
FSEC Energy Research Center®

9-1-2010

Evaluation Of Bias Issues Within Regression-Based Inverse Modeling Methods Against Climate And Building Characteristics Using Synthetic Data

Florida Solar Energy Center

Jamie Cummings
Florida Solar Energy Center



Part of the [Energy Systems Commons](#)

Find similar works at: <https://stars.library.ucf.edu/fsec>

University of Central Florida Libraries <http://library.ucf.edu>

This Contract Report is brought to you for free and open access by STARS. It has been accepted for inclusion in FSEC Energy Research Center® by an authorized administrator of STARS. For more information, please contact STARS@ucf.edu.

STARS Citation

Florida Solar Energy Center and Cummings, Jamie, "Evaluation Of Bias Issues Within Regression-Based Inverse Modeling Methods Against Climate And Building Characteristics Using Synthetic Data" (2010). *FSEC Energy Research Center®*. 304.
<https://stars.library.ucf.edu/fsec/304>



University of
Central
Florida

STARS
Showcase of Text, Archives, Research & Scholarship



FLORIDA SOLAR ENERGY CENTER®

Creating Energy Independence

Evaluation of Bias Issues within Regression-Based Inverse Modeling Methods Against Climate and Building Characteristics Using Synthetic Data

FSEC-CR-1863-10

July 2010

Authors

Jamie Cummings
Danny Parker
Karen Sutherland

Copyright ©2009 Florida Solar Energy Center/University of Central Florida
All Rights Reserved.

1679 Clearlake Road
Cocoa, Florida 32922, USA
(321) 638-1000

www.floridaenergycenter.org



A Research Institute of the University of Central Florida

Disclaimer

The Florida Solar Energy Center/University of Central Florida nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the Florida Solar Energy Center/University of Central Florida or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the Florida Solar Energy Center/University of Central Florida or any agency thereof.

Overview

Typically, a model of the energy use of a building is created using the building's characteristics and climate information for its location. However, models of existing homes can be created using monitored energy and climactic data in a process called inverse modeling. Such a technique has potential for those seeking to evaluate the savings of a home before and after a retrofit (Meier, Busch, & Conner, 1988).

Several programs use an inverse modeling technique to model a home's energy use based on total home energy use, heating and cooling energy use, and indoor/outdoor temperature data from a home (Kissock, Haberl, & Claridge, 2003). For example, PRISM, the Princeton Scorekeeping Method, was used widely in the 1980s and 90s to evaluate the performance of residential energy efficiency improvements (Fels, 1986).

In 2006, researchers from Building America evaluated several homes using a least-squares regression technique on monitored data from these homes. The researchers compared the performance of low-energy homes working towards the Building America goal of 70% whole-house efficiency to homes built to minimum code requirements. The 2006 study found a correlation between the efficiency of a home to the performance reported by the least-squares regression; however it asserted that more analysis into the method was needed in order to ascertain its accuracy, especially regarding floor type, climate, and house size (Chasar, et al., 2006).

This paper seeks to evaluate the accuracy of the regression technique used in 2006 by using the same least-squares regression analysis on data created from hourly energy simulation software. Synthetic data allows researchers to inexpensively and quickly obtain more data for analysis and to isolate the effects of single characteristics on a home's performance in a regression analysis model.

Least-Squares Regression

This model uses the relationship between cooling or heating energy and the average daily difference in temperature between outdoors and indoors. The difference between indoor and outdoor daily average temperatures is plotted on the x-axis of a graph, and the daily cooling or heating energy used per 1,000ft² of finished floor area is plotted on the y-axis. A line of least squares is fitted onto the data. Researchers calculate the area underneath the curve along a specified interval, making the assumption that this area corresponds to the amount of energy used in the home. If so, the area can be compared with a baseline home to determine the home's energy savings.

This method seeks to account for differences in energy consumption that result from different weather conditions and thermostat set points and from differing house size by using the outdoor-indoor delta T and by using the energy use per 1,000 square feet of floor area. Ideally, this method should evaluate a building's energy performance without being affected by differing building size, number of stories, weather, and location.

Example: North Port, FL

Consider a near-zero energy home in North Port, FL. The 1,446ft², one story home has a HERS index of 24. Its energy efficient features include photovoltaic panels, solar thermal water heating, and an 18.4 SEER/9.1 HSPF heat pump.

First, the researcher plots the home’s daily average outdoor-indoor delta T and cooling energy use per 1,000ft². A line of best fit on the data is imposed (Figure 1). This line is compared to the line from a code-minimum home in Lakeland, Florida (Figure 2). The area underneath the North Port line on the interval -5°F to 10°F, 115, represents the energy use of the home. Compared to the area underneath the Lakeland control trend line, 248, the North Port home shows 46% energy savings over minimum code (Figure 3).

In addition to calculating the area under a line of least squares, simply comparing the line of one house to another can visually reveal the differences in efficiency. The steeper a line of regression, the more energy a home will use. In the following graph, the standard-construction Lakeland Control Trendline is much steeper than the other homes that feature significant energy-efficiency improvements (Figure 4). The Lakeland Low-Energy Trendline and the NZ1-Gainesville trend line show that the two homes have similar energy performance.

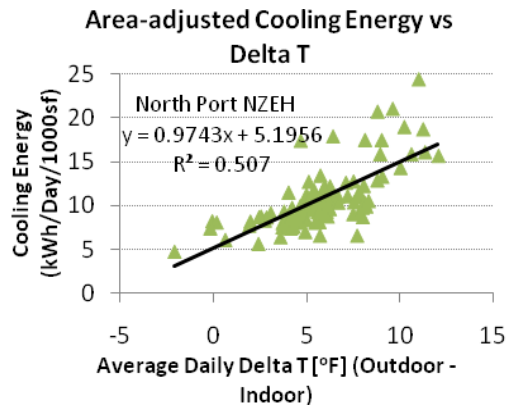


Figure 1. The regression analysis plot of a near-zero energy home in North Port, FL. It compares the area-adjusted cooling energy data to the difference between indoor and outdoor temperature data. The area under the line of best fit corresponds to the energy efficiency of the building.

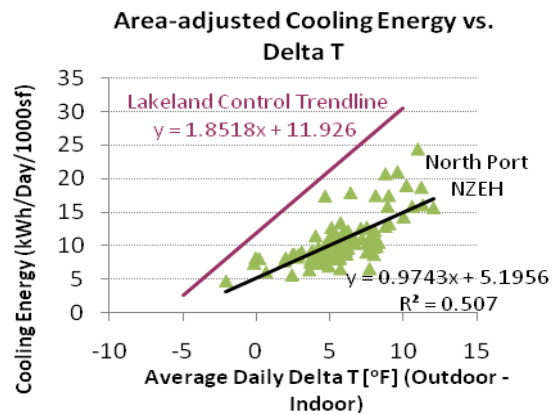


Figure 2. The line of best fit for the North Port home compared to a baseline data set of a standard efficiency Lakeland, FL home over five summers.

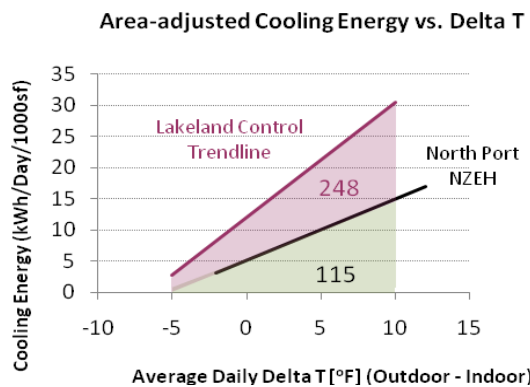


Figure 3. The area under the lines of best fit, from -5°F to 10°F, show that the North Port NZEH uses 46% less cooling energy than the Lakeland Control.

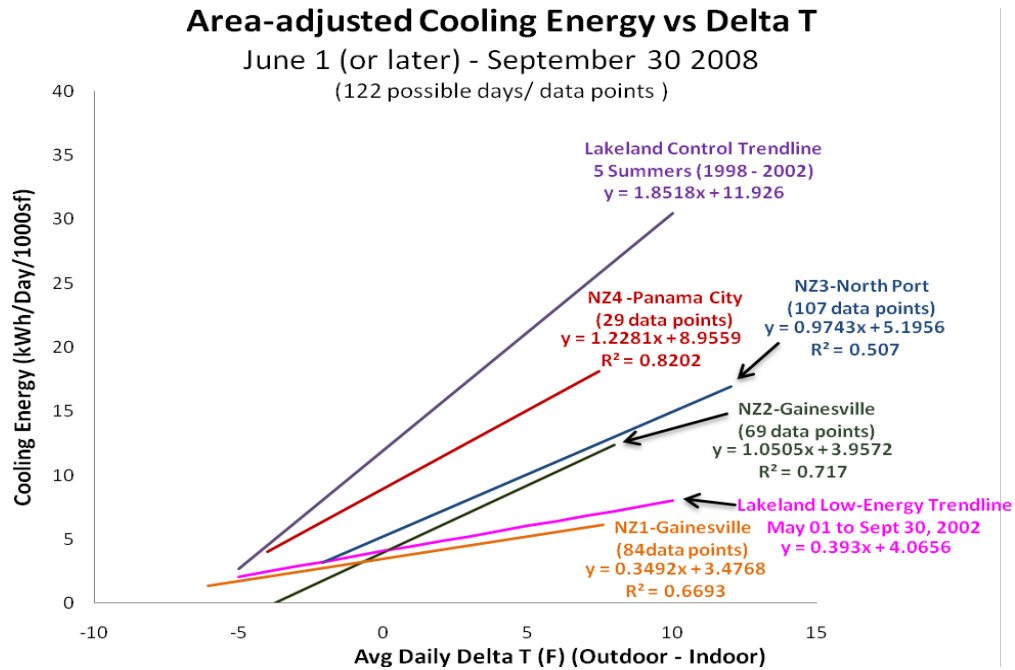


Figure 4. Lines of regression for four near zero energy homes measured in Florida and a high-efficiency and standard-efficiency home to serve as low- and high-efficiency comparison points. The more shallow the line's slope, the less energy the home uses.

The study found that this analysis accurately reflects increases in energy savings due to efficiency improvements. There is a close correlation between regression analysis savings and energy efficiency performance calculated by existing energy rating methods, such as the Building America Benchmark (Figure 5a-b), for several identical homes with differing levels of efficiency.

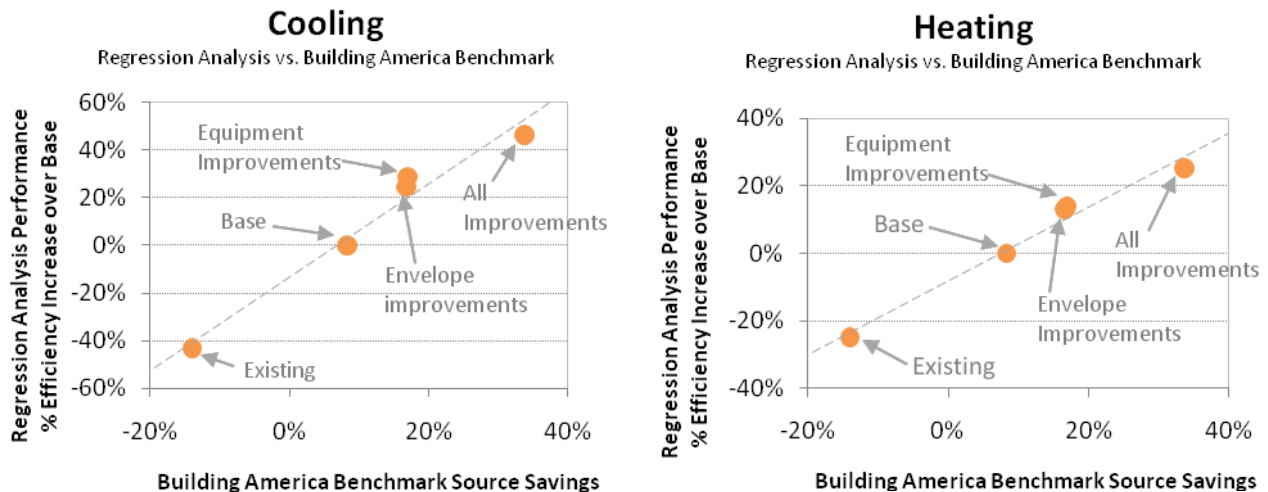


Figure 5a-b. Differing efficiencies show a promising correlation between the regression analysis method and Building America Benchmark rating method. The more efficiency measures added to a home, the more savings in both Building America Benchmark and regression analysis techniques.

Potential Problems

Although this method appears satisfactory for the several homes analyzed in the 2006 report, questions still remained regarding how this method would be affected by differing climates, house size, and occupant concentration.

The regression technique uses the change in indoor/outdoor temperature to normalize for weather conditions; however, climatic conditions other than air temperature, such as solar radiation and ground temperature, affect the cooling and heating energy of a building. These influences could cause considerable error in regression analysis comparisons across climates or days with different cloud cover.

The 2006 report anticipated discrepancies due to different floor types. Most of the homes studied had slab-on-grade foundation types, but researchers hypothesized that the homes with basements would show increased efficiency and those with crawlspaces would show decreased efficiency.

Biases due to house size and geometry are possible, even after adjustments for floor area. Researchers proposed that one home in the 2006 study would perform worse than the code-minimum baseline despite its energy-efficient design because it was significantly smaller (34%-62%) than the other research homes.

Synthetic Regression

To assess the validity of the regression analysis method for determining the energy savings of a home, researchers used the method on 30 simulated homes. They used energy simulation software to create synthetic temperature and cooling/heating energy data for all of the homes. They also simulated an 1824ft² Atlanta base home (Table 1) and changed several key parameters affecting efficiency: house size, climate, number of stories, and foundation type (Table 2).

Table 1. Base House Characteristics.

	Description	Value/Efficiency
Floor	Slab	1824ft ²
Roof	Medium Shingle, Vented Attic	R-30 Ceiling Insulation
Walls	Wood Frame	R-13
Windows	Double-Pane Metal Frame	U=0.447, SHGC=0.547
Ventilation	Mechanical Exhaust	54.4cfm (100% ASHRAE 62.2)
Infiltration		ACH ₅₀ =9.84
A/C	A/H in Attic	3-ton, SEER 13
Heating	Natural Gas Furnace	80% AFUE, 43kBtu/hr
Ducts		R-4.2, Q _n =0.10
Water Heating	40-gal Natural Gas	0.59 AFUE

Table 2. Changes made to the base home.

	Base Home Characteristics	Changes Made (Each line is a new simulation)
Efficiency	NREL Base (8% BA Savings)	Existing Home (-13% BA Savings) Equipment Improvements (17% BA Savings) Envelope Improvements (17% BA Savings) Envelope and Equipment Improvements (34% BA Savings)
Size	1,824ft ²	2,736ft ² 3,648ft ² 5,472ft ²
Type	Two Story	One Story ¹
Foundation	Slab	R-19 Insulated Crawlspace Basement (R-19 Ceiling)
Climate	Atlanta, GA (TMY2)	Miami, FL Houston, TX Olympia, WA Omaha, NE Helena, MT

Parametric Comparison

For the regression technique to be a valid, the area under the regression curve should be the same for identical homes of the same efficiency level when evaluating home energy performance. In other words, the method should produce the same results for all parameters changed in addition to efficiency.

Efficiency

Efficiency is the only parameter where changes in performance should occur. Researchers simulated four different changes to efficiency in the base house (Table 3). The first simulated home includes features found in many existing homes today. The other three homes feature efficiency improvements: equipment, envelope, and combined equipment, envelope, and appliance upgrades.

Table 3

Existing Home Characteristics	Base Home Characteristics	Equipment Improvements	Envelope Improvements	All Improvements
R-11 Walls Single Metal Windows SEER 10 A/C 75% AFUE Furnace 54% AFUE Water Heater Refrigerator: 1000kWh/yr	R-13 Walls Double Metal Windows SEER 13 A/C 80% AFUE Furnace 59% AFUE Water Heater	SEER 18 A/C 94% AFUE Furnace	R-60 ceiling R-19 walls Low-e double windows, U=0.218, SHGC=0.28	Envelope & Equipment Improvements Energy Star Appliances 100% Fluorescent Lighting
BA Benchmark Savings: -14%	BA Benchmark Savings: 8%	BA Benchmark Savings: 17%	BA Benchmark Savings: 17%	BA Benchmark Savings: 34%

The simulations showed a close correlation between the level of efficiency of a home and regression analysis performance. Figure 6 shows the cooling and heating results.

¹ Each size house was simulated with one story and two stories. To change the home to two stories, researchers doubled the first floor area while keeping the length-width ratio of the house and the window area the same.

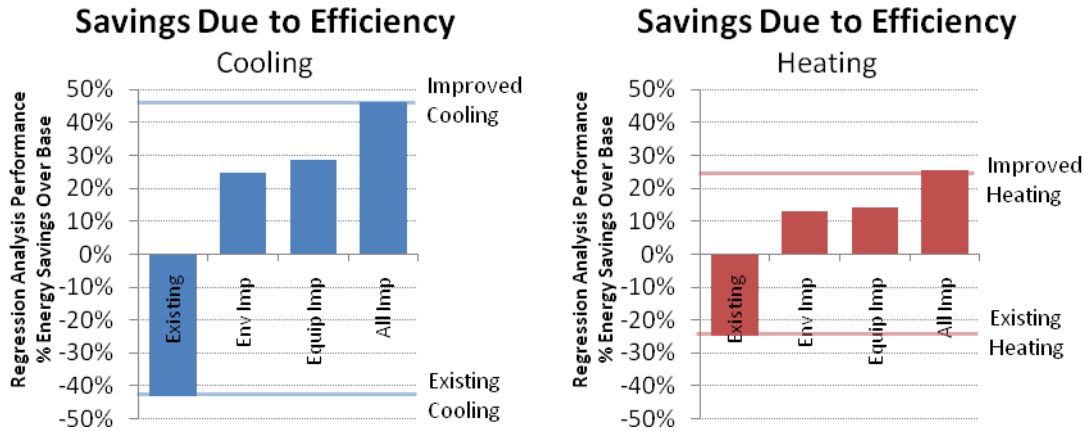


Figure 6. Heating and cooling regression analysis savings in homes with differing efficiencies.

The existing home and the home with both equipment and envelope improvements (“All Improvements”) are used in the rest of this document as reference points to show the magnitude of discrepancies in savings reported for the other parameters (Table 4 and Table 5).

House Size

The size of a home should not affect the relative performance of a home with similar envelope and equipment features. The base house size was increased, keeping the same proportions but increasing the finished floor area and window area by 150%, 200%, and 300%. The changed specifications are listed below in Table 6.

Table 4. Existing Home

Existing Home	Cooling	Heating
Linear Fit Equation	$y=1.3x+14.4$	$y=-0.10x-0.72$
R²	0.84	0.89
Area under curve	263	52
% efficiency increase over base	-43%	-25%

Table 5. Home with Equipment and Envelope Improvements

All Improvements	Cooling	Heating
Linear Fit Equation	$y=0.56x+5.2$	$y=-0.064x-0.51$
R²	0.80	0.90
Area under curve	99	31
% efficiency increase over base	46%	25%

Table 6. Size Parameter Specifications

Simulated Home	Finished Floor Area [ft ²]	Ceiling/Floor Area [ft ²]	Perimeter [ft]	Window Area [ft ²]
Base	1,824	912	122	328
1.5x Sized	2,736	1,368	152	492
2x Sized	3,648	1,824	176	656
3x Sized	5,472	2,736	215	984

The size of a home affected the regression significantly. Although no pattern emerges, different sized bases for the same home made a large impact on the slope of the regression line (Figure 7).

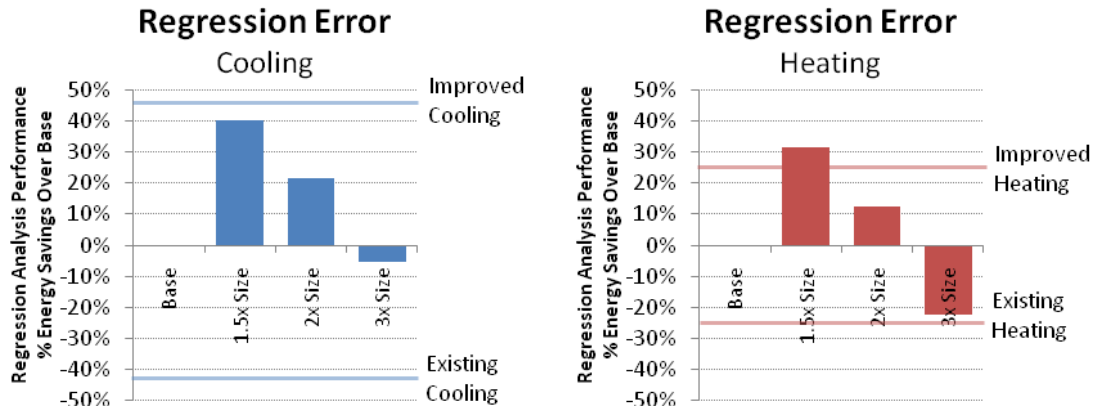


Figure 7. Cooling and heating regression analysis savings for different sized homes.

For example, the 2,736ft² home (150% of the base size), showed a cooling and heating efficiency increase of 32% and 40%, respectively (see Table 7), when it should show no savings. These numbers make the home seem significantly more efficient than it actually is—its savings are comparable to the savings from the “All Improvements” home with significant energy-efficiency improvements.

Table 7. Results for the home 150% larger than the base.

1.5x Sized	Cooling	Heating
Linear Fit Equation	$y=0.54x+6.0$	$y=-0.052x-0.33$
R²	0.80	0.87
Area under curve	110	29
% calculated savings over base	40%	32%

Number of Stories

The number of stories in a building has the potential to skew regression results. Researchers simulated a one-story version of the base, 1.5x-, 2x-, and 3x-sized homes with the same finished floor area and similar features. The effect of one-story versus two-story buildings proved to be insignificant, with the lines of regression only slightly different.

Foundation Type

The regression analysis method normalizes for air temperature but does not consider ground temperature in its calculations. Because foundation type changes the effect of ground temperature on building performance, researchers anticipated more discrepancies with slab and basement foundations. Ground temperature should affect the cooling and heating loads more than a crawlspace foundation. For this reason, researchers simulated homes with a basement and a crawlspace foundation type; however, slab, basement, and crawlspace performance were very similar in the Atlanta home.

Climate

The effect of climate on building performance using regression analysis is of special concern. This method does not normalize for climate factors other than air temperature, such as ground temperature, solar radiation, wind, and humidity. To understand the effects of climate on performance reported by regression analysis, researchers simulated the same base home in five other climate zones (Table 8).

Climate proved to have a significant effect on the calculations, overwhelming the effects of efficiency improvements (Figure 8). The most extreme examples, Montana for cooling and Florida for heating, have such small cooling and heating seasons, respectively, that in an absolute sense they are insignificant. Still, heating climates showed large errors—positive error in the warmer climates and negative error in the cooler climate.

Table 8

City	Climate Region
Miami, FL	Hot-Humid
Atlanta, GA (Base)	Mixed-Humid
Houston, TX	Hot-Dry
Olympia, WA	Marine
Helena, MT	Cold
Omaha, NE	Cold

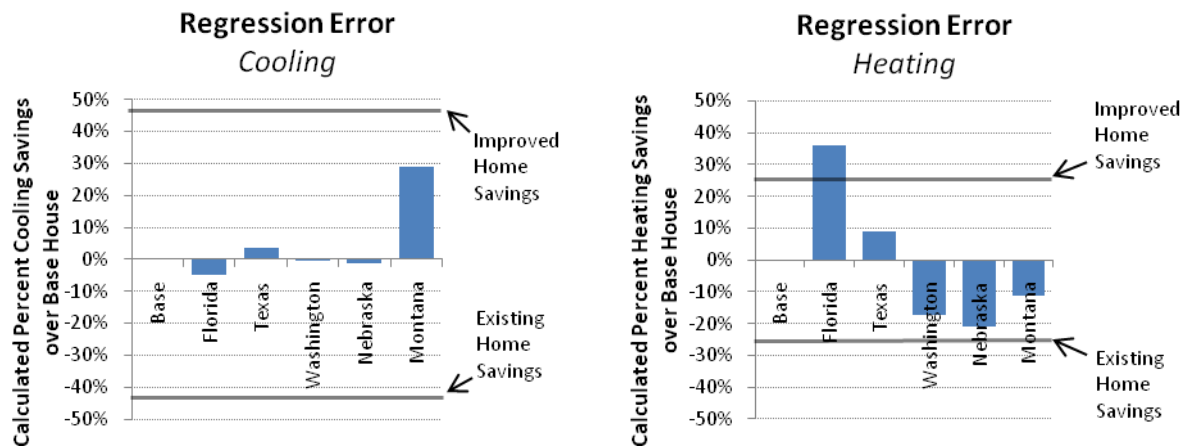


Figure 8. Different climates showed vastly different heating performance. Cooling performance was similar, with the exception of Montana, whose cooling season is small enough that the percentage change is insignificant.

The heating error is most likely caused by differences in ground temperature and the heat gain/loss through the home’s foundation (Figure 9).

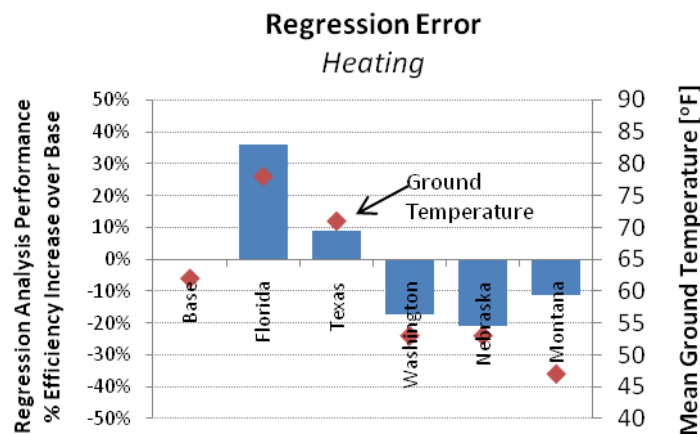


Figure 9. Heating performance error is similar to the mean ground temperature for that climate, suggesting that the error is based primarily on the effects of ground coupling. (Ground temperature data from “Geothermal Heat Pump Manual” by McQuay International, 2002)

Further investigating this hypothesis, researchers simulated the houses with crawlspace and basement foundations in the five other climates (Figure 10). The slab foundation type showed higher error for each climate: significantly lower heating performance in the three colder-climate areas: Nebraska, Montana, and Washington, and significantly higher heating performance in Atlanta and Miami (Figure 10). The basement and crawlspace homes both performed somewhat better in the warmer climates and somewhat worse in cooler climates.

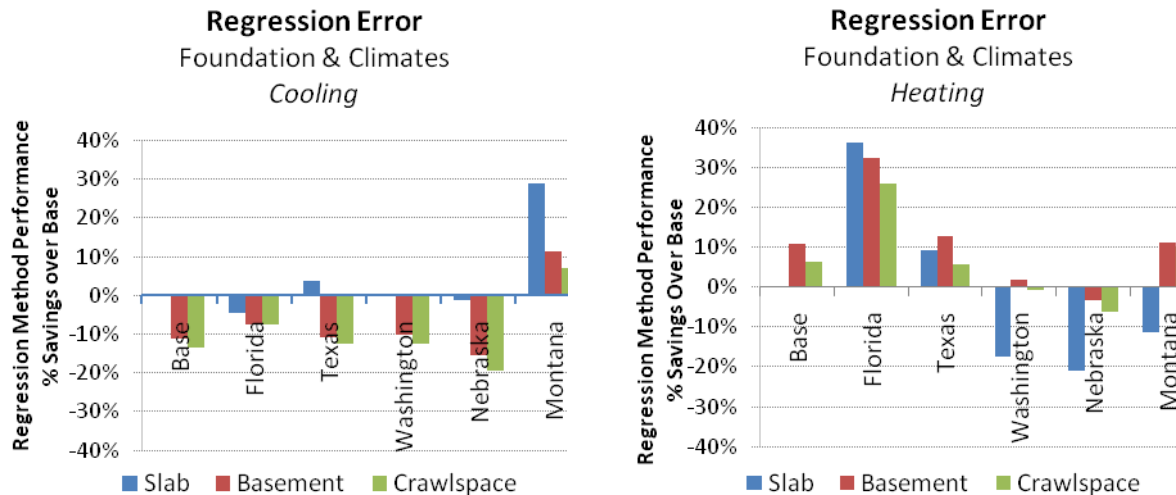


Figure 10. A slab foundation type shows greater heating savings variation across climates than crawlspace and basement foundation types.

Conclusions

Regression analysis appears to effectively compare home energy performance across similar-sized homes in the same climate. However, comparisons between homes in different climates or with different finished floor area amounts introduce considerable error in heating savings (Figure 11). Slab foundations, in contrast with crawlspace and basement foundation types, increased the amount of heating savings error. This result suggests that the heating error is largely due to interactions between the house and the ground, although other factors may be involved, most notably solar radiation.

The other large factor affecting the regression analysis calculation method is house size. The size of a home had a large impact on cooling and heating, with no pattern emerging (Figure 12).

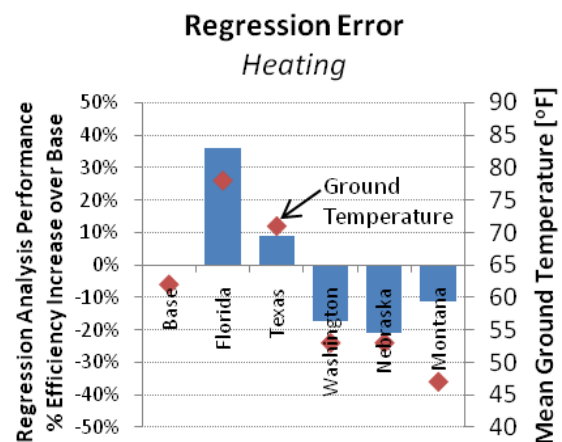


Figure 11. Heating performance error is similar to the mean ground temperature for that climate, suggesting that the error is based primarily on the effects of ground coupling. Ground temperature data from “Geothermal Heat Pump Manual” by McQuay International, 2002.

For researchers seeking to identify the energy efficiency of a home, the regression analysis method does not succeed. It is heavily affected by house size and climate. For example, the 1824ft² base home modeled with equipment efficiency improvements (SEER 18/94% AFUE) showed 29% heating savings. However, the 2736ft² home (150% of the base home size) with standard equipment (SEER 13/80% AFUE) showed 40% heating savings. Hence, this standard efficiency, 2736ft² home's performance is artificially increased to appear more efficient than a smaller home with high-efficiency equipment.

In general, this method is not effective for comparing house performance. However, for those comparing similarly-sized homes in similar climates, this method allows researchers to draw conclusions about the efficiency of a home based purely on its monitored data.

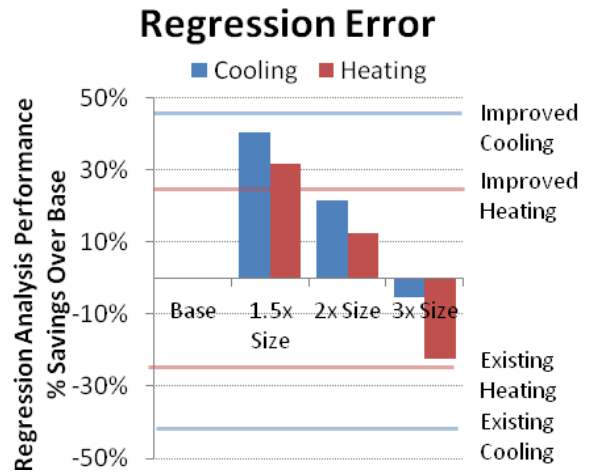
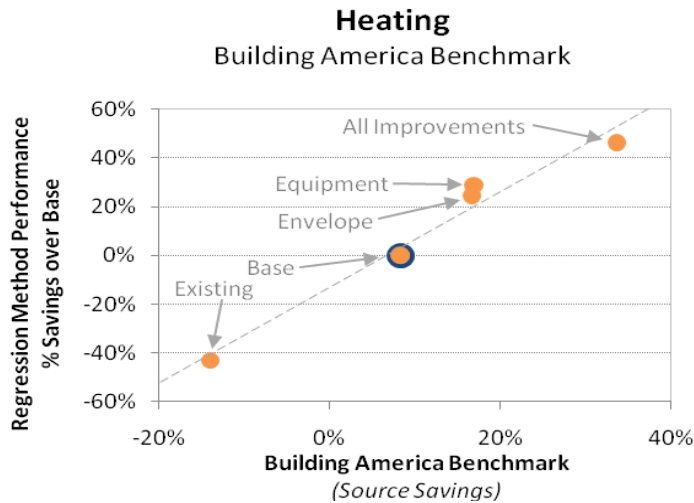


Figure 12. The variation in performance due to changing the size of the home clearly dwarfs the effect that significant energy efficiency measures have on building performance.

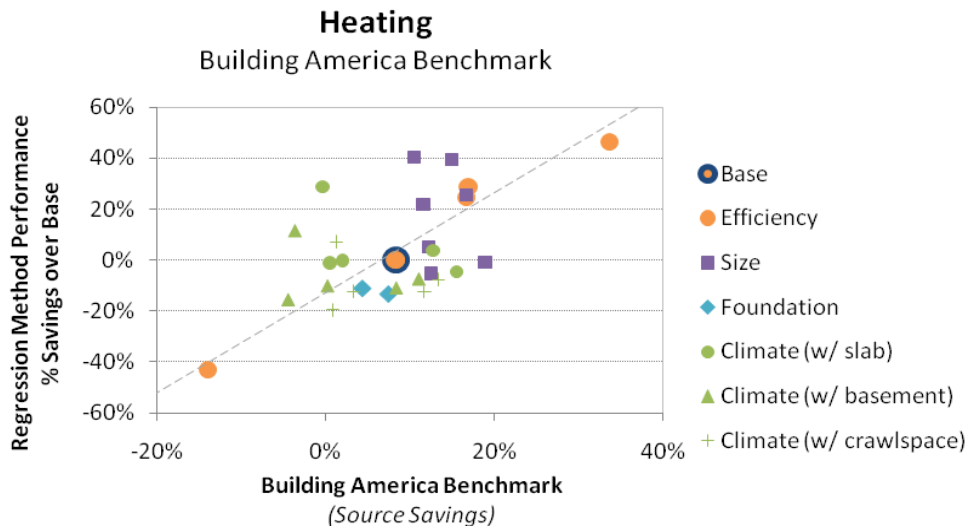
Appendix A - Base House Characteristics

Base House Characteristics		
City		Atlanta, GA (TMY2)
Stories		2
Bedrooms		3
Bathrooms		2
Floor Area		1824ft ²
Av. Wall Height		8ft
Volume		14,592ft ³
Floor Type		Slab
% tile-wood-carpet		0-20-80
Roof Material		Composition shingles
Roof Solar Absorptance		0.92
Attic Ventilation		Vented
Radiant Barrier		None
Ceiling Insulation		R-30
Wall Type		Frame
Wall Insulation		R-13
Window Type		Double (Clear)
Window Frame		Metal
U-value		0.447
SHGC		0.547
% Win/Floor area		20.0%
Mechanical Vent.		54.4cfm exhaust (100%)
ACH50		9.84
A/C:	SEER	13
Capacity		39.1 kBtu/hr
Heating:	Type	Natural Gas Furnace
Capacity		43.3 kBtu/hr
Efficiency		0.8
Ducts:	Insulation	R-4.2
Leakage		10%
Location		Attic
Air Handler Location		Attic
Water	Type	Natural Gas
Heater:		
Location		Attic
Efficiency		0.59
Capacity		40 gal
Programmable		No
Thermostat		
Ceiling Fans		No
% Fluorescent Lighting		14%
Refrigerator		671 kWh/yr

Appendix B



This graph shows the savings results created by the least-squares regression analysis compared to the energy savings calculated by the Building America Benchmark. Ideally, regression analysis should produce homes for which changes in efficiency are the only factors changing the energy savings. In the graph above, the orange dots show a strong correlation between increasing energy efficiency and better regression analysis performance.



Other parameters, however, showed great effects on the calculated performance of the homes. Differing climates and sizes dwarfed the effects of the efficiency changes, making this method invalid for homes of different size or homes in different climate zones. However, this method still works for comparisons of before and after home retrofits that do not alter the size of the house.

References

- Chasar, D., Chandra, S., Parker, D., Sherwin, J., Beal, D., Hoak, D., et al. (2006, July 24-26). Cooling Performance Assessment of Building America Homes. *Fifteenth Symposium on Improving Building Systems in Hot-Humid Climates* .
- Fels, M. (1986). PRISM: An Introduction. *Energy and Buildings* , 5-18.
- Kissock, J. K., Haberl, J., & Claridge, D. (2003). Inverse Modeling Toolkit: Numerical Algorithms. *ASHRAE Transactions - Research* , 109, Part 2, 425-434.
- McQuay International. (2002). *Geothermal Heat Pump Design Manual*.
- Meier, A. K., Busch, J., & Conner, C. C. (1988). Testing the Accuracy of a Measurement-based Building Energy Model with Synthetic Data. *Energy and Buildings* , 12, 77-82.