CS-BIGS 1(2): 153-163 http://www.bentley.edu/csbigs/vol1-2/erinjeri.pdf

Statistical Analysis of Three Sets of Residential Duct Leakage Measurements

Jinson J. Erinjeri
Louisiana Tech University, USA

Raja Nassar
Louisiana Tech University, USA

Myron Katz
Wisznia Associates, USA

Norman M. Witriol
Louisiana Tech University, USA

Residential HVAC (Heating Ventilation and Air Conditioning) duct-leakage-to-outside is the leakage of conditioned air from the HAVC ducts to the unconditioned space in residential homes mainly measured in cubic feet per minute (CFM). In this paper, the three most commonly used test methods (Manual, APT, and Combined) to determine residential duct-leakage-to-outside are compared using parametric and non-parametric statistical techniques. The objective of this study is to determine if there exist any differences between the various test methods and to predict the results of a test method from the other. Line of Equality, Linear Regression Analysis, Paired-t and Wilcoxon paired-sample tests were employed to compare the three methods. These results were then used to determine the mean duct-leakage-to-outside in north Louisiana, USA using a weighted sum approach. The exposition is accessible to a wide range of readers; a basic knowledge of statistics is needed. Some exposure to regression models is useful.

Introduction

The three most commonly employed methods utilized to determine residential Heating Ventilation and Air Conditioning (HVAC) associated duct-leakage-to-outside in the USA are:

(1) Manual (Manually collected Modified Subtraction): uses a blower door to depressurize a home, (testing personnel have to manually operate the blower's fan speed controller and record the CFM and pressure data). A digital manometer (DG-3 gauge) and blower door are used in a manual data collection algorithm that utilizes only a single depressurization test pressure. This is the most commonly used method in Louisiana for measuring duct-leakage.

(2) APT (Automated Performance Testing with Modified Subtraction): is a variant of the Manual method wherein hundreds of data are automatically collected at different pressures (Automated Performance Testing Manual, 2000, Minneapolis Blower Door Operation Manual, 1993). Since 1999, the ASTM[1] house-leakiness testing standard requires automated data collection equipment (ASTM, 1999). The APT method differs from the Manual method in its use of APT™ hardware

[1] American Society for Testing and Materials

and Tectite[TM] software to control the blower door's fan speed and automatically collect data at a series of pressures.

(3) Combined (Duct Blaster assisted by blower door): uses a fan (e.g., Duct Blaster[TM]) to directly pressurize the duct system to a set pressure and a blower door to pressurize the home to the same pressure. In 1994, ASTM accepted this method as the duct-leakage-to-outside testing standard (ASTM, 1994).

All the above three test methods used in determining residential duct-leakage-to-outside are measured in cubic feet per minute (CFM), the unit of flow rate of air in air-conditioning systems. The three test methods described above were attempted on 55 homes in northern Louisiana, USA, and, using the various algorithms associated with each method, measured the same parameter, duct-leakage-to-outside (Witriol et al., 2003). However in actuality, each of the tests above gave different values for a given home. Duct-leakage-to-outside measurements obtained by the APT, Manual and Combined methods were statistically analyzed to determine the error incurred by estimating one measurement from the other in the following cases:

- APT from Manual
- Combined from APT
- Combined from Manual

Agreements between the means of these pairings of the three datasets were investigated. Both parametric and non-parametric tests were used in this analysis. The importance of this comparison is to bring forth the issue whether or not we are measuring the same quantity – in this case duct-leakage-to-outside. If differences exist between the three methods, it becomes important to obtain the duct-leakage-to-outside by all three methods, since the research community may not be sure of which test method is the most accurate. Energy auditors perform only one of the above mentioned test methods; the test method is mainly based on requirements by the state or agency.

On the other hand, a predictive model will enable us to estimate the measurements of a method given another. This model will enable us to create a database of results (some estimated) from all three tests for comparison purposes at US level.

Statistical Comparisons: Depressurized APT vs. Depressurized Manual

Depressurization measurements in cubic feet per minute (CFM) were obtained at a pressure of 50 Pascal's (Pa) using the APT and Manual methods. Forty-two (42) homes were successfully tested using both the depressurized APT and Manual tests. These two sets of measurements were then statistically analyzed using Line of Equality and Regression analyses. Moreover, the means of the Manual and APT data sets were statistically compared using the Paired t as well as a non-parametric test.

Line of Equality

For both the Line of Equality and the Regression, the analysis starts with data plotted in the same manner as in Figure 1. For each home, a point is plotted on a graph where the abscissa (viewed as the independent variable, i.e., horizontally) is the Manual datum and the ordinate (viewed as the dependent variable, i.e., vertically) is the corresponding APT datum.

The Line of Equality is a 45-degree line drawn through the origin as presented in Figure 1. A point lying exactly on the line indicates that the values corresponding to the two axes are in agreement. The further the point lies from the Line of Equality, the greater the difference is between the set of values. Figure 1 displays the Line of Equality graph that associates APT vs. Manual data within the 42-home sample; by definition, the slope of the Line of Equality is always one. Figure 1 indicates that the APT and Manual methods are in reasonable agreement when the duct-leakage-to-outside measurements are below 1000 CFM but that outliers are more likely for higher CFM values. The causes for these abnormal variations may be attributed to errant conditions, such as: unusual wind speed, untimely changes in other building environmental parameters, human errors, as well as test conditions that may have resulted in the inconsistent opening/closing of building airflow valves (Katz et al., 2004).

Regression Analysis

The goal of a regression analysis is to present a (hopefully) "close" and linear approximation of the relationship between two variables (Montgomery et al., 2001). A Regression analysis finds the line of best fit and statistically evaluates the quality of that approximation. The line of best fit is a line which "best" approximates the relationship between two variables.

In this study Regression analysis was used to investigate and model the relationship between:

1. APT and Manual
2. Combined and APT
3. Combined and Manual

It is important to note that the Combined procedure is the most widely used test method for determining duct-leakage-to-outside, followed by APT and Manual. It is interesting to get an estimate of Combined results from both APT and Manual tests. In addition, estimates of APT from Manual tests will also be beneficial in order to develop a database on all these values.

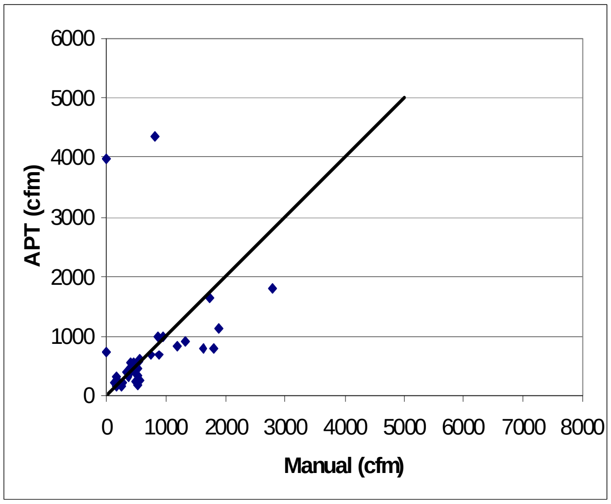

Figure 1. Line of Equality: Duct-Leakage-To-Outside APT vs. Manual

Determining Outliers and Influential Points

Loosely speaking, outliers are defined as observations that are far from the rest of the sample, or observations that have large residuals. A residual is defined as the difference between the actual measured value and the predicted value. Since the magnitude of the residuals depends upon the unit of measurement of the observed variables (i.e., for example, the magnitude of the residuals would change size if the units of duct leakage were "cubic feet per second" instead of "cubic feet per minute"), one cannot identify outliers by defining an upper bound on the magnitude of acceptable residuals. For this reason, a standardized residual is defined to be a residual divided by its estimated standard error. Unlike residuals, standardized residuals have magnitudes that are independent of the units of measurement. In particular, standardized residuals provide a "statistical" metric for judging the size of a residual. A standardized residual with magnitude greater than 2.75 is large and represents an outlier.

A plausible rationale for this approach is that any datum that has a standardized residual with magnitude in excess of 2.75 must be associated with a residual that is more than 2.75 standard deviations from the mean of the residuals. Therefore, if one assumes that the data's errors are normally distributed, that datum is statistically more unusual than 98% of the other data and should be recognized as an outlier. Procedurally, such a datum is removed from the data and the Regression analysis repeated on the remaining data.

Influential points on the other hand are data points with extreme values that greatly affect the slope of the regression line. It is important to note that influential points do not necessarily reduce the coefficient of determination. In this paper, outliers as well as influential points were carefully investigated before exclusion from the data. Data points that were invalid were deleted whereas those that were unusual where scrutinized to justify their validity/invalidity.

Sources of Error in Duct Leakage Measurements

This paper concerns itself with duct-leakage-to-outside as opposed to total duct leakage (leakage to inside the conditioned space as well as to outside), because duct-leakage-to-outside measures a quantity which is more relevant to energy loss in buildings. However with increasing total duct leakage there will be greater degradation of the integrity of the duct system, and thus a larger degradation of the error in the measurement should be expected. Various sources of error in measurements contribute to outliers and influential points. Possible sources of errors in determining duct leakages using these three methods include:

- Valves, which may open during a pressurized test and close or not open in a depressurized test and vice-versa. In such cases, we should expect disparity between depressurized and pressurized measurements.
- Modeling errors, which are implicit in the theoretical design of the test. For example, recently published work (ASTM, 1994) indicates that Modified Subtraction will give less accurate results when the attic pressure is significantly different from the outside pressure; this situation is not uncommon. Combined tests definitions have changed in response to this problem as well (Minneapolis Duct Blaster Operation Manual, 2000, 2003).
- Internal pressure differences in the ducts, which may affect these tests in unexpected ways. For example, combined tests rely upon the assumption that there is negligible pressure drop across an evaporator coil, which may be questionable. However, Modified Subtraction is less likely to be affected by these issues.
- Wind, attic fans and other sources of pressure fields external to ducts will also affect these tests inconsistently. Modified Subtraction is generally

more sensitive than the Combined test to such problems.

Outliers and influential points may cause variations between the test methods. Therefore, the detection of outliers and influential points are important in this study, as these issues need to be addressed in future investigations.

Regression Analysis with 95% Prediction Intervals

Regression analysis was used to investigate and model the relationship between APT and Manual measurements. Figure 2 demonstrates a Regression analysis for APT and Manual data (with outliers/influential points removed) graphically. The resulting regression provides 95% confidence and prediction intervals in addition to the regression line. A Regression analysis provides a prediction.

The results rely on assumptions such as the independence and normality of random errors. Confirmation of these assumptions was ascertained via Figure A1 for the independence of errors and Figure A2 for normality, in Appendix A.

Note that the observation order of the data is the home numbers (experiment number) after outliers are removed. The residual plot resembles a random scatter – thereby supporting the independence of errors.

The closeness of the points to the line plotted in Figure A2 indicates the normality of the residuals. A Normal Probability Plot graphically displays standardized residuals, but uses a varying scale for the ordinate. Like a logarithmic scale which causes an exponential function to be graphed as a straight line, the scale chosen for use in a normal probability plot causes a normally distributed set of standardized residuals to lie on a straight line. That is, a normal probability plot of experimentally obtained standardized residuals will display a "closeness" of the distribution to a normal distribution by noting the closeness to the line of the graphed data (Weisberg, 1985, Wonnacott et al., 1977). For a more definitive and numerical approach than the visual method provided by a Normal Probability Plot, statisticians often supplement the analysis with the Anderson-Darling test. An Anderson Darling p-value of the residuals of 0.061 confirms normality at level 5% (Minitab, 1998).

We note that three outliers were removed from the data, leaving a 39-home sample. Home numbers 46, 48 and 1 had invalid observations. For example home numbers 1

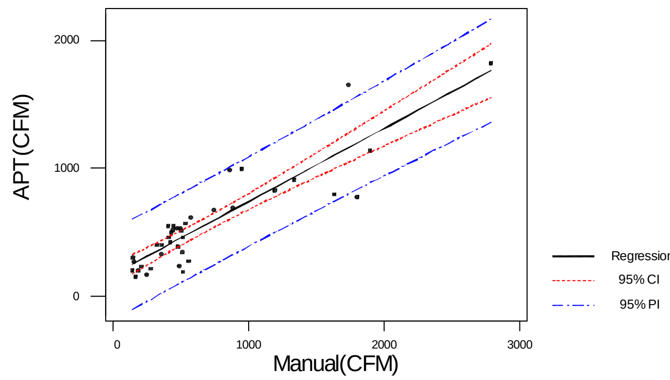

Figure 2. Regression Plot with 95% Confidence and Prediction Intervals: APT vs. Manual

and 48 have duct leakages 0 and 1.4 CFM respectively for the Manual methods, which is physically incorrect. Home number 46 showed inconsistent values in regards to flow exponent values. It is important to note that these observations had residuals greater than 2.75 and were also influential. Other observations, which were influential, were not considered invalid thus were not deleted.

The regression equation is

$$APT = 168 + 0.57 \text{ Manual}. \qquad (1)$$

Namely, the value of the APT reading can be predicted from the Manual reading using Equation 1. From equation 1, we see that the regression coefficient is positive (=0.57) which means that the higher the value of Manual is, the greater the corresponding value of APT is. The slope of the line is significant. For a given Manual value, the red dashed lines in Figure 2 display the interval that includes the true APT mean value with probability 0.95. Also displayed in the output (we used the statistical package MINITAB) is the coefficient of determination, r^2, with a value of 0.79; thus indicating a high correlation between the two measuring methods (the correlation coefficient r = 0.89, the square root of the coefficient of determination r^2). A Confidence Interval (CI) is useful as an interval estimate for a population parameter (for example the mean APT for a given Manual value) but it does not provide any information about the range of possible individual APT values for a given Manual value. On the other hand, the Prediction Interval (PI) is useful in determining the range of a future APT individual value based on a current Manual value; the PI is displayed with blue dashed lines on Figure 2.
Table 1 shows the predicted APT values for each Manual reading and their 95% prediction intervals. Note that the prediction interval is bounded by the values in columns 5

and 6 and centered about the predicted APT value given in column 4. From Table 1, we see that the predicted value for the response (APT) varies by an average of 20% of the actual response. However, the prediction intervals are wide because the estimates of prediction errors used in determining prediction intervals take into account both the uncertainty in the estimation of the regression line, and the fact that points do not exactly line on the line due to random errors. Note that the prediction error is larger when the predictor data are far from their calibration-period means (Weisberg, 1985).

Table 1. 95% Prediction Intervals for Duct-Leakage-To-Outside values: APT vs. Manual

Observation/ (Home Number)	APT (CFM)	Manual (CFM)	Predicted (CFM) APT	95% Prediction Interval	
1 (2)	987.2	861	732.09	382.64	1081.54
2 (3)	907	1337.9	686.27	337.37	1035.16
3 (4)	1818.4	2789.2	1207	842.29	1571.71
4 (5)	340.3	508.4	362.48	12.73	712.23
5 (6)	170.1	245.5	265.23	0	616.86
6 (8)	795	1630.5	622.27	273.87	970.68
7 (11)	421.8	418.7	409.04	59.93	758.16
8 (12)	149.7	166.7	253.58	0	605.48
9 (13)	530.8	474.1	471.32	122.79	819.85
10 (14)	1133.8	1895.6	815.85	464.95	1166.74
11 (16)	232.5	209.6	300.89	0	651.74
12 (18)	674	742.5	553.14	204.91	901.37
13 (19)	775.7	1800.8	611.25	262.9	959.6
14 (20)	492.4	427.6	449.38	100.68	798.08
15 (21)	301.5	148	340.31	0	690.42
16 (26)	326.8	354.5	354.77	4.9	704.64
17 (27)	270.6	559.4	322.66	0	673.08
18 (28)	266.3	153.8	320.2	0	670.67
19 (29)	400.5	361.1	396.87	47.61	746.14
20 (30)	385.38	477.3	388.24	38.86	737.61
21(31)	1647.9	1732	1109.58	749.37	1469.8
22(33)	615.3	569.6	519.6	171.31	867.89
23 (34)	510.1	502.3	459.5	110.88	808.11
24 (35)	190.9	515.2	277.12	0	628.48
25 (36)	549.8	408.4	482.18	133.72	830.64
26 (37)	517.4	438.3	463.67	115.08	812.25
27 (38)	546.6	445.3	480.35	131.88	828.82
28 (39)	401.1	327.2	397.22	47.96	746.47
29 (40)	198.2	182.8	281.29	0	632.55
30 (42)	216.1	278.3	291.52	0	642.56
31 (43)	457.4	517	429.38	80.5	778.27
32(44)	994.5	950.2	736.26	386.75	1085.77
33 (45)	566.2	535	491.55	143.14	839.95
34 (47)	459.4	409	430.53	81.65	779.4
35 (51)	828.3	1192	641.3	292.78	989.82
36 (52)	236.6	486	303.23	0	654.03
37 (53)	530.5	496.5	471.15	122.62	819.68
38(54)	690.4	883.3	562.51	214.28	910.74
39 (55)	199.1	140.9	281.8	0	633.06

Statistical Comparison of Means: APT Mean compared to Manual Mean

To confirm that the two methods tend to measure the same mean duct leakage, a comparison of the means was performed using the Paired t-test. The paired t-test was performed on the 42-home sample using the MINITAB software.

Paired t-test

The paired t-test was used to determine whether the mean results of the APT and Manual methods are not significantly different. We note that the applicability of the paired t-test depends upon the data following a normal distribution or a large enough sample size (30 usually suffices) and on the data consisting of pairs of observations as is the case here: APT and Manual measurements for each home. The paired t-test procedure involves calculating a "difference score" for each subject home; i.e., for a given home a differences core, D, is taken to be “APT measurement – Manual measurement”. The differences D are then treated as a single sample. The advantage of using the differences is that the matched samples give a much more precise (narrow) 95% interval. The average sample difference $\overline{D}$ was calculated. The 95% confidence interval was then calculated via

$$\Delta = \overline{D} \pm t_{.025}\frac{s_D}{\sqrt{n}}$$

where Δ is the 95% confidence interval, s_D is the standard deviation, n is the number of data points, and $t_{.025}$ is the critical point, obtained from a table. (We note that $t_{.025}$ decreases to 1.96 as n approaches ∞).

The 95% confidence interval, Δ, for the mean difference between APT and Manual was determined to be between –197 CFM and 360 CFM. Since this range includes zero, the test indicates that there is no statistically significant difference in the average measured readings of the APT and the Manual methods.

Non-Parametric test

A non-parametric test was also performed because the assumption of normality was violated in the paired t-test. Normality tests are performed to check whether the data follows a normal distribution or not. In this case, the set of paired differences (APT – Manual) failed the normality tests. Non-Parametric tests or Distribution-free methods are methods used when the assumption of normality is seriously violated. The Wilcoxon paired-sample test is the non-parametric analog to the paired t-test, and is very efficient when there are serious departures from normality (Walpole et al., 1998). We note that the paired t-test is still usable in our case since the sample size is large enough, but it is interesting to see whether its results agree with the non-parametric test.

The Wilcoxon paired-sample test was used to compare the medians of the APT and Manual measurements. This test utilizes both direction and magnitude of the differences between the two measured values of duct leakage. The p-value of the Wilcoxon statistic for the difference between the medians of the APT and Manual methods is computed to be 0.532. Since this value of 0.532 is greater than the level of significance of 0.05, we can conclude that the means of the APT and Manual measurements of duct-leakage-to-outside are not significantly different.

The conclusion reached by the parametric test (paired t-test) test was similar to that reached by the Non-parametric test (Wilcoxon paired-sample test), in part because of our sufficient sample size.

Statistical Comparisons: Pressurized Combined vs. Depressurized APT

Measurements obtained at 50 Pa in the depressurized state for APT and the pressurized Duct Blaster assisted by Blower Door methods produced two experimental data sets. These sets are now statistically compared using the Line of Equality and Regression analyses, as well as a paired t-test and a non-parametric test, as done in the previous analysis.

Line of Equality

Although the sample originally contained 55 homes, only 34 homes had sufficient data to compare the depressurized APT and pressurized Combined methods. Figure 3 illustrates the Line of Equality for Combined vs. APT for the resulting 34-home sample.

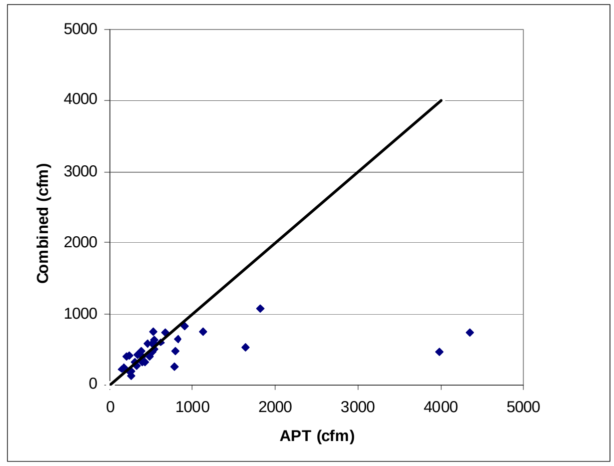

Figure 3. Line of Equality: Combined vs. APT

Outliers and influential points were checked while performing the Regression analysis. Assumptions of independence of errors and normality were confirmed via Figure A3 and Figure A4 respectively, in Appendix A. In addition, normality was confirmed using the Anderson-Darling test. All the standardized residuals were within +/-2.75. However, there were influential points: home numbers 46 and 48 were excluded from the analysis since they had inconsistent values in regards to flow exponent values. Standardized Residuals vs. Observation Order are graphed in Figure A3. The scatter indicates that the errors are independent. The data along the straight line in Figure A4 yields a p-value of 0.521 for the Anderson-Darling test, thereby indicating that the residuals follow a normal distribution.

Regression Analysis with 95% Prediction Intervals

In this case, the response variable is taken to be Combined and the predictor is taken to be APT.

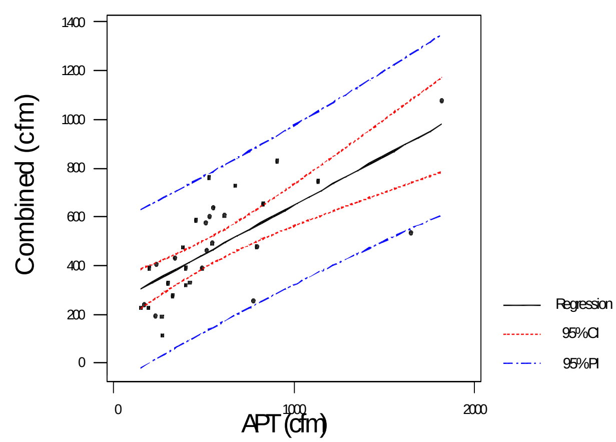

Figure 4. Regression Plot with 95% Confidence and Prediction Intervals: Combined vs. APT

The Regression analysis was performed on the 34-home sample (Figure 4). This figure illustrates the 95% confidence and prediction intervals. The regression equation was determined to be:

$$\text{Combined} = 244.53 + 0.40\ \text{APT} \qquad (2)$$

Using Equation 2 and the prediction interval, the Combined reading is predicted from the APT reading with a probability of 0.95. From equation 2, we note that the as the value of APT increases the value of Combined also increases. The slope of the line is significant. The r-square value of 0.52 indicates a moderately high correlation (r = 0.72) between the two measuring methods.

Table 2 shows the predicted Duct Blaster assisted by blower door values for each APT reading and their 95% prediction intervals. From Table 2, we see that predicted fit for the response (combined) varies by an average of 23% of the actual response.

Statistical Comparison of Means: Combined vs. APT

To confirm that the two methods tend to measure the same mean value of duct-leakage-to-outside, a comparison of the means was performed using the paired t-test on the 34-home sample.

Paired t-test

The 95% confidence interval, Δ, for the mean difference between Combined and APT was determined to lie between –16 CFM and 594 CFM; since this range includes zero, the test indicates that there is no statistically significant difference in the average measured readings of the APT and Combined methods. However, non-parametric tests give more reliable results when normality conditions are not satisfied, although our sample size is likely to still suffice.

Non-Parametric test

The Wilcoxon paired-sample test mentioned previously was performed on the data. The p-value of the Wilcoxon statistic for the difference between the medians of the APT and Combined methods is computed to be 0.161. Since this value of 0.161 is greater than the level of significance of 0.05, we can conclude that the APT and Combined methods are not significantly different in their measurements of mean duct-leakage-to-outside.

Table 2. 95% Prediction Intervals for Duct-Leakage-To-Outside: Combined vs. APT

Observation /(Home Number)	APT (CFM)	Combined (CFM)	Predicted (CFM) Combined	95% Prediction Interval (CFM)	
1 (3)	907	827.51	578.765	255.809	901.72
2 (4)	1818.4	1077.7	679.819	350.284	1009.35
3 (5)	340.3	430.57	418.44	97.544	739.34
4 (6)	170.1	239.94	341.443	17.827	665.06
5 (8)	795	477.23	437.287	116.692	757.88
6 (11)	421.8	329.3	377.534	55.49	699.58
7 (12)	149.7	226.34	335.949	12.048	659.85
8 (13)	530.8	602.19	487.756	167.253	808.26
9 (14)	1133.8	747.74	546.546	224.841	868.25
10 (16)	232.5	195.72	323.58	0	648.17
11 (18)	674	729.22	539.063	217.589	860.54
12 (19)	775.7	254.13	347.175	23.843	670.51
13 (20)	492.4	387.05	400.86	79.554	722.17
14 (21)	301.5	327.51	376.812	54.741	698.88
15 (26)	326.8	273.64	355.054	32.091	678.02
16 (27)	270.6	112.83	290.103	0	616.84
17 (28)	266.3	191.86	322.024	0	646.7
18 (29)	400.5	320.04	373.794	51.611	695.98
19 (30)	385.38	475.92	436.754	116.153	757.36
20 (31)	1647.9	533.57	460.039	139.615	780.46
21 (33)	615.3	605.8	489.216	168.701	809.73
22 (34)	510.1	574.98	476.767	156.333	797.2
23 (35)	190.9	226.93	336.188	12.299	660.08
24 (36)	549.8	638.49	502.419	181.748	823.09
25 (37)	517.4	463.92	431.908	111.242	752.57
26 (38)	546.6	490.46	442.628	122.093	763.16
27 (39)	401.1	390.9	402.417	81.153	723.68
28 (43)	457.4	586	481.218	160.762	801.67
29 (51)	828.3	654.46	508.87	188.097	829.64
30 (52)	236.6	404.51	407.913	86.787	729.04
31 (53)	530.5	758.73	550.984	229.132	872.84
32 (55)	199.1	388.02	401.252	79.957	722.55

Statistical Comparisons: Combined vs. Depressurized Manual

Measurements obtained at 50 Pa in the depressurized state for the Manual method and measurements obtained from the pressurized Combined method produced two data sets. These data were analyzed with the same techniques as described in previous sections.

Line of Equality

From the 55-home sample, only 35 could be used for comparing the Manual and Combined methods. While running the Regression analysis outliers in the data were checked. Using the standardized residuals > 2.75 test, two observations were deleted (homes 46 and 48) – leaving a 33-home sample for the final regression model.

The assumptions of independence (Figure A5, Appendix A) and normality (Figure A6, Appendix A) of errors were checked and found to hold satisfactorily, and the normality test was performed on the residuals using the Anderson-Darling test (yielding a p-value of .210).

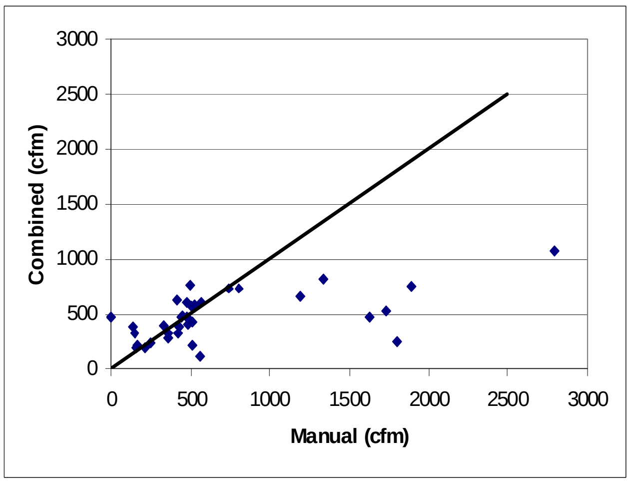

Figure 5. Line of Equality: Combined vs. Manual

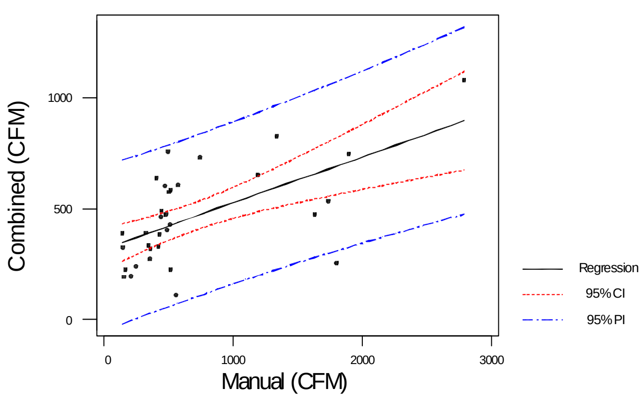

Figure 6. Regression Plot with 95% Confidence and Prediction Intervals: Combined vs. Manual

Regression Analysis with 95% Prediction Intervals

In this analysis the response variable was taken to be the measurements from the Combined method and the predictor variable consisted of measurements of duct leakage by the Manual method. The Regression analysis, performed on the 35-home sample, is displayed in Figure 6. This figure illustrates the 95% confidence and prediction intervals. The regression equation is:

$$\text{Combined} = 319.90 + 0.21\ \text{Manual} \qquad (3)$$

The slope of the line (equation 3) is significant. Using Equation (3), the value of the Combined reading can be predicted from the Manual reading as discussed previously. The r-squared value of 0.358 indicates a moderate correlation (r =0.60) between the two measuring methods.

Table 3 shows the predicted Combined values for each Manual reading and their 95% prediction intervals. An individual Combined value, given a particular Manual value, falls within the prediction interval with a probability of 0.95.

Table 3 reveals that predicted values for the response (combined) vary by an average of 32% of the actual response.

Statistical Comparison of Means: Combined vs. Manual

To confirm that the two methods tend to measure the same mean duct leakage, a comparison of the means was performed using the paired t-test on the 35-home sample.

Paired t-test

The 95% confidence interval, Δ, for the mean difference between the Combined and Manual measurements was determined to be between 24 CFM and 384 CFM; since this range does not include zero, the test indicates that there is a statistically significant difference in the average measured readings of APT and Combined method. However, it is necessary to test the results with a distribution free method since the conditions of normality are not satisfied, although the sample size is probably high enough.

Non-Parametric test

The Wilcoxon paired-sample test was then performed. The p-value of Wilcoxon statistic for the difference between the medians of the APT and Manual methods is computed to be 0.171. Since this value of 0.171 is greater than the level of significance of 0.05, we cannot reject the hypothesis that the Combined and Manual methods measure the same mean amount of duct-leakage-to-outside. It is possible that the paired t-test was sensitive to outliers in the differences, while the Wilcoxon test, based on medians, is more robust to these outliers.

Comparison of Mean Duct-Leakage-to-Outside on Full and Reduced Data (after removal of outliers)

The Mean Duct Leakage prior to statistical analyses is presented in Table 4; its rows and columns summarize the source duct leakage measurements. Mean values are presented in the first six rows of Table 4.

Table 3. 95% Prediction Intervals for Duct-Leakage-To-Outside values: Combined vs. Manual

Observation /(Home Number)	Manual (CFM)	Combined (CFM)	Predicted (CFM) Combined	95% Prediction Interval (CFM)	
1 (3)	1337.9	827.51	491.201	126.838	855.564
2 (4)	2789.2	1077.7	542.991	176.782	909.2
3 (5)	508.4	430.57	409.033	43.97	774.097
4 (6)	245.5	239.94	369.572	2.595	736.55
5 (8)	1630.5	477.23	418.693	53.942	783.443
6 (11)	418.7	329.3	388.069	22.115	754.023
7 (12)	166.7	226.34	366.757	0	733.909
8 (13)	474.1	602.19	444.559	80.345	808.772
9 (14)	1895.6	747.74	474.688	110.542	838.834
10 (16)	209.6	195.72	360.418	0	727.984
11 (18)	742.5	729.22	470.854	106.732	834.975
12 (19)	1800.8	254.13	372.51	5.71	739.31
13 (20)	427.6	387.05	400.024	34.613	765.435
14 (21)	148	327.51	387.699	21.726	753.671
15 (26)	354.5	273.64	376.548	9.982	743.113
16 (27)	559.4	112.83	343.26	0	712.074
17 (28)	153.8	191.86	359.62	0	727.24
18 (29)	361.1	320.04	386.152	20.102	752.202
19 (30)	477.3	475.92	418.42	53.661	783.178
20 (31)	1732	533.57	430.353	65.9	794.807
21 (32)	345.1	336.82	389.626	23.747	755.504
22 (33)	569.6	605.8	445.307	81.102	809.512
23 (34)	502.3	574.98	438.927	74.634	803.22
24 (35)	515.2	226.93	366.879	0	734.024
25 (36)	408.4	638.49	452.073	87.933	816.214
26 (37)	438.3	463.92	415.936	51.102	780.769
27 (38)	445.3	490.46	421.43	56.758	786.103
28 (39)	327.2	390.9	400.822	35.444	766.2
29 (43)	517	586	441.208	76.949	805.466
32 (51)	1192	654.46	455.379	91.259	819.5
33 (52)	486	404.51	403.639	38.373	768.904
34 (53)	496.5	758.73	476.963	112.798	841.128
35 (55)	140.9	388.02	400.225	34.822	765.628

Table 4. Mean Duct-Leakage-to-Outside

Tests	Number of Tests	Mean Duct-Leakage-to-Outside (CFM)	Standard Deviation (CFM)
1. APT Depressurized (50 Pa)	42	733	859
2. Manual Depressurized (50 Pa)	43	645	578
3. APT Pressurized (50 Pa)	16	459	261
4. Manual Pressurized (50 Pa)	16	654	334
5. Combined (25 Pa)	43	313	154
6. Combined (50 Pa)	43	474	233
7. Weighted Mean of All tests at 50 Pa	160	604	503
8. Weighted Mean of All tests at 25 Pa	160	399	332

The 7th row of Table 4 presents a weighted mean of duct leakage measurements at 50 Pa (which includes rows 1- 4 and 6). The 8th row presents a weighted mean of the duct leakage measurements at 25 Pa; it was obtained from the 7th row by multiplying by $(25/50)^{0.60}$ (Witriol et al., 2003).

Table 5 displays the revised mean values of the duct leakage measurements calculated after the removal of outliers from each of the original data sets. The 7th row is a weighted mean of rows 1-4 together with the 6th row. The 8th row is the conversion of the weighted mean in row 7 to 25 Pa by multiplying by $(25/50)^{0.60}$. The 9th row is a conversion of the 6th row to 25 Pa. Comparing rows 7 and 8 in tables 4 and 5, it is seen that after outliers and influential points were removed the weighted mean of all tests decreased by about 14%.

6. Conclusions

Neither parametric (paired-t) nor non-parametric (Wilcoxon paired-sample) tests performed on APT (depressurized) vs. Manual (depressurized) and Combined vs. APT (depressurized) measurements showed any statistically significant difference in mean between the two measurement methods compared. However, the paired-t-test showed a statistically significant difference in mean measurements between the Combined and Manual (depressurized) methods. On the other hand, it is important to note that the assumptions of normality were violated in the paired-t-tests between Combined and Manual (depressurized) methods; at the same time sample sizes were satisfactory. However, a non-parametric test did not show any significant difference between these two methods.

Table 5. **Mean Duct-Leakage-to-Outside Post-Regression**

Test	Number of Tests	Mean Duct Leakage to Outside (CFM)	Standard Deviation (CFM)
1. APT Depressurized (50 Pa)	39	557	374
2. Manual Depressurized (50 Pa)	39	681	585
3. APT Pressurized (50 Pa)	16	459	261
4. Manual Pressurized (50 Pa)	16	654	334
5. Combined (50 Pa) (vs. APT)	34	467	218
Combined (50 Pa) (vs. Manual)	32	463	216
6. Combined (50 Pa) Weighted Mean	33	465	217
7. All tests (50 Pa) Weighted Mean	143	530	349
8. All tests (25 Pa) Weighted Mean	143	350	290
9. Combined (25 Pa)	33	307	379

The regression equations for the respective predictive models were as follows:

APT = 168 + 0.57 Manual
Combined = 244.53 + 0.40 APT
Combined = 319.90 + 0.21 Manual

All of the respective models have moderate to high correlation with all the assumptions of linear regression satisfied.

The mean duct-leakage-to-outside, based on the weighted mean of all tests is equal to 350 CFM. Therefore, if we assume that the average air conditioner size of the various samples was near 3 tons (corresponding to 1200 CFM), then the mean percent duct leakage for the State of Louisiana would be 350/1200 = 29%.

Acknowledgements

This work was funded by the Louisiana Department of Natural Resources Interagency Agreement No. PVE29-01-12 and the U.S. Department of Energy. During the course of this research project, we received valuable support and guidance from Wade Byrd, and Buddy Justice. We would also like to acknowledge the support and assistance of other members of the DNR Energy Section staff; in particular Paula Ridgeway, Harvey Landry and Tangular Williams.

Correspondence: jinson_je@yahoo.com

REFERENCES

Automated Performance Testing Manual, (2000), The Energy Conservatory, Minneapolis, MN, www.EnergyConservatory.com, 29-40.

Minneapolis Blower Door Operation Manual, (1993), The Energy Conservatory, Minneapolis, MN, www.EnergyConservatory.com, 49-50.

ASTM E779-99, (1999), "Standard Test Method for Determining Air Leakage Rate by Fan Pressurization," *American Society for Testing and Materials*, West Conshohocken, PA.

ASTM E1554-94, (1994), "Standard Test Methods for Determining External Air Leakage of Air Distribution Systems by Fan Pressurization," *American Society for Testing and Materials*, West Conshohocken, PA.

Witriol, N., J. Erinjeri, M. Katz, R. Mc Kim, and R. Nassar, (2003), *Testing HVAC Duct Leakage in Existing Residential Buildings in North Louisiana*, Trenchless Technology Center, Louisiana Tech University, Ruston, LA.

Katz, M., N. Witriol, and J. Erinjeri (2004), "The Generalized Subtraction Correction Algorithm for Measuring Duct Leakage Using a Blower Door," *Journal of Testing and Evaluation*, 32 (6), West Conshohocken, PA.

Montgomery, D.C., E. A. Peck, and G. G. Vining, (2001), *Introduction to Linear Regression Analysis*, 3rd edition, New York: John Wiley & Sons.

Minneapolis Duct Blaster Operation Manual, (2000), The Energy Conservatory, Minneapolis, 26-28.

Minneapolis Duct Blaster Operation Manual, (2003), The Energy Conservatory, Minneapolis, www.EnergyConservatory.com, 28-30.

Mendenhall, W. and T. Sincich, T., (1988), *Statistics for the Engineering and Computer Sciences*, 2nd edition, California: Dellen Publishing Company.

Walpole, R.E., R. H. Myers, and S. L. Myers, (1998), *Probability and Statistics for Engineers and Scientists*, 6th edition, New Jersey: Prentice Hall Inc.

MINITAB User's Guide for Data Analysis and Quality Tools, (1998), Minitab Inc.

Weisberg, S., 1985, *Applied Linear Regression*, 2nd ed., John Wiley, New York, 324

Wonnacott, T. and R. Wonnacott, R., (1977), *Introductory Statistics*, 3rd edition, New York: John Wiley & Sons.

Appendix A

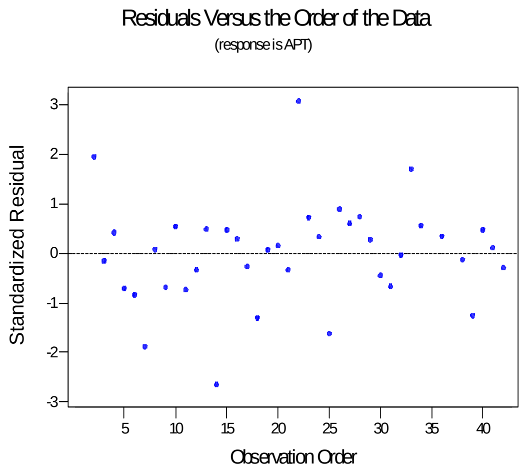

Figure A1. Standardized Residuals' Plot: APT vs. Manual

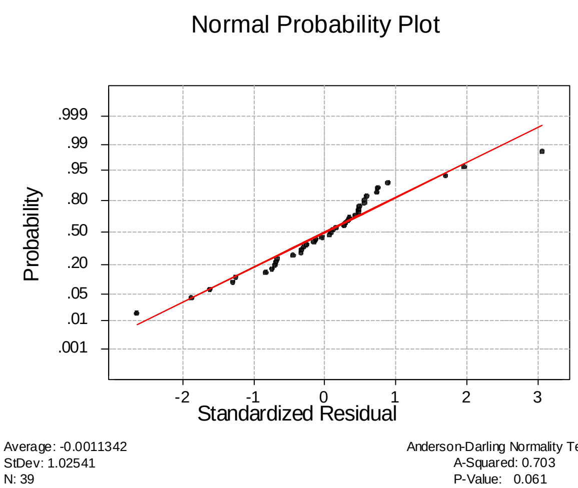

Figure A2. Normal Probability Plot of Standardized Residuals: APT vs. Manual

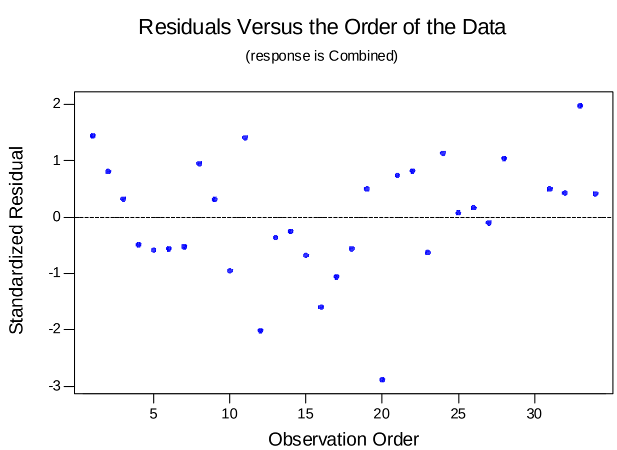

Figure A3. Standardized Residuals' Plot: Combined vs. APT

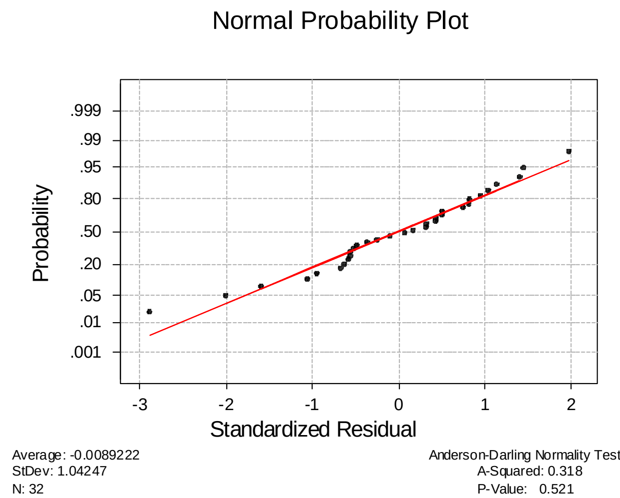

Figure A4. Normal Probability Plot of Standardized Residuals: Combined vs. APT

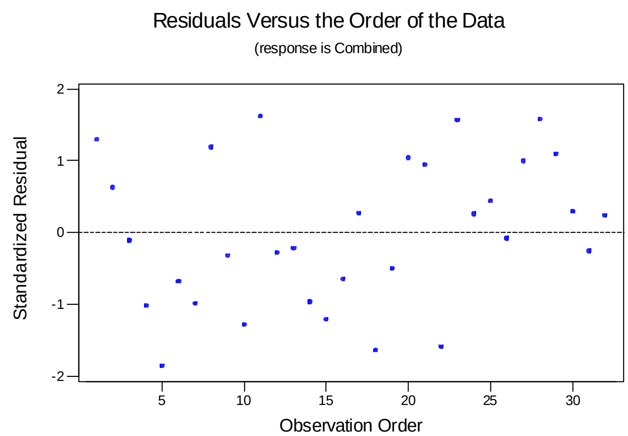

Figure A5. Standardized Residuals' Plot: Combined vs. Manual

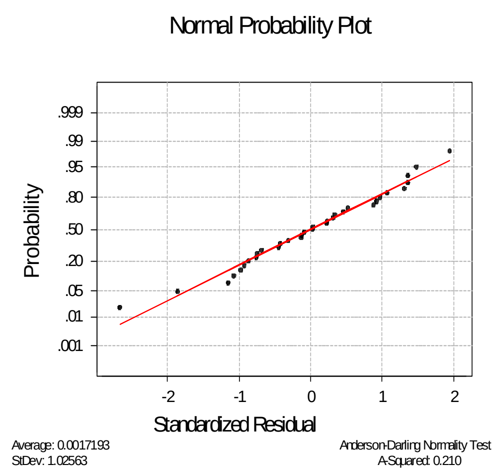

Figure A6. Normal Probability Plot of Standardized Residuals: Combined vs. Manual