Identifying Potential Air Sealing Candidates Without Blower Door Testing

Jake Millette, Michaels Energy

ABSTRACT

The EPA estimates that US homes can save an average of 15% of heating and cooling energy use by improving air sealing and insulation. Individual houses can have higher levels of air infiltration than others, but there is no way of knowing how leaky a house is and whether it is prudent to do comprehensive air sealing without going onsite and conducting a blower door test, which costs time and money.

To identify homes with the greatest savings potential from air sealing, we analyzed blower door test data and common housing characteristics from over 1,000 sites in six states. Using this data, we ran multiple regressions to identify the building characteristics that best correlate with a home's air infiltration rate (CFM50). Homes with relatively high air infiltration rates generally have the highest potential for savings through improved air sealing and insulation.

This paper is a proof of concept of how energy efficiency program administrators can use available building characteristics to not only better target homes for air sealing, but to more accurately estimate the potential savings of the homes and broader residential building stock.

Introduction

Reducing air leakage in homes can result in many occupant benefits, including lower energy costs, improved comfort, and improved indoor air quality. Energy savings occur through the more effective operation of HVAC equipment as well as the ability to install smaller and more efficient systems to meet the needs of lower heating and cooling loads. The EPA estimates that US homes can save an average of 15% of heating and cooling energy use by improving air sealing and insulation. (US Department of Energy 2026b)

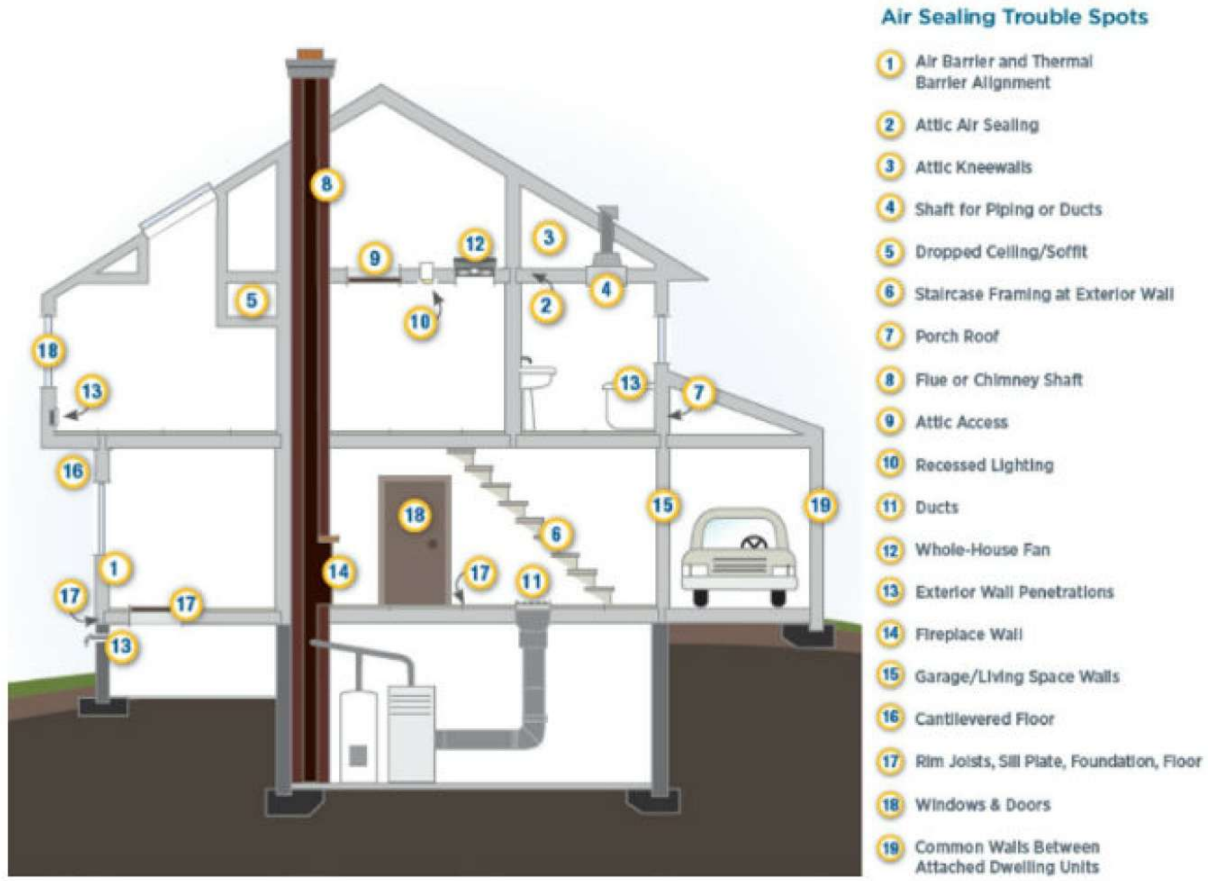
Home energy professionals measure air infiltration rates using a blower door test. This test involves depressurizing a home and then measuring the rate of air coming back in through gaps in the building envelope (air infiltration). Blower door tests often measure the building's airtightness in terms of CFM50 or the airflow in Cubic Feet per Minute that is required to maintain the home's internal pressure of 50 Pascals. The higher the airflow CFM, the more leaky the house. According to The Energy Conservatory (2026): "An old and uninsulated, two-story, wood framed house would likely produce a CFM50 test result between 4,000 and 8,000 [CFM], which is very leaky. A new modern house built to a strict airtightness standard would likely produce a test result between 600 and 1,000 [CFM], which is very tight." Blower door testing is recommended for air sealing to not only identify the extent of leaks but to set a baseline from which an air infiltration reduction target can be established.

To help program administrators understand the potential for energy savings from air infiltration, Michaels Energy has completed blower door tests as part of several recent baseline and potential studies. While doing so, we hypothesized that these administrators may be able to use commonly available home characteristics to predict a home's level of air infiltration (i.e., its potential for savings from air sealing measures). This paper explores that hypothesis, using information from site visits from residential baseline studies to understand the common characteristics of homes with high air infiltration. The data for this paper comes from over 1,000 blower door tests conducted by Michaels Energy as part of residential baseline studies conducted from 2021 to 2025 in six states in the Pacific Northwest, Midwest, and Mountain Region. Program administrators can then use these building characteristics to not only better target homes for air sealing, but to more accurately estimate the potential savings of the homes and broader residential building stock.

Methodology

As a first step, Michaels Energy identified the building characteristics that most contribute to a home's air infiltration (shown in the figure below) and compared that to the data available from the baseline studies included in this analysis.

Figure 1. Common Residential Air Infiltration Areas



Source: U.S. Department of Energy 2026a

Michaels Energy identified building characteristics related to home air infiltration available in the data collected in residential baseline studies included in this analysis. The table below lists those variables as well as their justification for inclusion.

Table 1. Variables Related to Home Air Infiltration

Variable	Justification
Building Age	Newer homes were subject to modern building codes and construction techniques and materials. Newer homes also tend to be larger. Older homes may have increased air infiltration due to settling or damage or damage to the building envelope.
Conditioned Area (SF)	Larger homes will have more surface area for potential air infiltration.
Number of Windows	Windows are a common cause of air leakage through moving parts and poor sealing around window frames.

Variable	Justification
Presence of Ductwork	Ducts may result in air leakage if they penetrate the building envelope into unconditioned spaces.
Presence of Finished/Conditioned Basement	Finished and conditioned basements may have greater air sealing than unfinished basements or crawlspaces.
Presence of Attic	Attics may cause air leakage through access doors or poorly sealed plumbing or ducts.
Number of Bathrooms	Air leaks around plumbing penetrations and bathroom ventilation fans are common.
Presence of Wall Cavity Insulation	Presence of wall cavity insulation may indicate better air sealing.
Presence of Whole House Fan	Whole house fans may cause large amounts of air infiltration in unconditioned spaces.
Annual Energy Usage	Total annual weather-normalized electric and gas energy usage (kBtu) will be higher in homes with higher air infiltration.

After compiling a data set of the blower door test data and prioritized building characteristics from the baseline studies included in this analysis, the study team ran a series of multiple linear regressions to understand the relative impact of each independent variable on a home's air infiltration, measured in CFM50.[1] In the analysis, the study team reviewed the coefficient of determination (R^2) for each regression. This value represents the proportion of the variance in the dependent variables that is predictable from the independent variables. In other words, the R^2 value shows how much of the dependent variable (the level of air infiltration of a home) is explained by the independent variables (from the table above) in the regression instead of other factors. The higher the R^2 value, the smaller differences between the observed data and the predicted values.

Results

The table below shows the inputs and R^2 values for each of the regressions run for this analysis. Note that the variables shown in red indicate that they are not statistically significant ($p>0.05$).

[1] Air changes per hour at 50 Pascals (ACH50) is often considered a more useful measurement of air infiltration than CFM50 because it normalizes for the volume of the home and better incorporates the surface area through which air infiltration could occur. However, the baseline studies from which data was used for this analysis varied in the level of information collected to determine homes' volume. Because of concerns about data completeness and accuracy, the analysis used CFM50 as the measurement of air infiltration. In preliminary analyses, the results using CFM50 and ACH50 as the dependent variable were largely comparable.

Table 2. Inputs and Results of Regressions

Run	Input 1	Input 2	Input 3	Input 4	Input 5	Input 6	Input 7	Input 8	R^2	Number of Observations
1	Conditioned Area	Home Vintage	# of Windows	Ductwork	Conditioned Basement	# of Bathrooms	Wall Cavity Insulated	Annual Energy Usage	0.526	190
2	Conditioned Area	Home Vintage	# of Windows	Ductwork	Conditioned Basement	Annual Energy Usage	Presence of Attic	--	0.519	204
3	Conditioned Area	Home Vintage	# of Windows	Ductwork	Conditioned Basement	Annual Energy Usage	--	--	0.514	204
4	Conditioned Area	Home Vintage	# of Windows	Conditioned Basement	Annual Energy Usage	--	--	--	0.502	204
5	Conditioned Area	Home Vintage	Conditioned Basement	Annual Energy Usage	--	--	--	--	0.496	204
6	Conditioned Area	Home Vintage	Annual Energy Usage	--	--	--	--	--	0.481	204
7	Home Vintage	Annual Energy Usage	--	--	--	--	--	--	0.377	204
8	Conditioned Area	Home Vintage	# of Windows	Ductwork	Conditioned Basement	# of Bathrooms	Wall Cavity Insulated	--	0.365	809
9	Conditioned Area	Home Vintage	# of Windows	Ductwork	Conditioned Basement	# of Bathrooms	--	--	0.365	872
10	Conditioned Area	Home Vintage	# of Windows	Ductwork	Conditioned Basement	Presence of Attic	--	--	0.365	865

Run	Input 1	Input 2	Input 3	Input 4	Input 5	Input 6	Input 7	Input 8	R^2	Number of Observations
11	Conditioned Area	Home Vintage	# of Windows	Ductwork	Conditioned Basement	Whole House Fan	--	--	0.364	874
12	Conditioned Area	Home Vintage	# of Windows	Ductwork	Conditioned Basement	--	--	--	0.355	1005
13	Conditioned Area	Home Vintage	# of Windows	Conditioned Basement	--	--	--	--	0.351	1005
14	Conditioned Area	Home Vintage	# of Windows	--	--	--	--	--	0.341	1005
15	Conditioned Area	Home Vintage	--	--	--	--	--	--	0.313	1005
16	Annual Energy Usage	--	--	--	--	--	--	--	0.165	204
17	Home Vintage	--	--	--	--	--	--	--	0.147	1005
18	Conditioned Area	--	--	--	--	--	--	--	0.095	1005

Note: Red text indicates p>0.05

As shown in the table above, the regression with the highest R^2 (0.514) and with all variables statistically significant ($p<0.05$) was for "Run 3." This regression included:

- Conditioned Area of the Home – CFM50 increases with home size
- Home Vintage – CFM50 increases with increased age
- Number of Windows – Higher for homes with larger number of windows
- Presence of Ductwork – Higher for homes with ducts
- Presence of Conditioned/Finished Basement – lower for homes with finished/conditioned basements
- Annual Energy Usage – higher energy use results in higher CFM50

The table below shows the results from this regression.

Table 3. Results of "Run 3" Regression

	Coefficients	Standard Error	t Stat	P-value
Intercept	39,321.07	3,885.9010	10.1189	1.2E-19
Conditioned_Area	0.4245	0.1037	4.0929	6.21E-05
Home_Vintage	-19.6524	1.9756	-9.9475	3.78E-19
# of Windows	13.7680	9.9185	1.3881	0.1667
Ductwork	288.4209	131.0340	2.2011	0.0289
Conditioned Basement	-339.3330	147.4017	-2.3021	0.0224
Annual Energy Usage	0.0045	0.00140	3.2430	0.0014

There were two models with slightly higher R^2 values, but these included variables that were not statistically significant.

Another model, "Run 6," resulted in a similar R^2 value (0.481) had fewer variables, all of which could be easily compiled by a program administrator through public information (e.g., real estate listings) and utility billing data. These include:

- Conditioned Area of the Home
- Home Vintage
- Annual Energy Usage

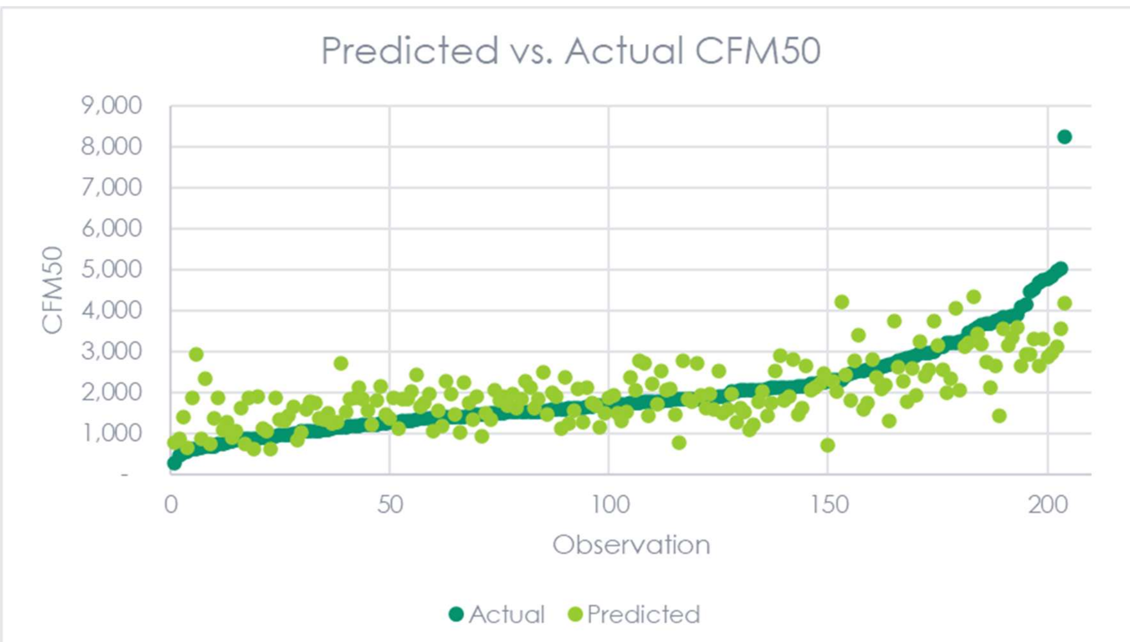
The table below shows the results from this regression.

Table 4. Results of “Run 6” Regression

	Coefficients	Standard Error	t Stat	P-value
Intercept	37,432.31	3,564.954	10.5001	8E-21
Conditioned_Area	0.4959	0.0785	6.3192	1.68E-09
Home_Vintage	-18.6315	1.8135	-10.274	3.73E-20
Annual Energy Usage	0.0056	0.0014	4.0453	7.46E-05

The figure below shows how the predicted CFM50 value from the regression compares to the actual value measured in the field. Based on this figure, it appears that there may be some bias in the model as it seems to overpredict the air infiltration of sites with low CFM50 values and underpredict sites with high CFM50 values.

Figure 2. Predicted vs. Actual CFM50 Value



Results by State

To understand how results may differ by state, the study team ran state-specific regressions of the regression incorporating all six states that had the highest R^2 (0.355). As shown in the table, the R-squared value for the states ranged from 0.215 to 0.433. Notably, for each state, at least one of the dependent variables was not statistically significant ($p>0.05$).

Table 5. Inputs and Results of Regressions by State

State	Input 1	Input 2	Input 3	Input 4	Input 5	R^2	Number of Obs.
1	Conditioned Area	Home Vintage	# of Windows	Ductwork	Conditioned Basement	0.422	65
2	Conditioned Area	Home Vintage	# of Windows	Ductwork	Conditioned Basement	0.427	243
3	Conditioned Area	Home Vintage	# of Windows	Ductwork	Conditioned Basement	0.277	66
4	Conditioned Area	Home Vintage	# of Windows	Ductwork	Conditioned Basement	0.215	117
5	Conditioned Area	Home Vintage	# of Windows	Ductwork	Conditioned Basement	0.423	254
6	Conditioned Area	Home Vintage	# of Windows	Ductwork	Conditioned Basement	0.433	260
Overall	Conditioned Area	Home Vintage	# of Windows	Ductwork	Conditioned Basement	0.355	1,005

Note: Red text indicates p>0.05

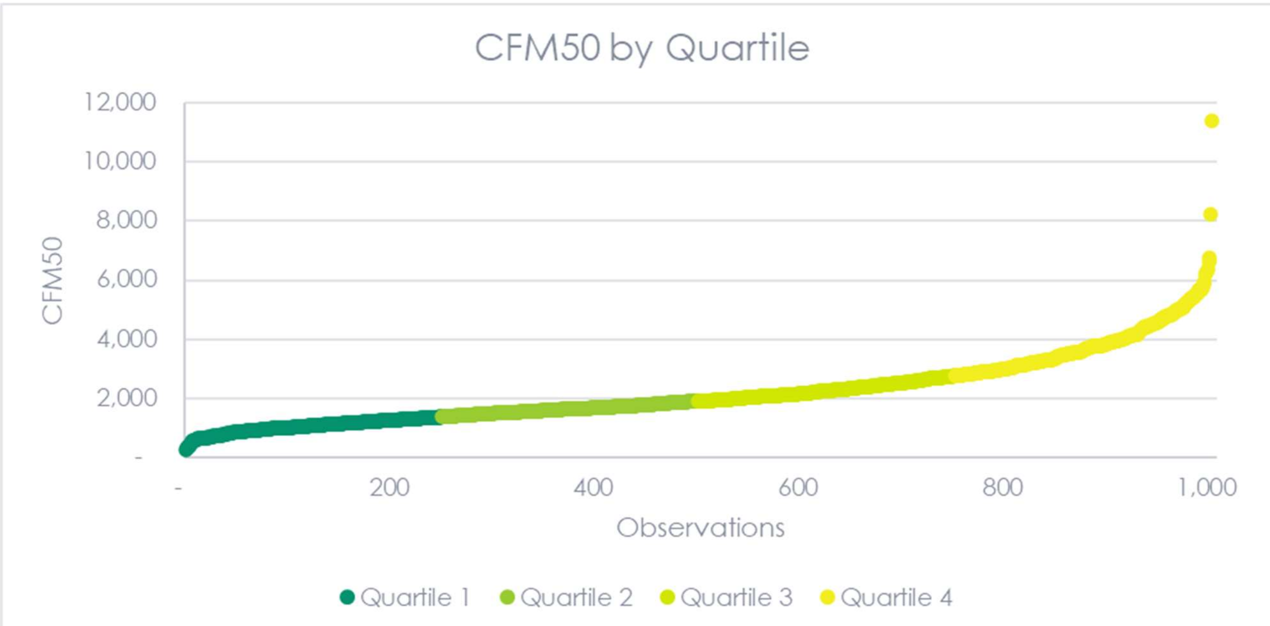
The wide range of results by individual states suggests that other factors, such as climate zones, building code adoption and enforcement, and history of energy effciency programs, may also factor into the predictive ability of these variables. The relatively small sample sizes for each state may also affect the results.

Quartiles Analysis

Our analysis shows that the best regressions made with commonly available data only explain about half of a home's air leakage, expressed as CFM50. But for program administrators, that may be enough. Rather than needing to predict an exact level of air leakage, programs only need to identify homes that are likely to have high levels of air infiltration that will make air sealing helpful.

Program administrators will likely set their own thresholds of air leakage values to target based on expected savings and cost effectiveness, but for this exercise the study team reviewed the 1,005 sites in our data set and established quartiles for measured CFM50. As shown in the figure below, the cutoff for the first quartile is 1,369 CFM50, the cutoff for the second quartile (i.e., the median) is 1,899, and the cutoff for the third quartile is 2,771.

Figure 3. CFM50 Values by Quartile



After establishing these quartiles, the study team reran the preferred regressions using the quartile of air leakage as the independent variable instead of CFM50. As shown in the table below, the R^2values for these regressions were similar but slightly lower than the same regressions using CFM50.

Table 6. Inputs and Results of Regressions using CFM50 Quartiles

Run	Input 1	Input 2	Input 3	Input 4	Input 5	Input 6	R^2	Number of Obs.
1	Conditioned Area	Home Vintage	# of Windows	Ductwork	Conditioned Basement	Annual Energy Usage	0.452	204
2	Conditioned Area	Home Vintage	Annual Energy Usage	--	--	--	0.406	204

When analyzing how the predicted quartiles compared to the actual quartiles, we found mixed success. Less than half of predicted quartiles (43%) matched the actual quartile, while similar shares of observations were either greater or less than the actual values.

Table 7. Comparison of Predicted and Actual Results by Quartile

	Observations	Share
Actual < Predicted	57	28%
Actual = Predicted	88	43%
Actual > Predicted	59	29%
	204	100%

If programs are only interested in identifying the worst performing homes (i.e., those above the median), then the regression performed better. Of the 64 homes that the model predicted were in the two quartiles with the highest CFM50 values, 77% were actually in that half. Similarly, the model correctly predicted that a home was in the worst half 75% of the time.

Table 8. Results by CFM50 Quartiles

	Predicted to be in lowest performing half	Predicted to NOT be in the lowest performing half	Total
In the lowest performing half	49	31	80
Not in the lowest performing half	15	109	124
Total	64	140	204

The table below shows the mean values of the three dependent variables used in this regression by quartile.

Table 9. Mean Home Characteristics by CFM50 Quartile

	1st Quartile	2nd Quartile	3rd Quartile	4th Quartile
Mean Conditioned Area (SF)	1,575	1,1786	1,930	2,461
Mean Home Vintage	1991	1982	1975	1955
Mean Annual Energy Use (electric and gas, kBtu)	68,328	74,197	89,749	122,365

Conclusions

Regressions using commonly available information about homes is a reasonably good predictor of homes' leakiness. However, approximately half of the CFM50 value was explained by the variables included in regressions and result from other variables. These likely involve factors related to the specific construction of each home. Every home is different and their nuances cannot be included in the simplifying assumptions used in this analysis.

Results by state vary considerably, suggesting that other factors, such as climate zones, building code adoption and enforcement, and history of energy effciency programs, may also drive results. Individual, state-specific regressions may need to be run for each program to identify the best independent variables to include and their coefficients.

Ultimately, if program administrators use a regression approach to identify homes that are the best candidates for air sealing, then they do not need to predict an exact CFM50 value, only that the home is likely to have relatively high air infiltration compared to other homes. In this proof of concept case, using only the conditioned area of the home (SF), the year of construction, and the annual energy use, the regression currently predicted the worst performing homes in more than three-quarters of cases.

References

O'Connor, B. and S. Narsiah. 2022. "A Deep Dive into Air Sealing Best Practices and Market Potential." In *Proceedings of the 2022 ACEEE Summer Study*. https://www.aceee.org/sites/default/files/summerstudy/2022/event-data/pdf/catalyst_activity_32446/catalyst_activity_paper_20220810190525138_d0d3b97f_64b9_46ed_b8f3_ebcdd14bf0f3.

The Energy Conservatory. "Blower Door Test Results: How Energy Auditors Measure a Home." https://www.energyconservatory.com/blog/blower-door-test-results-how-energy-auditors-measure-a-home/. Accessed 5/15/26.

US Department of Energy. "Air Sealing Your Home." https://www.energy.gov/energysaver/air-sealing-your-home. Accessed 5/14/26.

US Department of Energy. "Methodology for Estimated Energy Savings." https://www.energystar.gov/saveathome/seal_insulate/methodology. Accessed 5/15/26.