U.S. DEPARTMENT OF ENERGY | Energy Efficiency & Renewable Energy

BUILDING TECHNOLOGIES OFFICE

Predicting Envelope Leakage in Attached Dwellings

O. Faakye, L. Arena, D. Griffiths
Consortium for Advanced Residential Buildings

July 2013

Building America
U.S. Department of Energy

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

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/ordering.htm

Printed on paper containing at least 50% wastepaper, including 20% postconsumer waste

U.S. DEPARTMENT OF ENERGY | Energy Efficiency & Renewable Energy

Predicting Envelope Leakage in Attached Dwellings

Prepared for:
The National Renewable Energy Laboratory
On behalf of the U.S. Department of Energy's Building America Program
Office of Energy Efficiency and Renewable Energy
15013 Denver West Parkway
Golden, CO 80401
NREL Contract No. DE-AC36-08GO28308

Prepared by:
O. Faakye, L. Arena, D. Griffiths
Steven Winter Associates, Inc.
of the
Consortium for Advanced Residential Buildings (CARB)
61 Washington Street
Norwalk, CT 06854

NREL Technical Monitor: Cheryn Metzger
Prepared under Subcontract No. KNDJ-0-40342-03

July 2013

[This page left blank]

Contents

List of Figures

Unless otherwise noted, all figures were created by CARB.

List of Tables

Unless otherwise noted, all tables were created by CARB.

Definitions

CARB	Consortium for Advanced Residential Buildings
CFM50	Cubic feet per minute of airflow needed to create a change in building pressure of 50 Pascal
ESA	Exposed surface area of building envelope
FG	Fully guarded blower door value (CFM50)
LTO	Leakage to outside (CFM50)
MVLR	Multivariable linear regression
PCC	Partial correlation coefficient
RCWTSA	Ratio of common wall to total surface area
RESATSA	Ratio of exposed surface area to total surface area
RF	Random Forest
R_{FS}	Ratio between fully guarded and solo blower door values
RWATESA	Ratio of window area to total exposed surface area
SO	Solo blower door value (CFM50)
TSA	Total surface area of building envelope

Acknowledgments

CARB would like to thank those who have enthusiastically supported this effort. A special thanks for the input of project advisors Michael Lubliner, Terry Brennan, Gary Nelson, Collin Olson, and Iain Walker. We are also grateful to Anthony Cox and Chase Counts of the New River Center for Energy Research and Training for sharing a treasure trove of data.

Executive Summary

The most common method for measuring air leakage is to use a single blower door to pressurize and/or depressurize the test unit. In detached housing, the test unit is the entire home and the single blower door measures air leakage to the outside. In attached housing, this “single unit,” “total,” or “solo” test method measures both the air leakage between adjacent units through common surfaces as well air infiltration or exfiltration across the exterior, noncommon surfaces of the enclosure. Measuring and minimizing this total leakage are recommended to avoid indoor air quality issues between units, reduce energy losses to the outside, reduce pressure differentials between units, and control stack effect. However, two significant limitations of the total leakage measurement in attached housing are:

- For retrofit work, if total leakage is assumed to be all to the outside, the energy benefits of air sealing can be significantly overpredicted.
- For new construction, the total leakage values may result in failing to meet an energy-based house tightness program criterion.

A practical method needs to be developed to quantify the fraction of total leakage for an attached dwelling that is to the outside. The consensus of a Building America Expert Meeting held in March 2012 confirmed the need for research to develop such a method.

Using blower door test data available from four multifamily projects, the framework for a simple algorithm based upon a solo blower door test result and a few basic dwelling unit characteristics has been outlined. While the subject dataset is very limited, preliminary analyses suggest that statistically significant predictors are present and can support the development of an algorithm. The key next step is to collect additional data for analyses and algorithm development so they may be more broadly applied with confidence.

The scope of this research is to investigate an approach for developing a viable simplified algorithm that can be used by contractors to assess energy efficiency program qualifications and/or compliance based upon solo test results. This report describes the work that has been done thus far and is intended for the building science research community that is familiar with blower door test protocols. The algorithm would not replace the more rigorous, and more accurate, guarded blower door method that is appropriate for building science research. Also, this research effort does not intend to suggest appropriate targets for maximum air leakage values.

1 Introduction and Background

Guarded blower door testing, or fully guarded (FG) testing, is a pressurization method that is often recommended for measuring air leakage to the outside. This method utilizes multiple blower doors to pressurize or depressurize adjacent spaces to the same level as the unit being tested; maintaining a neutral pressure across common walls, ceilings, and floors acts as a “guard” against air exchange between units. The measured air leakage in the test unit is air leakage to the outside. While preferred for assessing energy impact, this method is often not implemented because the equipment and labor requirements can be daunting. In retrofit situations where adjacent units may be occupied, simultaneous access can be logistically difficult. These challenges have been noted by other Building America researchers as well (Lyons 2013; Neuhauser et al. 2012; Ueno et al. 2012; Wytrykowska et al. 2012).

Whole-building testing is analogous to guarded blower door testing because both methods attempt to measure energy-related exterior envelope leakage. In terms of energy efficiency and utility bill savings, exterior envelope leakage is the key. But, like guarded blower door testing, whole-building blower door tests require substantial resources in equipment, personnel, and time (Hynek 2011).

The simpler and more common method for measuring air leakage in attached dwellings is to use a single blower door to pressurize and/or depressurize the test unit. This “single unit,” “total,” or “solo” (SO) test method measures the combination of air leakage between adjacent units through common surfaces as well as air leakage to the outside. Minimizing total leakage, or compartmentalization, is good practice to avoid indoor air quality issues between units, reduce energy use, reduce pressure differentials between units, and control stack effect. However, two significant limitations of the SO leakage test are:

- For retrofit work, if total leakage is assumed to be all to the outside, the energy benefits of air sealing can be significantly overpredicted.
- For new construction, the total leakage values may result in failing to meet an energy-based house tightness program criterion.

As described in a recently completed study by the Heschong Mahone Group (HMG 2012), the high cost and building occupant disruption of blower door testing deters utility program participants from pursuing building envelope improvements.

“Until a protocol is established that provides a clear path with limited costs and risks for participation, building owners will continue to be hesitant to participate in whole building upgrade (retrofit) programs.”

In March, a Building America Expert Meeting was held to discuss the various methods, barriers to implementation, and current practices for measuring air changes and/or envelope leakage in attached dwellings. Attendees included representatives from federal government agencies, weatherization industry trainers and practitioners, building science researchers, and national laboratories. Presentations on multifamily dwelling air leakage were given by several building science researchers to encourage discussion on understanding why and how air leakage is measured, the challenges and costs associated with the measurements, and the need to conduct research to improve the process.

A key objective of this meeting was to provide a forum for researchers and service providers to share information on their current practices and challenges and exchange ideas on opportunities for improvement. There was definitely a consensus for a simpler, but adequate, approach to assessing air leakage, both total and external.

A practical method needs to be developed to quantify the fraction of total leakage that is to the outside, and that is the objective of this research. With this simplification, measurement protocols can focus on the more practical single-dwelling unit total leakage test.

2 Research Objectives

The objective of the current research effort is to determine the viability of developing an algorithm that estimates the quantity of air leakage to the outside based upon a total leakage measurement and limited building and/or dwelling unit characteristic information.

The primary questions to be addressed by this research are:

- Conceptually, how would a method for estimating outside air leakage from a total leakage measurement work?
- Do preliminary analyses suggest that a simplified, reliable method can be developed?
- What additional research is needed to develop the concept into a viable, industry-accepted method?

3 Technical Approach

The theoretical concept for the simplified method is that a SO (total) leakage test result and a limited amount of building and/or apartment characteristic information can define a multiplier that provides the fraction of the SO (total) leakage that is to the outside.

The Expert Meeting in March was held to confirm interest and provide early guidance to the methodology development. Several meeting participants agreed to serve on an advisory panel, and an initial meeting of the panel was held in conjunction with the summer ASHRAE meeting in San Antonio. Suggestions for approaching the analysis were provided.

To demonstrate the viability of an outdoor-to-total leakage algorithm, analysis was performed on a limited set of data. Air leakage and building characteristic data were obtained from internal project resources at Steven Winter Associates, Inc. and Community Housing Partners' New River Center for Energy Research and Training. Datasets include building characteristic information, dwelling unit information, and blower test results, both SO and guarded.

FG and SO blower door test results were obtained from four multifamily low-rise projects. Table 1 shows a summary description of the projects. While the projects are not geographically diverse, they do represent different types of attached housing and different ages.

Table 1. Summary of Building Used

Project	Location	Type of Building	Number of Buildings	Number of Units	Year of Construction
1	Winchester, Virginia	Garden-style apartments	2	14	1989
2	Newport News, Virginia	Garden-style apartments	5	35	1980
3	Roanoke, Virginia	Rowhouses	4	22	1970
4	Staten Island, New York	Duplex rowhouses	4	41	2010

Projects 1 and 2 are both complexes of garden-style apartments in two- and three-story buildings. Construction is wood frame on slab. Project 3 is a complex of 100 row houses with two to four bedrooms in 10 two-story wood frame-on-slab buildings. Project 4 is a newly constructed ENERGY STAR® Home project using steel framing that presented air sealing challenges. Extensive blower door testing was done to achieve performance goals. These data were included in the analysis to provide a different type of construction. The two-family row houses or townhomes have a one-bedroom unit on the ground floor with a two-story, three-bedroom unit above.

4 Model Development Methodology

Three models were examined with primary emphasis on a multivariable linear regression (MVLR) model. Random Forest (RF) method and a simple ratio method of exposed surface area to total surface area (RESATSA) were the other models. The two models were studied with the purpose of confirming results from the MVLR model and exploring a simpler way of predicting leakage to outside from a SO test result. For all three models, the target response was the ratio of FG to SO test value. With this ratio, one can obtain the FG test value given the SO test value and vice versa.

The following section describes each model and their results. The tool used for modeling was R. R is a language and environment for statistical computing and graphics (R Development Core Team 2008).

4.1 Multivariate Linear Regression Analysis

Statistical analyses were performed to create a model that would predict FG blower door or leakage to outside (LTO) values, from SO blower door results. An MVLR model was chosen:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X + \varepsilon \qquad (1)$$

where

- Y = R_{FS} = predicted ratio of FG to SO.
- X_p = variables/predictors describing various building specifications, and p is the number of predictors being considered.
- β_p = partial correlation coefficient (PCC). It represents the change in R_{FS} associated with a unit increase in the value of X_p (p^{th} variable) when all other variables are kept constant.
- β_o = intercept. Geometrically, it represents the value of R_{FS}, where the regression surface (or plane) crosses the Y axis. Substantively, it is the expected value of R_{FS} when all the variable or predictors are equal to 0.
- ε = the deviation of the value R_{FS} from the mean value of the distribution given X_p. This error term may be conceived as representing (1) the effects on R_{FS} of variables not explicitly included in the equation, and (2) a residual random element in the dependent variable. The basic idea behind creating a linear accurate model is to minimize ε for each prediction.

Thirty-two variables were considered as possible candidates for creating the model. However, not all 32 variables could be used, since the goal of this research was to create a strong MVLR model with the smallest number of predictor variables that are also significant. Figure 1 shows a flowchart describing the procedure taken to identify the most significant predictors. The data sample used consisted of 112 observations containing 32 building specifications per apartment unit and measured R_{FS}.

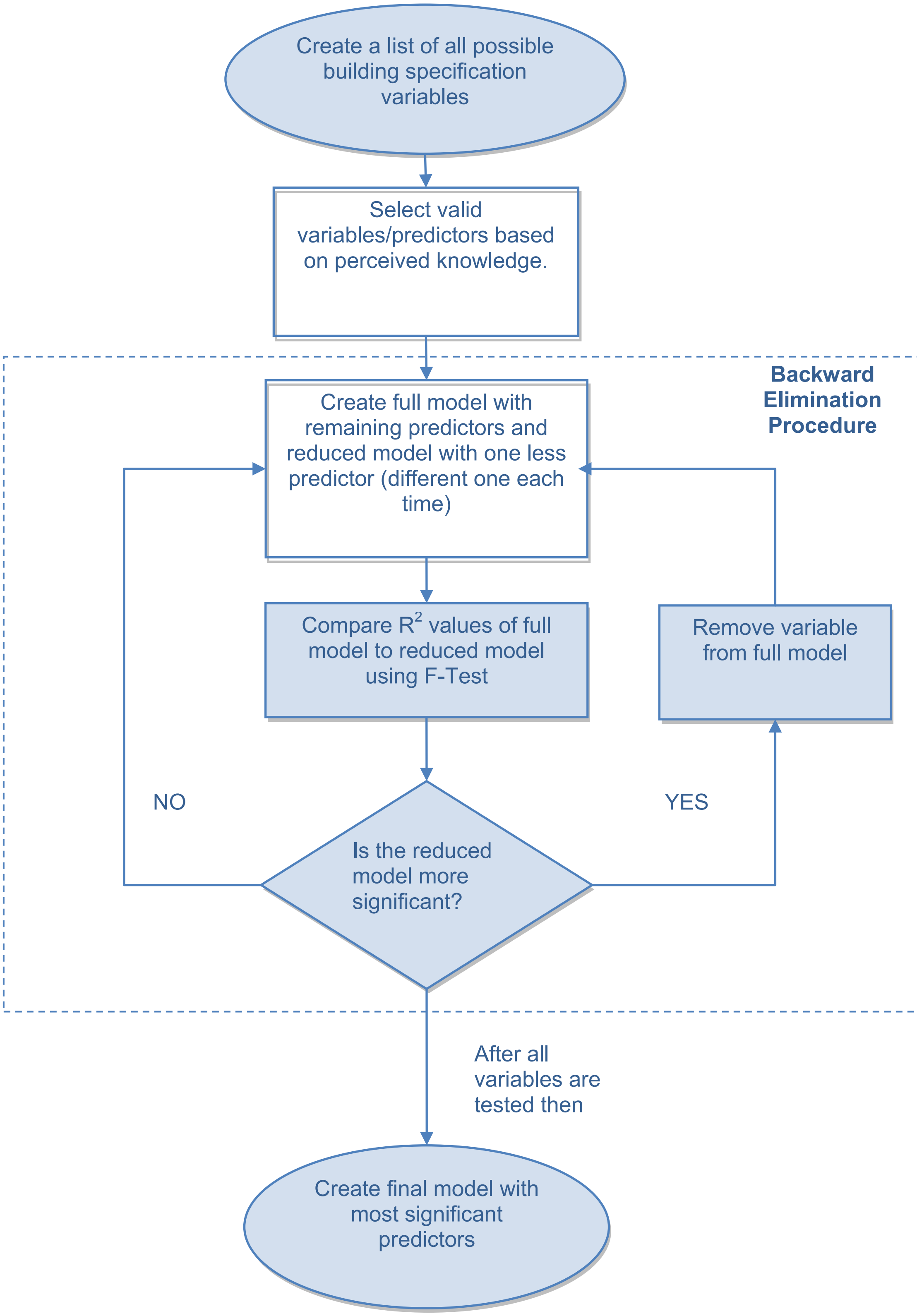


Figure 1. Process for creating model

The model was created using an iterative process of identifying and eliminating less significant variables. A full model was created from all valid variables, and a reduced model was created using one less predictor. The F-test was then used to investigate the significance of the variable removed by comparing the R^2 value of the full model to that of the reduced model. R^2 is a measurement of how much variation in the data is explained by a model. An appropriate null hypothesis for the F-test is:

$$H_0(\beta_1=0) \tag{2}$$

against the alternative that the predictors $\beta_{2\ldots p} \neq 0$. Thus, the reduced model in this case is:

$$R_{FS} = \beta_2 X_2 + \ldots + \beta_q X_q \tag{3}$$

The F value is calculated as:

$$F = \frac{\left(R_p^2 - R_q^2\right)/\left(p-q\right)}{\left(1-R_p^2\right)/\left(n-p-1\right)},\ \ d.f. = p-q, n-p-1 \tag{4}$$

where

R_p = the sample multiple correlation coefficient that is obtained when the full model with all p variables, is fitted to the data.

R_q = the sample multiple correlation coefficient when the reduced model is fitted with q specific variables.

n = number of observations

p = number of variables in full model

q = number of variable in reduced model

The calculated or observed F value from Equation 4 is compared to tabulated F values with degrees of freedom of p-q and n-p-1 at a 95% confidence level. If the calculated F is greater than the tabulated F, the null hypothesis, Equation 2, is rejected. This means the coefficient $\beta_1 \neq 0$ and is therefore significant to the model. If not, the model is recreated without the variable in question, β_1.

The process is then repeated for all variables in the model. This method of iterating through equations and dropping least significant predictors is also called the backward elimination procedure. Obtaining a limited number of predictor variables helps isolate the most important variables and provide a simple and reusable model for builders, raters, engineers, and architects performing blower door tests. A final model was then created using the most significant predictors from the statistical analysis.

In order to make the statistical analysis valid, the following necessary assumptions were made:

- Predictors are independent and normally distributed with mean zero and constant variance.
- Explanatory variables are nonstochastic: the values of the predictors were measured or obtained in advance.
- Predictors were measured without errors.

4.2 Predictor Variables Considered

A list of all variables considered can be found in Appendix A. In order to get a holistic sample of variables both quantitative and qualitative predictors were selected. Figure 2 shows an illustration of some of the variables considered for the model.

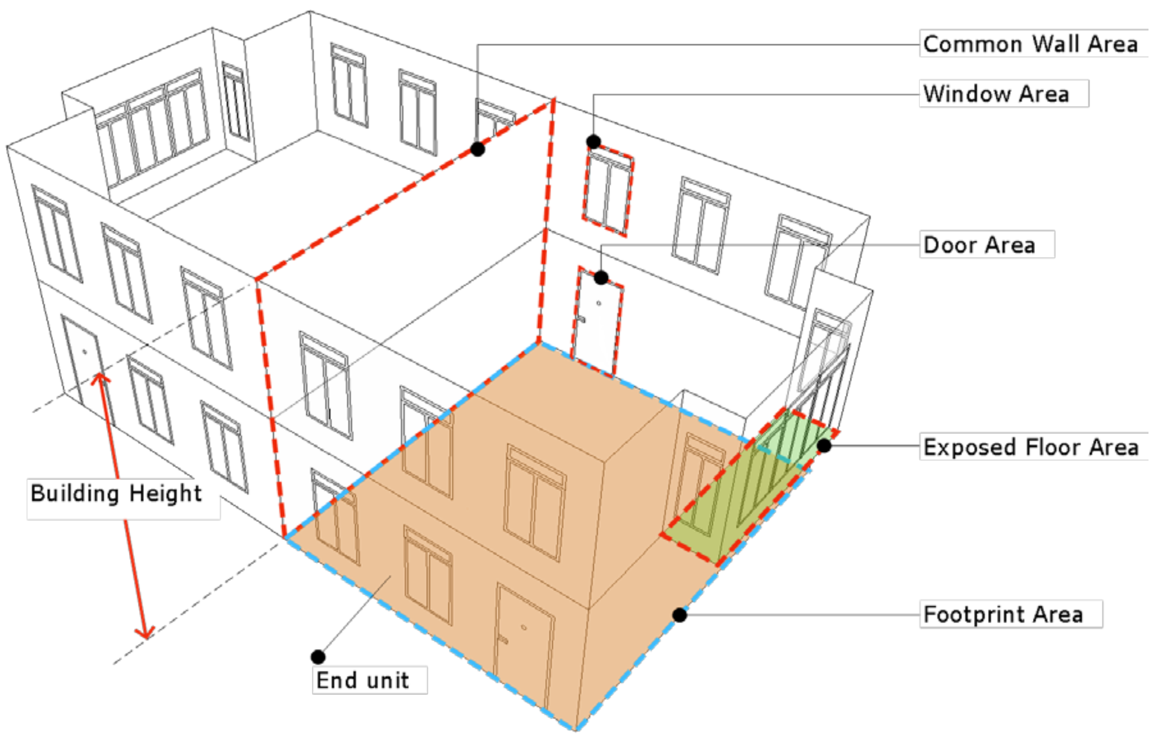


Figure 2. Illustration of typical building specification

Quantitative predictors describe building specifications that can be quantified such as footprint area, total wall area, and exposed wall area. Qualitative predictors describe categorical building specifications like duct location, type of wall insulation, and unit location. For every categorical variable used, a set of dummy variables was created to describe each level in the variable by using a dummy coding scheme (UCLA: Statistical Consulting Group). This coding scheme assigns a value of 1 or 0 to the dummy variable by comparing each level of the categorical variable to a fixed reference level. The numerical values of the dummy variables are not intended to represent quantitative ordering, but only serve to identify the levels in a categorical variable (Chatterjee and Bertram 1938). Table 2 shows an example of dummy variables and corresponding values created for Ductwork Location.

Table 2. Coding Scheme for Categorical Variables

Dummy Variables Number	Dummy Variables for Duct Location	Dummy_Var_1 Versus Dummy_Var_2	Dummy_Var_1 Versus Dummy_Var_3
1	Ductwork Location-None	0	0
2	Ductwork Location-Conditioned Space	1	0
3	Ductwork Location-Unconditioned Space	0	1

The fixed level in the dummy variable for Ductwork Location is "None" (Dummy_Var_1). This is case where a unit has no ducts. The reference point is seen as the point where all dummy variables are set to zero. As such, dummy variable Ductwork Location-None is represented when the two other Ductwork Location levels are set to zero. There were no units with ducts located in both conditioned and unconditioned space within the data analyzed. If there were such units, another level in the dummy variable would have been created for Ductwork Location-Conditioned Space and Ductwork Location Unconditioned Space.

4.3 Major Predictive Variables

The Consortium for Advanced Residential Buildings (CARB) considered 32 variables, Appendix A, that are likely to affect infiltration values and tried to populate them for each unit given the preliminary information received. Due to several variables having missing values and/or one value, the full list was trimmed to the 13 variables shown in Table 3. Examples of variables with one value include window types were all double hung and building framing types were all platforms, thus two variables would not be beneficial to the model.

Table 3. Initial MVLR Model Results

Variable Number	Variable
1	Ductwork location
2	Unit floor level
3	Unit location
4	Total exposed area
5	Window area
6	Exposed floor area
7	Exposed wall area
8	Footprint area
9	Total surface area
10	Ceiling area
11	Total shared surface area
12	Volume
13	Common wall area

Most of the variables shown in Table 3 are highly correlated with each other or collinear. For example, if footprint area is increased, the volume of the apartment unit also increases. This issue could cause redundant variables within a model. Thus, to reduce collinearity, new variables were created by creating ratios between variables. For example, the ratio of common wall to total surface area (RCWTSA) was created with the idea of capturing the relationship between common wall, standardized by total surface area, and R_{FS}. Another ratio created was the ratio of window areas to total exposed surface area (RWATESA). Several ratios could be created, such as total exposed area to total surface area (RTESATSA), total surface area to volume (RTSAV), and exposed floor area to footprint area (REFAFPA). These ratios create more interactions between variables in addition to the additive interaction facilitated in MVLR models. Different interactions between variables often improve the strength and accuracy of the model. By using the above-mentioned ratios, variables 4 to 13 from Table 3 were safely dropped. For the sake of having a concise description of the process of selecting the most significant variables, only two ratios along with the remaining valid categorical variables were selected. Future analyses will include more ratio variables and other interactions by multiplying, squaring, adding, or subtracting variables.

The remaining variables after elimination are:

1. RCWTSA, X_1
1. RWATESA , X_2
2. Unit Location
    - End, X_3
    - Interior, X_4
3. Duct Location
    - Duct Location – Conditioned Space, X_5
    - Duct Location – Unconditioned Space, X_6
    - Duct Location – None, X_7
4. Unit Level
    - Bottom Floor, X_8
    - Top Floor, X_9
    - Middle Floor, X_{10}

The first two predictors, RCWTSA (X_1) and RWATESA (X_2), are quantitative variables, whereas Unit Location, Duct Location and Unit Level are qualitative/categorical variables. Ductwork Location indicates whether the unit's ductwork is in conditioned or unconditioned space or not present at all. Unit Location indicates whether a unit is an end or interior unit. Unit Level indicates whether a unit is on the top, bottom, or middle floor of an apartment building. Due to the small sample size, two-story units were described as top units in order to reduce the number of variable predictors used for such a small data sample. Future modeling with more data will certainly include another level in the dummy variable describing multistory units. As mentioned above, categorical variables are usually represented with dummy variables to denote each level in the variable. As such, Unit Location uses two dummy variables (X_3 and X_4) and Ductwork Location (X_5, X_6 and X_7) and Unit Level (X_8, X_9 and X_{10}) use three dummy variables each, resulting in a total of ten predictors.

Thus, the MVLR equation becomes:

$$R_{FS} = \beta_0 + X_1\beta_1 + X_2\beta_2 + \ldots + X_{10}\beta_{10} \quad 5$$

Knowing the predicted value of R_{FS} from Equation 5, FG or LTO at CFM50 can be predicted using Equation 6, which multiplies predicted R_{FS} by the measured SO blower door value:

$$FG = R_{FS} \times SO \quad 6$$

4.4 Multivariable Linear Model Results

A summary of the MVLR model results is shown in Table 4. The Pr(>|t|) column is used to assess the significance of each variable. Pr(>|t|) is the probability that the PCC of a variable is equal to zero, according to the *t* test. The *t*-test, is used in this analysis to determine the significance of a variable's coefficient. For dummy variables, the Pr(>|t|) value is the probability that the difference between the dummy variable and its reference point is zero.

As a general rule, a 5% or less probability is the acceptable benchmark to reject the null hypothesis that a predictor variable's PPC is zero; that is, the variable is insignificant. From Table 4, the probability that the difference between Ductwork Location-Conditioned Space and None is zero, as such there is no significant difference between the two Ductwork Location dummy variables. Since Ductwork Location is a variable that often affects infiltration values, CARB decided to leave it in the model; with a larger data sample the significance of dummy variable Ductwork Location will likely be realized. The intercept is the least significant nondummy predictor and it is not surprising that it has the lowest PCC and confidence level. Insignificant variables will be dropped at this point, however, since this reports seeks to demonstrate more about the statistical process than the production of a finalized model, the intercept will not be discarded.

Table 4. MVLR Model Results

Predictors/Variables		Estimate	Pr(>|t|)	Confidence Level
	Intercept	0.179	0.373	< 90.0%
X_1	RCWTSA	0.712	0.008	95.0%
X_2	RWATESA	–0.904	0.009	99.0%
X_3	Unit Location-End	0.139	0.003	95.0%
X_5	Ductwork Location-Conditioned Space	0.026	0.704	< 90.0%
X_6	Ductwork Location-Unconditioned Space	0.156	6.63e-08	99.9%
X_8	Unit Floor Level: Bottom	0.304	1.20e-06	99.9%
X_9	Unit Floor Level: Top	0.230	3.51e-06	99.9%

After evaluating the significance of each variable's coefficient, the backward elimination procedure was performed to assess the significance of each variable, to the model. As discussed in Section 4.1, the F-test was done to determine the significance of the reduced model with respect to the full model at a 95% confidence level. The F-test is a measure of how relevant a variable is to the model (Chatterjee and Bertram 1938). The F-value calculated for each reduced model was less than the corresponding tabulated F-values. As such, each null hypothesis was rejected, implying that all variables in the full model are significant to the model. It must be noted that the significance of these predictors could change with a different sample data.

The model was finally recreated and shown in Equation 7:

$$R_{FS} = 0.18 + 0.71X_1 – 0.90X_2 + 0.14X_3 - 0.03X_5 + 0.16X_6 + 0.30X_8 + 0.23X_9 \qquad 7$$

As mentioned above, dummy variables were created using a coding scheme, which chooses a level in the categorical variable as a reference point. The reference point is seen as the point when all dummy variables are set to zero. As such, dummy variable Unit Location-Interior (X4) is represented when Unit Location-End (X3) is zero, Ductwork Location-None (X7) is represented when the two other Ductwork Location levels (X5 and X6)) are set to zero, and Unit Floor Level is middle (X10), when Top (X9) and Bottom (X8) are set to zero. Consequently, depending on the values of the categorical variables for each input observation, the MVLR model will change as shown in Table 5.

Table 5. MVLR Model Equations

Unit Location	Duct Location	Unit Floor Level	MVLR Model (Unit and Duct Location), R_{FS}
Interior (X_4)	None (X_7)	Middle (X_{10})	$0.18 + 0.71X_1 – 0.90X_2$
End (X_3)	None (X_7)	Middle (X_{10})	$0.18 + 0.71X_1 – 0.90X_2 + 0.14X_3$
Interior (X_4)	Ductwork Location-Conditioned Space (X_5)	Middle (X_{10})	$0.18 + 0.71X_1 – 0.\ 90X_2 – 0.03X_5$
End (X_3)	Ductwork Location-Conditioned Space (X_5)	Middle (X_{10})	$0.18 + 0.71X_1 – 0.\ 90X_2 + 0.14X_3 – 0.03X_5$
Interior (X_4)	Ductwork Location-Unconditioned Space (X_6)	Middle (X_{10})	$0.18 + 0.71X_1 – 0.90X_2 + 0.16X_6$
End (X_3)	Ductwork Location-Unconditioned Space (X_6)	Middle (X_{10})	$0.18 + 0.71X_1 – 0.90X_2 + 0.14X_3 + 0.16X_6$

Figure 3 shows the final model predicted values against measured values. The red diagonal line represents an accurate predictive model. The closer the data points are to the red line the more accurate the model, thus the MVLR model is fairly accurate.

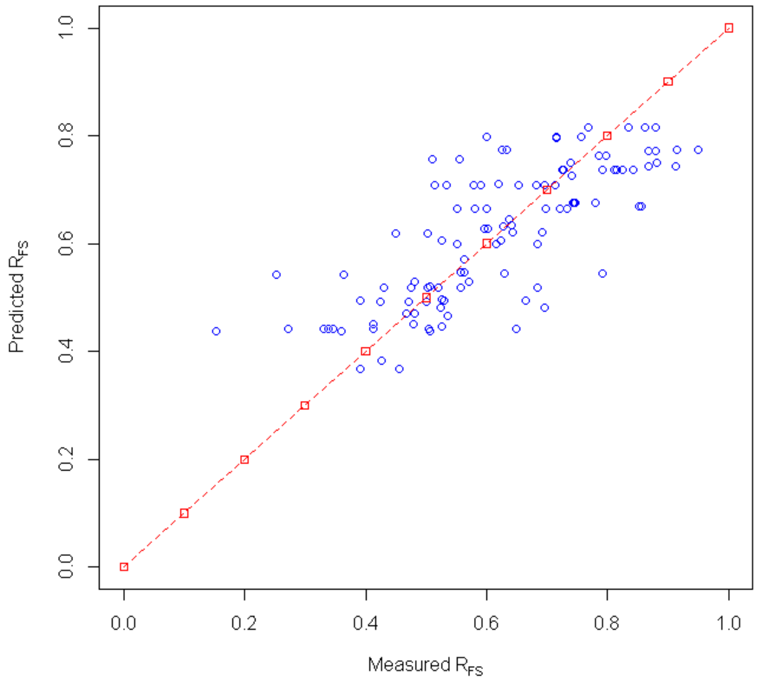


Figure 3. Plot of MVLR model fitted against measured R_{FS}

Statistical tests and confidence intervals used above are based on the assumption of normal errors. Figure 4 shows four plots that validate the assumptions made. The top left plot shows residuals against the fitted values with data points randomly distributed around the dotted horizontal line at 0 and lie between ± 0.3 on the vertical axes. This shows that there is no discernible pattern to the distribution of residuals, supporting the assumption of normality in the data. The scale-location plot in the lower left shows the square root of the standardized residuals (square root of relative error) as a function of the fitted values. Again, for normally distributed data, there should be no obvious trend and that is also confirmed in this plot. The normal Q-Q is an informal graphical test of the hypothesis that a data sequence is normally distributed. If the plotted data points fall exactly on dotted diagonal line, the normal distribution assumption holds (Khattree and Naik 1999). In this case, data points align fairly close to the diagonal line with a few outliers at the ends, thus the assumption of normality still stands. Finally, the plot in the lower right shows each data point's leverage, which is a measure of its importance in determining the regression result. Superimposed on the plot are contour lines for Cook's distance, another measure of the importance of each observation to the regression (Van Steen et al. 2001). Smaller distances mean that removing the observation has little effect on the regression results. Distances larger than 1, (not shown because of limited axis range) are suspicious and suggest the presence of possible outliers or a poor model. This MVLR model shows only three obvious outliers (94th, 95th, and 101st observations) out of 112 observations, as such the importance each observation is evenly spread.

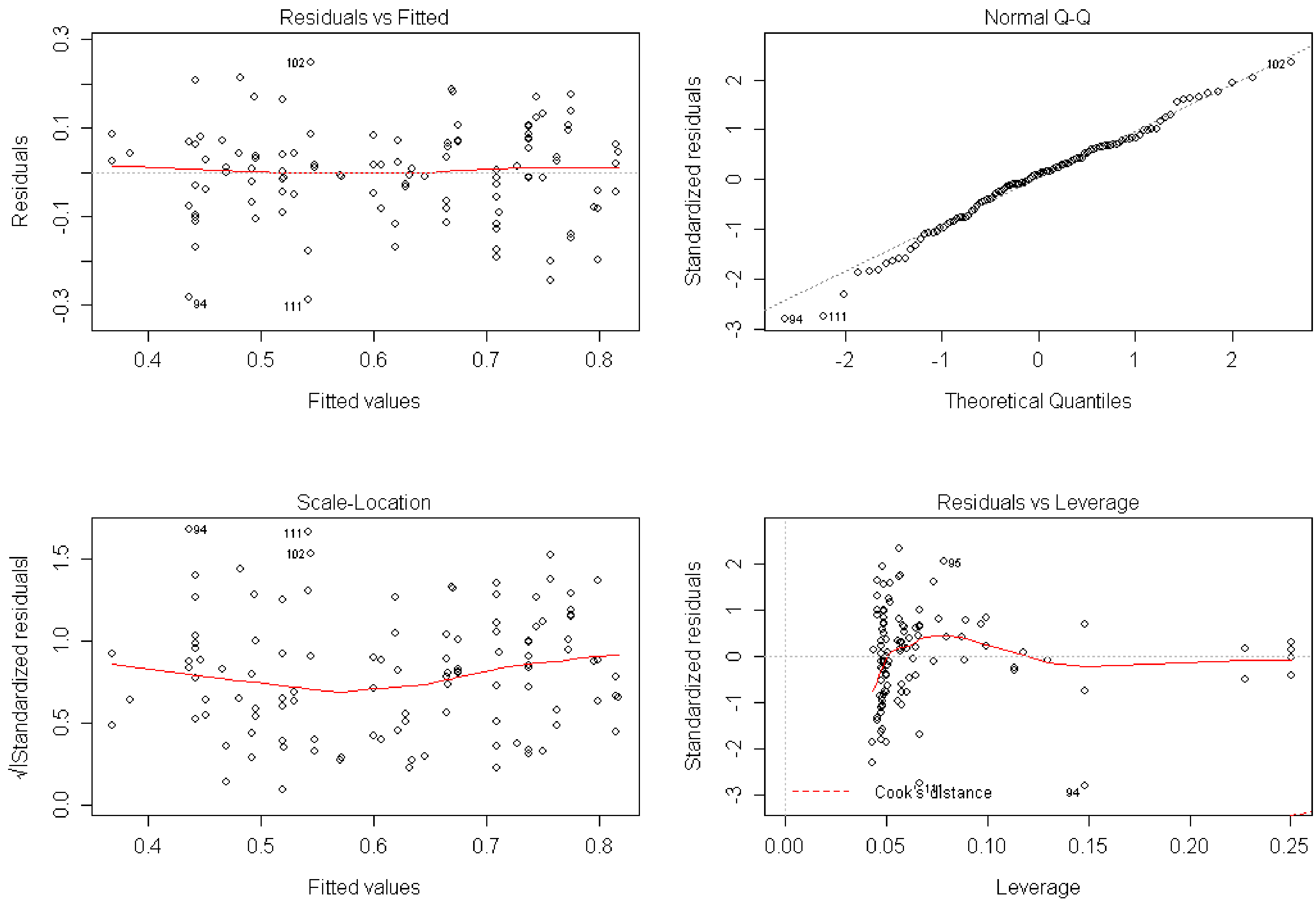


Figure 4. Plots validating linear model

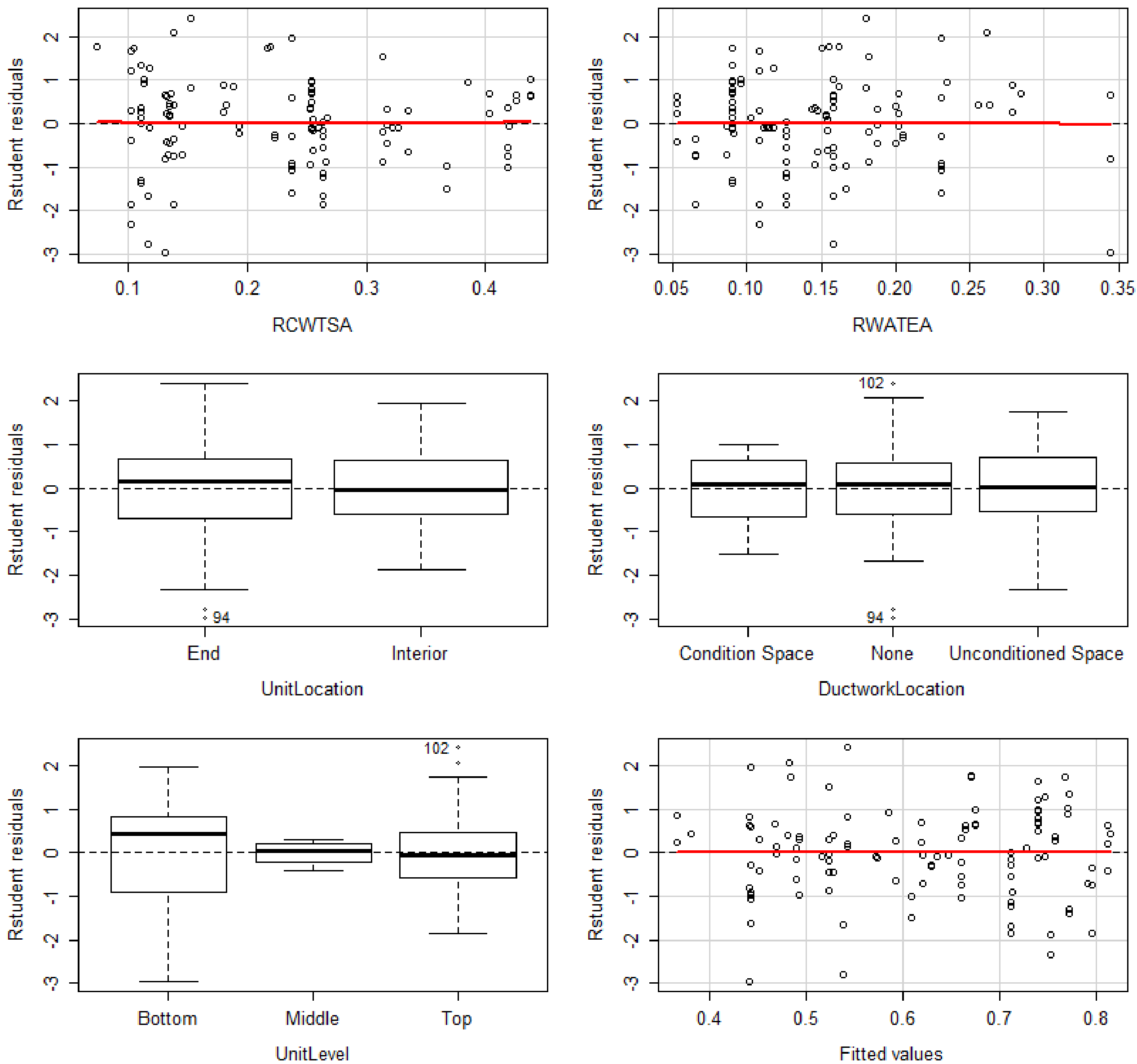


Figure 5. Plot of Rstudent residuals against predictors

Figure 5 displays the Rstudent residuals against each predictor variable. Rstudent residuals, also known as studentized residuals, is a standard way of measuring each predictor variable's contribution to the residuals of a model (Cook and Weisberg 1982). These graphs above are used to determine nonlinearity between a variable and predicted values, R_{FS}, as well as outliers in data points. For quantitative variables RCWTSA (X_1) on the top left, RWATESA (X_2) on the top right, and fitted values on the bottom right, their data points are evenly distributed across the zero line and show no systematic pattern. As such, it is concluded that there is a linear relation between selected predicted variables and the response, R_{FS}. For categorical variables not much can be said about linearity, however, one can notice distinct outliers. These outliers could be the result of discrepancies in the model predictions. Removing these points will slightly improve the model; however, due to limitations in sample size, this research focuses more on the analytical methodology of arriving at an accurate and reusable model than on creating a finalized model. Next CARB investigated the possibility of having a predicting model with other modeling types.

4.4.1 Random Forest Approach

Random Forest is a statistical modeling tool that implements Breiman's random forest algorithm for classification and regression.

> It produces a regression tree which is built recursively from the data sample into more and more homogeneous groups, until a terminal node is reached. Each split is based on the values of one variable and is selected according to a splitting criterion. Once a tree has been built, the response for any observation can be predicted by following the path from the root node down to the appropriate terminal node of the tree. Based on the observed values for the splitting variables, the predicted response value is simply the average response in that terminal node (Grömping 2009).

As with other highly computational procedures, RF does not have a simple representation such as a formula (e.g., linear regression model) for the relationship between the predictor variables and the predicted values (Cutler et al. 2007). This makes interpretation and field application difficult. Appendix B shows an example of a regression tree that describes the RF model for this research. Random Forest was used in this research only to validate the results from the MVLR model. For more details about this modeling tool see Liaw and Weiner (2002).

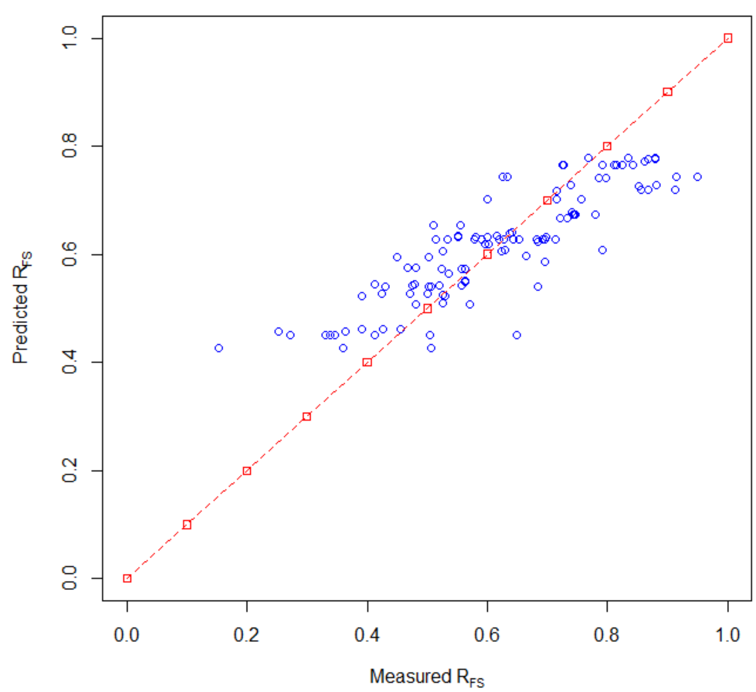


Figure 6. RF model results

A plot of predicted against measured R_{FS} is shown in Figure 6, and once again a comparison of the data points with the red line gives a visual of the accuracy of the model. Since the data points are evenly spread and close to the diagonal red line, it is concluded that the RF model is also fairly accurate.

4.4.2 Area Weighted Approach

The area weighted approach has been suggested as a very simple method that may be good enough. The available sample data were tested against the validity of this approach. The method basically describes R_{FS} as being equal to the ratio of exposed surface area, ESA, to total surface area, TSA, of the apartment envelope:

$$R_{FS} = \frac{ESA}{TSA} = \frac{FG}{SO} \tag{8}$$

ESA is the area of building including walls, windows, doors, and roofs in contact with ambient conditions. The reason behind exploring this model was to investigate a much simpler approach of predicting the R_{FS}.

Predicted *FG* can then be calculated as:

$$FG = \frac{ESA}{TSA} \times SO \tag{9}$$

$$FG = RESATSA \times SO \tag{10}$$

where

$RESATSA$ = Ratio of exposed surface area to total surface area

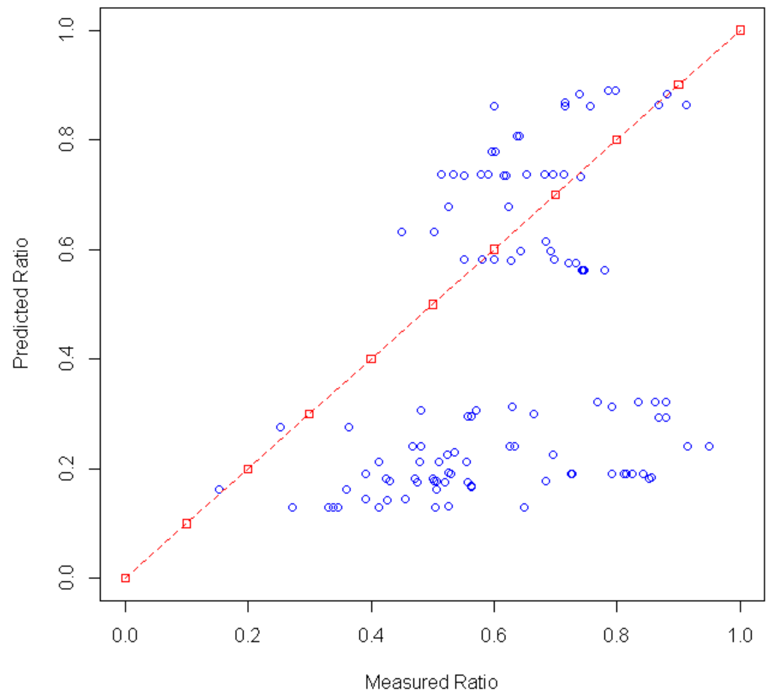


Figure 7. RESATSA model results

Figure 7 shows the RESATSA model results by plotting predicted against measured R_{FS}. The graph shows a wider spread in data points that are also relatively far from the red diagonal line. Thus, the predictive ability of this RESATSA model is limited. A detailed comparison of the three models' results and their relative accuracy is described below.

5 Comparing Models

Each model, with the exception of RF, has a simple linear equation that can be used to calculate predicted FG by multiplying SO by the predicted R_{FS}. There are several ways to measure quality of a model. In this research two parameters are used (Table 6).

Table 6. Summary of Regression Accuracy

Model	Residual Standard Error	Adjusted R^2
RF	0.011	0.580
MVLR	0.109	0.553
RESATSA	0.249	0.065

Residual standard error explains the discrepancy between the measured and predicted R_{FS}, taking into consideration the degrees of freedom, which is the difference between the numbers of observations in the data sample and the number of predictor variables, including the intercept. The lower the residual standard error is, the stronger the model. RF shows the lowest residual standard error, 0.011, meaning one should expect the smallest error between predicted and measured R_{FS} when using RF model. This is followed by MVLR model and then RESATSA model.

Adjusted R^2 is obtained from R^2 value however, adjusted R^2 accounts for the degrees of freedom in the model. The value of R^2 ranges from 0 to 1 and the closer the value to 1 the more accurate the model is. A larger adjusted R^2 value implies that a greater proportion of the variance in the data can be explained by the model. MVLR shows an adjusted R^2 value, 0.553. Adjusted R^2 of 0.55, for a model created from measured values that have a wide range of uncertainly, is statically justifiable but certainly a higher R^2 will be most preferred. Random Forest model had the highest R^2 value, 0.58, but due to its high complexity, the model's practicality in the field may not be feasible. With the lowest R^2 value, REASTSA model comes out to be the worst, explaining only 6.5% of the variance in measured response.

The final MVLR model includes variables RCWTSA (X_1), RWATESA (X_2), Unit Location-End (X_3), Unit Location-Interior(X_4), Ductwork Location-Conditioned Space (X_5), Ductwork Location-Unconditioned Space (X_6), Unit Floor Level-Top (X_8) and Unit Floor Level-Bottom (X_9). Although the model's variables and coefficients could change depending on size and variation in the sample dataset, the selected variables and coefficients, in this case are arguably credible.

The predictability of this MVLR model is limited to the data sample and cannot be applied in all situations. Sample data in this research lack variety in terms of climate zone (only mixed-humid locations), foundation type (only slab-on-grade), and wall insulation (only dense pack and loose fiber glass), to mention a few. As such the model is expected to do poorly with data from other climate zones with different foundation and wall insulation types. More data with more variations would help obtain a stronger and more reliable model. There is also the possibility that the significance of predictor variables would change.

Moreover, a good variable selection procedure will result in several sets rather than a so called single "best" set of predictor variables (Chatterjee and Bertram 1938). With a larger, well-varied data sample, various sets of adequate equations would be obtained. The best choice of equation will boil down to which equation has the predictor variables easily obtained.

In this research, CARB basically demonstrated the possibility of having a correlation between building geometric characteristics, location of duct, and the relationship of the dwelling to adjacent dwellings and $R_{FS.}$ This was achieved by creating three models with all 112 observations and analyzing the correlation between select significant variables and the response, R_{FS}. To varying extents, all three models prove a predictable relationship between building specifications and R_{FS}. Considering the ease of usability in the field, the most practical model was the MVLR model. The result of this research was certainly limited by the relatively small and less varied sample data size. CARB's goal is to create a model that has a strong predictive power in different climate zones, construction types, and building specifications, as such this model is still in development stages. The objective of this research was to demonstrate the MVLR model's viability and development process.

6 Preliminary Test Case

A preliminary test of the MVLR model was run on blower door results for four attached rowhouses from the King County Housing Authority in Washington (Figure 8).

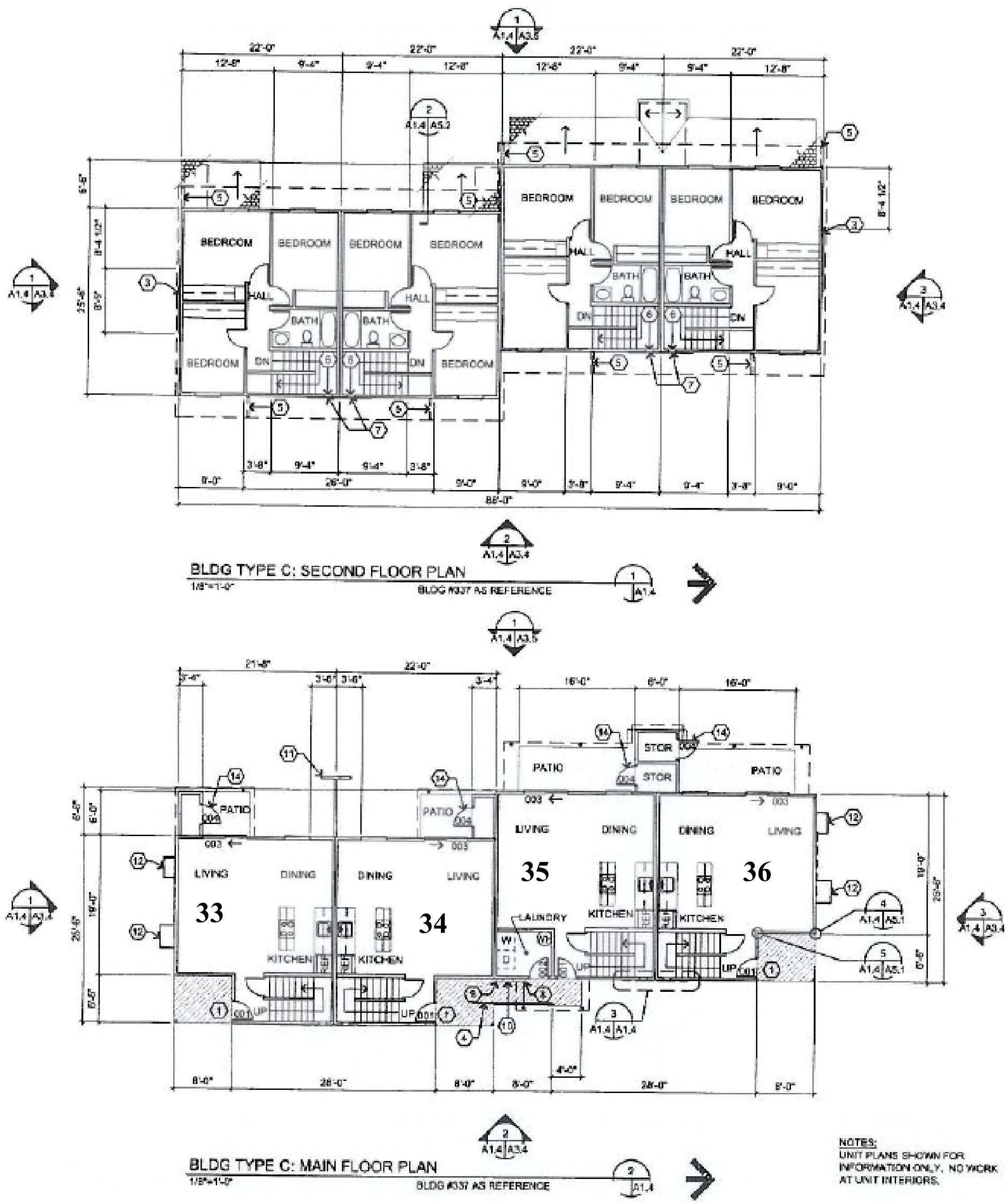


Figure 8. Floor plan of King County Housing Authority apartment units

The data consisted of SO test results for each unit, pre- and post-retrofit, and a whole-building leakage test using multiple blower door systems. While we did not have SO and FG tests for each unit, we thought these data could still be used for demonstration purposes. Ideally, the sum of FG tests on each unit would be equivalent to the whole-building leakage test result.

For this demonstration, we used the SO test data (left side of Table 7) to predict FG for each unit (right side) using the preliminary model, Equation 7, and then summed the predicted FGs and compared the results to the measured whole-building leakage values.

Table 7. MVLR Model Predictions Versus Sample Data

Unit No.	Measured SO (CFM 50)		Predicted R_{FS}	Predicted FG (CFM 50)	
	Pre-Retrofit	Post-Retrofit		Pre-Retrofit	Post-Retrofit
A	930	580	0.71 ± 0.07	660 ± 70	410 ± 40
B	900	720	0.78 ± 0.05	700 ± 40	560 ± 40
C	930	620	0.92 ± 0.09	850 ± 100	570 ± 60
D	900	590	0.85 ± 0.06	760 ± 50	500 ± 30
Measured LTO Whole Bldg. (All 4 Units)	**3,340**	**1,930**	**Predicted FG Whole Bldg. (All 4 Units)**	**2,980 ± 260**	**2,040 ± 170**

Measured post-FG for the whole building, 1,930 CFM, falls within the range of predicted FG 2,040 ± 170 CFM, whereas for pre-infiltration, 3,340 CFM is just above the upper limit of the predicted range of 2,980 ± 260.

The result above is debatably a decent approximation, bearing in mind the limited dataset used to develop this model. Note the differences indicated between the pre- and post-retrofit SO and whole-building test results are not intuitive. It is not relevant to this exercise, but does suggest that there is uncertainty in blower door testing and its interpretation. The point of this exercise was to demonstrate the prediction methodology proposed.

7 Discussion

The analyses to date are encouraging, but given the very limited dataset studied, it is too early to draw conclusions. To apply the model to most any type of low-rise attached housing, the dataset from which the model is developed needs to include variety in age, construction materials, and configuration.

Our next step will be to identify and gather significantly more data from organizations that have conducted both SO and FG or whole-building testing on projects. This somewhat redundant testing is not commonly done, but we believe that research, weatherization training, and utility program design and evaluation firms may be sources. We also know of instances where SO tests were initially performed, but when compliance targets were not achieved, a whole-building test was performed.

Through continued research we will attempt to answer the following questions:

- Does the rigor of the predictive model hold up through analysis of more buildings, different types of construction, different climate regions, etc.? As the dataset becomes more varied, does it become more difficult to predict with acceptable accuracy the leakage to outside for a particular dwelling?
- Do the predictive variables change? Are more predictive variables needed to achieve acceptable accuracy? Is the model still simple enough to be useful?
- Are different models necessary for predicting pre- and post-retrofit conditions? It’s quite possible that retrofit activity will address with some consistency certain sources of leakage. This possible bias might warrant a separate model or could be within the noise of a single model. If enough data can be collected for both pre- and post-retrofit testing, this will be examined.
- How good is good enough? A question that was discussed without resolution during the Expert Meeting is, How accurate does the estimate of outside air leakage need to be? The value is used for many purposes, including determining the cost effectiveness of air sealing and the adequacy of fresh air for ventilation. Perhaps the algorithm would be most appropriate as a screening tool. Outcomes would be yes, no, or maybe. The “maybe” outcomes would require more rigorous testing. CARB will seek input on this question from project advisors and efficiency program managers.

8 Conclusion

The research performed to date was successful with the following key findings:

- Conceptually, how would the new method work? The framework for a simple algorithm based upon a SO blower door test result and a few basic dwelling unit characteristics has been outlined (Section 4.4, Table 5).
- Do preliminary analyses suggest that a simplified, reliable method can be developed? Yes, while the subject dataset is very limited, preliminary analyses suggest that statistically significant predictors are present and can support the development of an algorithm.
- What additional research is needed to develop the concept into a viable, industry-accepted method? The next step is to collect additional data for analysis and algorithm development so they may be more broadly applied with confidence.

References

Chatterjee, S.; Bertram, P. (1938). *Regression Analysis by Example*. 2nd Edition. New York: John Wiley & Sons, Inc.

Cook, R. D.; Weisberg, S. (1982). *Residuals and Influence in Regression*. New York: Chapman and Hall.

Cutler, D.R.; Edwards, Jr., T.C.; Beard, K.H.; Cutler, A.; Hess, K.T.; Gibson, J.; Lawler, J.J. (2007). "Random Forests for Classification in Ecology." *Ecology*, 88(11):2783–2792.

Grömping, U. (2009). "Variable Importance Assessment in Regression: Linear Regression Versus Random Forest." *The American Statistician* 63(4):308–319.

HMG. (2012). *Study of Multi-Family Air Leakage Testing Strategies*, Final Report, HMG Project Number 1014. Heschong Mahone Group, Inc., August 10, 2012.

Hynek, D. (2011). "Blower Door Testing in Multifamily Buildings," *Home Energy Magazine* September/October 2011. www.homeenergy.org/show/article/id/1711.

Khattree, R.; Naik, D.N. (1999). Applied Multivariate Statistics With SAS Software. SAS Press/John Wiley & Sons.

Liaw, A.; Wiener, M. (2002). "Classification and Regression by randomForest." *R news* 2(3):18–22. Accessed November 26, 2012: www.webchem.science.ru.nl:8080/PRiNS/rF.pdf.

Lyons, J. (2013). *Short-Term Test Results: Multifamily Home Deep Energy Efficiency Retrofit*. Golden, CO: NREL/SR-5500-56460. www.nrel.gov/docs/fy13osti/56460.pdf.

Neuhauser, K.; Bergey, D.; Osser, R. (2012). *Leveraging Limited Scope for Maximum Benefit in Occupied Renovation of Uninsulated Cold Climate Multifamily Housing*. http://apps1.eere.energy.gov/ buildings/publications/pdfs/building_america/renovation_uninsul_multifamily.pdf

R Development Core Team. (2008). R: A Language and Environment for Statistical Computing, Vienna, Austria: R Foundation for Statistical Computing. Available at www.R-project.org, ISBN 3-900051-07-0.

UCLA: Statistical Consulting Group. Coding for Categorical Variables in Regression Models, www.ats.ucla.edu/stat/r/modules/dummy_vars.htm. Accessed November 24, 2012.

Ueno, K.; Lstiburek, J.; Bergey, D. (2012). *Multifamily Ventilation Retrofit Strategies*. Golden, CO: National Renewable Energy Laboratory. NREL/SR-5500-56253. www.nrel.gov/docs/fy13osti/56253.pdf.

Van Steen, K.; Molenberghs, G.; Verbeke, G.; Thijs, H. (2001). "A Local Influence Approach to Sensitivity Analysis of Incomplete Longitudinal Ordinal Data." *Statistical Modeling* 1(2):125–142.

Wytrykowska, H.; Ueno, K.; Van Straaten, R. (2012). *Byggmeister Test Home: Cold Climate Multifamily Masonry Building Condition Assessment and Retrofit Analysis*. http://apps1.eere.energy.gov/ buildings/publications/pdfs/building_america/byggmeister_test_home.pdf%20.

Appendix A: Predictor Variables Considered

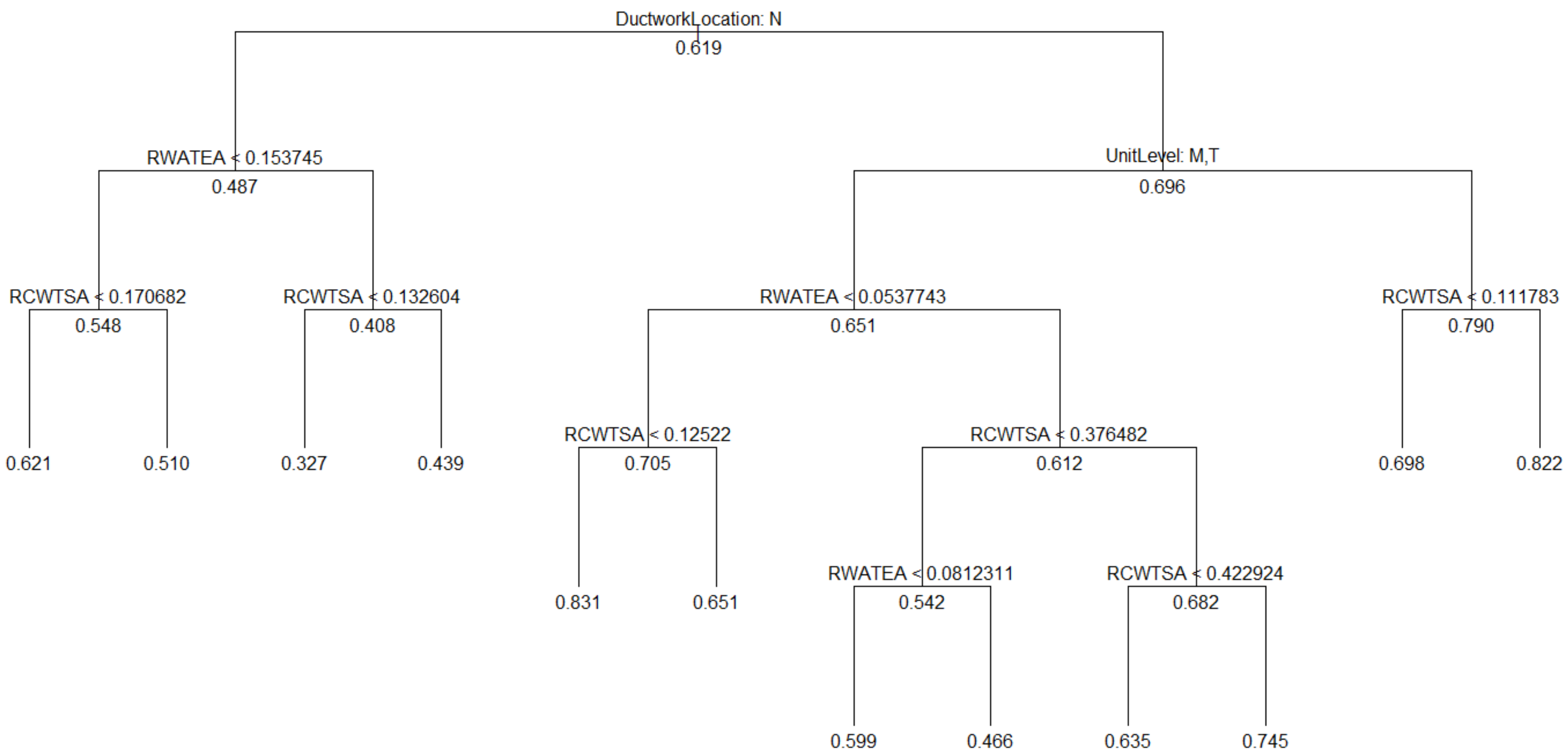
Variable	Description	Format
Building and Unit Information		
ApartmentComplex	Name of apartment complex	Text
BuildingNumber	Building Number	Text
BuildingType	Building Type usu. a digit or letter (optional)	Text
UnitConfiguration	How are the unit set up in the apt complex	Text
AgeOfBuilding	Age of building of the apartment complex	Number
Numberoffloors	Number of floors in the apartment complex	Number
Numberofunits	Number of units in the apartment complex	Number
Quantitative Specifications		
FootPrintArea	Foot Print Area of Unit	Number
CeilingArea	Ceiling Area of Unit	Number
ExposedFloorArea	Exposed flow area of Unit, ex cantilever	Number
ExposedCeilingArea	Exposed ceiling area usu. for a top floor	Number
ExposedWallArea	Exposed wall area i.e. wall in contact with outside	Number
TotalExposedArea	Same as wall area but includes ceiling for top unit	Number
InteriorUnitHeight	Height of Unit	Number
Volume	Volume of Unit	Number
CommonWallArea	Area of shared wall	Number
TotalSharedSurfaceArea	Total area of share surface include ceiling if there is a unit on top	Number
WindowArea	Total window area	Number
NumberofPanes	Number of panes in the windows	Number
TotalWallArea	Total wall area	Number
DoorArea	Total door area	Number

Qualitative Specifications

Unit Story Level	Top, Bottom or Middle	Top, Bottom or Middle
Unit Location	End, Interior	End, Interior
FoundationType	Type of foundation	Full Basement Crawl space Slab-on-Grade
ContructionMaterial	Material used for the framing of the building	Wood Steel Frame Masonry
BuildingFramingType	Building framing type	Balloon Platform
WallInsulation	Type of all insulation used	Blanket: batts and rolls Foam board Insulating concrete forms (ICFs) Loose-fill cellulose Loose-fill fiber glass Reflective/Bubble Rigid fibrous Sprayed foam Structural insulated panels (SIPs) Dense Pack
RoofInsulation	Type of roof insulation used	Same as WallInsulation
WindowType	Type of window is used	Single Hung Double Hung Casement Bay/Bow Awning Slider
WindowFrame	Window framing material	Metal Vinyl Fiberglass Wood
Siding	Wall siding material	Stucco Vinyl Aluminum Wood Stone Brick Mixed
CommonWallConstruction	Common wall construction material	Shaft Wall (2hr) Double Wall (2hr)

		Double Layer Drywall (1hr) Concrete/Brick (2hr)
Heating	Space heating system	Atmospheric Sealed Combustion Resistance
Cooling	Space cooling distribution system	Ducted Non-ducted
DuctworkLocation	Location of duct work	Conditioned Space Unconditioned Space

Appendix B: Random Forest Output



U.S. DEPARTMENT OF ENERGY | Energy Efficiency & Renewable Energy

DOE/GO-102013-3939 ▪ July 2013

Printed with a renewable-source ink on paper containing at least 50% wastepaper, including 10% post-consumer waste.