	x
51	C_WC_ewwf21wo	Binary	Yes = 1, No = 0	ewwf19wo	x	x
52	C_WC_ewps00wo	Binary	Yes = 1, No = 0	ewwf19wo	x	x
53	C_WC_ewps11wo	Binary	Yes = 1, No = 0	ewwf19wo	x	x
54	C_WC_ewps13wo	Binary	Yes = 1, No = 0	ewwf19wo	x	x
55	C_DL_none	Binary	Yes = 1, No = 0	conditioned space	x	x
56	C_DL_uncond_attic	Binary	Yes = 1, No = 0	conditioned space	x	x
57	C_DL_uncond_basement	Binary	Yes = 1, No = 0	conditioned space	x	x
58	C_state_OR	Binary	Yes = 1, No = 0	WI	x	x
59	C_state_MN	Binary	Yes = 1, No = 0	WI	x	x
60	C_state_NC	Binary	Yes = 1, No = 0	WI	x	x
61	C_state_TX	Binary	Yes = 1, No = 0	WI	x	x
62	C_number_bedrooms	Numeric	Univariate	N/A	x	x
63	C_num_floor_above_grade	Numeric	Univariate	N/A	x	x
64	C_floor_to_ceiling_height	Numeric	Univariate	N/A	x	x
65	C_conditioned_floor_area	Numeric	Univariate	N/A	x	x
66	C_envelope_leakage	Numeric	Univariate	N/A	x	x
67	C_roof_absorptance	Numeric	Univariate	N/A	x	x
68	C_skylight_area	Numeric	95th percentile = 1, zero = -1	N/A	x	x
69	C_total_window_area	Numeric	Univariate	N/A	x	x
70	C_wtmean_window_u_value	Numeric	Univariate	N/A	x	x
71	C_wtmean_window_shgc	Numeric	Univariate	N/A	x	x
72	C_HDD65	Numeric	Univariate	N/A	x	x
73	C_CDD65	Numeric	Univariate	N/A	x	x
74	C_age_years	Numeric	95th percentile = 1, zero = -1	N/A	x	x
75	C_Fnd_Flr_R_Value	Numeric	95th percentile = 1, zero = -1	N/A	x	x
76	C_Roof_R_Value	Numeric	Univariate	N/A	x	x
77	C_Wall_R_Value	Numeric	Univariate	N/A	x	x
78	C_NG_HeatingEff	Numeric	Univariate	N/A	x	
79	C_CoolingEff	Numeric	Univariate	N/A		x

Appendix D: Statistical Equations

Table 21 shows the mathematical equations used to populate Table 4 and Table 5 in Section 2.

Table 21. Mathematical Equations Used to Populate Table 4 and Table 5

Statistic	Description	Equation
Number of Observations	The number of observations in sample	n
Mean Measured	The mean value of the measured observations	$\frac{\sum_{i=1}^{n} m_i}{n}$
Mean Predicted	The mean value of the predicted observations	$\frac{\sum_{i=1}^{n} p_i}{n}$
Difference	Differences between predicted and measured observations (d)	$p - m$
Mean Difference	The mean value of the differences between predicted and measured observations $(\bar{d})$	$\frac{\sum_{i=1}^{n}(p_i - m_i)}{n}$
Median Difference	The median value of the differences between predicted and measured observations	The value for which 50% of errors are lower and 50% are higher.
Standard Deviation of Difference	The sample standard deviation of the differences	$\sqrt{\frac{1}{n-1}\sum_{i-1}^{n}(d_i - \bar{d})^2}$
Mean Absolute Difference	The mean value of the absolute differences	$\frac{\sum_{i=1}^{n}\lvert d_i \rvert}{n}$
Median Absolute Difference	The median value of the absolute differences	The value for which 50% of errors are lower and 50% are higher.
Mean Absolute Percent Difference	The mean value of the absolute differences	$\frac{\sum_{i=1}^{n}\left\lvert {}^{d_i}/_{m_i} \right\rvert}{n}$
Median Absolute Percent Difference	The median value of the absolute differences	The value for which 50% of errors are lower and 50% are higher.
RMSE	The square root of the mean value of the squared differences	$RMSE = \sqrt{\frac{\sum_{i=1}^{n} {d_i}^2}{n}}$
Percent RMSE	Normalized RSME value of the squared differences	$NRMSE = \frac{RMSE \times 100}{Mean\ Measured}$