Understanding and Reducing Uncertainties in the Delta-Q Test during Field Testing

Paul W. Francisco

Building Research Council
University of Illinois at Urbana-Champaign

March 31, 2009
Executive Summary

Duct leakage to outside has long been recognized as a major source of energy loss. However, many homes have ducts that do not leak appreciably, so efforts to diagnose the leakage from ducts in homes have become common so as to focus repair efforts on those homes for which they are warranted. The two most common test methods at the current time are the duct pressurization test and the pressure pan test.

The Delta-Q test for duct leakage was initially developed in 2001 (Walker et al. 2001). This test estimates duct leakage to outside, which is the leakage that results in substantial energy losses. It has several advantages over other residential duct leakage techniques. It requires only a blower door and a laptop computer; the supply and return registers and grilles do not need to be accessed; and the estimated leakage is based on the operating conditions of the forced-air distribution system.

The test also has undergone extensive laboratory and field evaluation, resulting in changes to both the sampling and analysis methodologies (Dickerhoff et al. 2004; Walker et al. 2004; Walker and Dickerhoff 2008). These efforts led to a 2006 study that evaluated various aspects of the “ramping” data collection technique for use in weatherization homes (Francisco 2006). The results of that study suggested significant promise for the Delta-Q test in this application, but recommended further evaluation of strategies to reduce the noise due to wind and if and how the test could be done successfully using only a single blower door ring instead of two. Working towards a quantitative uncertainty estimate was also cited as a primary need.

Since then further work has been done in the field to help characterize the uncertainty of the test method and how well it performs for other applications. For example, the Delta-Q test was performed in 19 new construction homes in Wisconsin, most of which were pursuing energy efficiency credits and which had most (but not all) of their ducts within the conditioned space (Pigg and Francisco 2008). This study compared the Delta-Q test results to other methods, including an advanced “nulling” method suitable only in research contexts that provided a comparison method that also evaluated duct leakage at operating conditions. This study also identified wind as a confounding factor, and took an initial attempt at developing a quantitative uncertainty estimate. Because of the nature of the homes in this study most had low leakage of both the homes and the ducts, making the results of limited application to weatherization homes.

Another study looked at repeatability of the Delta-Q test in several homes in California and Nevada (Dickerhoff and Walker 2008). Attempts to develop quantitative estimates of uncertainty for an individual house were of marginal success, though an estimate of 1% of house leakage rate at 50 Pascals (Pa) depressurization as an average value was proposed. The results of this study were consistent with those of Francisco (2006) with regard to the width of the pressure “bins” into which the data would be collected for analysis (5 Pa bins having been proposed by Francisco).
A third study evaluated repeated testing from several users around the U.S. to develop a means of estimating uncertainty by comparing the standard deviations among the repeated tests to the individual test standard errors of the leakage estimates (Olson 2008). Two methods of arriving at a multiplier to be used in conjunction with the standard errors of the leakage estimates were considered, with multipliers between 2 and 3 being most promising.

This project is a follow-up study to the 2006 study by Francisco, and also makes use of the work done by Dickerhoff and Walker (2008), Pigg and Francisco (2008), and Olson (2008). The primary objectives were to reduce uncertainty due to wind, evaluate single-ring testing options, and further the development of quantitative uncertainty estimates. The advanced “nulling” technique was conducted in as many homes as possible for comparison. Homes were selected such that most homes had the majority of the duct work outside the conditioned space. Comparisons to the common techniques of duct pressurization and pressure pan tests were also conducted so as to assess the extent to which the tests gave different signals with regard to whether a duct system warranted air sealing.

A number of valuable conclusions have resulted from this project.

- The ramping sampling protocol is preferable to the stations sampling protocol.
- There is no substantial benefit to placing the outdoor reference tap in a location other than the blower door side of the house, and in some cases is detrimental. Therefore the blower door side of the house is the recommended location for this pressure tap.
- The one-ring version of the Delta-Q test, with the ring chosen to be the smallest that can reach at least 30 Pa house depressurization, produces results comparable to the two-ring test. Under calm conditions the test should produce reasonable results using a single ramp, but a second ramp is recommended when there is noticeable wind.
- The one-ring test with one ramp has approximately 40% greater uncertainty than the two-ring test. It may be that this increased uncertainty can be removed with a dual-ramp test, and this possibility warrants further investigation.
- To provide uncertainty estimates for the Delta-Q leakage estimates, the results from these homes suggest that the standard error from the software be multiplied by about 2.2 for the two-ring tests or about 3.0 for the single-ring, single-ramp tests.
- The Delta-Q test performs better than both the duct pressurization and pressure pan tests at indicating the need to repair the ducts. This is especially true for the pressure pan tests using typical interpretations.
• The median value of the pressure pan tests had much better correlation with the Delta-Q and nulling leakage estimates, suggesting that if the pressure pan is to be used that the median value would be a preferable metric to use when determining whether to perform air sealing.

• The duct pressurization test is subject to errors in leakage estimates due to non-uniformities in the pressures within the ducts, which increases the uncertainty about the reliability of the estimates.
# Table of Contents

Executive Summary ............................................................................................................ 2  
Introduction ......................................................................................................................... 6  
Duct Leakage Methods ....................................................................................................... 7  
    Commonly Used Duct Leakage Testing Methods .......................................................... 7  
        Duct Pressurization ................................................................................................. 7  
        Pressure Pan ............................................................................................................ 8  
        Delta-Q Test Development ...................................................................................... 8  
        The Nulling Test ..................................................................................................... 11  
Methodology ...................................................................................................................... 11  
    Site Characteristics/Recruitment .................................................................................. 11  
    Duct Leakage Testing ................................................................................................. 12  
Analysis .............................................................................................................................. 15  
Results ............................................................................................................................... 16  
    Ramping vs. Stations .................................................................................................... 16  
    Choice of Outdoor Reference ....................................................................................... 21  
    One-Ring Test vs. Two-Ring Test ................................................................................ 32  
    Error Estimation ......................................................................................................... 37  
    Comparison to Duct Pressurization and Pressure Pan .................................................. 39  
Conclusions ....................................................................................................................... 44  
Acknowledgments .............................................................................................................. 45  
References .......................................................................................................................... 46
Introduction

Duct leakage to outside has long been recognized as a major source of energy loss. However, many homes have ducts that do not leak appreciably, so efforts to diagnose the leakage from ducts in homes have become common so as to focus repair efforts on those homes for which they are warranted. The two most common test methods at the current time are the duct pressurization test and the pressure pan test.

The Delta-Q test for duct leakage was initially developed in 2001 (Walker et al. 2001). This test estimates duct leakage to outside, which is the leakage that results in substantial energy losses. It has several advantages over other residential duct leakage techniques. It requires only a blower door and a laptop computer; the supply and return registers and grilles do not need to be accessed; and the estimated leakage is based on the operating conditions of the forced-air distribution system.

The test also has undergone extensive laboratory and field evaluation, resulting in changes to both the sampling and analysis methodologies (Dickerhoff et al. 2004; Walker et al. 2004; Walker and Dickerhoff 2008). These efforts led to a 2006 study that evaluated various aspects of the “ramping” data collection technique for use in weatherization homes (Francisco 2006). The results of that study suggested significant promise for the Delta-Q test in this application, but recommended further evaluation of strategies to reduce the noise due to wind and if and how the test could be done successfully using only a single blower door ring instead of two. Working towards a quantitative uncertainty estimate was also cited as a primary need.

Since then further work has been done in the field to help characterize the uncertainty of the test method and how well it performs for other applications. For example, the Delta-Q test was performed in 19 new construction homes in Wisconsin, most of which were pursuing energy efficiency credits and which had most (but not all) of their ducts within the conditioned space (Pigg and Francisco 2008). This study compared the Delta-Q test results to other methods, including an advanced “nulling” method suitable only in research contexts that provided a comparison method that also evaluated duct leakage at operating conditions. This study also identified wind as a confounding factor, and took an initial attempt at developing a quantitative uncertainty estimate. Because of the nature of the homes in this study most had low leakage of both the homes and the ducts, making the results of limited application to weatherization homes.

Another study looked at repeatability of the Delta-Q test in several homes in California and Nevada (Dickerhoff and Walker 2008). Attempts to develop quantitative estimates of uncertainty for an individual house were of marginal success, though an estimate of 1% of house leakage rate at 50 Pascals (Pa) depressurization as an average value was proposed. The results of this study were consistent with those of Francisco (2006) with regard to the width of the pressure “bins” into which the data would be collected for analysis (5 Pa bins having been proposed by Francisco).
A third study evaluated repeated testing from several users around the U.S. to develop a means of estimating uncertainty by comparing the standard deviations among the repeated tests to the individual test standard errors of the leakage estimates (Olson 2008). Two methods of arriving at a multiplier to be used in conjunction with the standard errors of the leakage estimates were considered, with multipliers between 2 and 3 being most promising.

The study presented here is a follow-up study to the 2006 study by Francisco, and also makes use of the work done by Dickerhoff and Walker (2008), Pigg and Francisco (2008), and Olson (2008). The primary objectives were to reduce uncertainty due to wind, evaluate single-ring testing options, and further the development of quantitative uncertainty estimates. The advanced “nulling” technique was conducted in as many homes as possible for comparison. Homes were selected such that most homes had the majority of the duct work outside the conditioned space. Comparisons to the common techniques of duct pressurization and pressure pan tests were also conducted so as to assess the extent to which the tests gave different signals with regard to whether a duct system warranted air sealing.

**Duct Leakage Methods**

**Commonly Used Duct Leakage Testing Methods**

The two most common duct leakage test methods in use in field applications are the duct pressurization test and the pressure pan test. Duct pressurization is used by rating agencies such as RESNET (RESNET 2006) and is the method used by ASHRAE Standard 152 (ASHRAE 2004), while pressure pan testing is the most common method employed by low-income weatherization assistance programs.

**Duct Pressurization**

The most direct test method for duct leakage involves pressurizing the ducts while the registers are temporarily sealed, analogous to blowing up a balloon. This method is commonly referred to as the duct pressurization test, and is one of the methods found in the ASTM Standard on duct leakage measurement (ASTM 2007). The amount of air required to pressurize the ducts to a certain level is an indication of their airtightness. In order to separate supply and return leakage, which can be important in estimating energy penalties, a temporary airtight barrier is placed between the return and supply sides and the test is done on each side separately. In order to estimate only the portion of the leakage that goes to outside the conditioned space, which is also important in estimating energy penalties, these tests must be done with a second fan concurrently pressurizing the house such that the pressure between the ducts and the house is zero.

This leakage diagnostic method, however, is effectively only measuring the amount of holes in the ducts. The pressures at the holes would be required to know the leakage under normal operation. Pressures vary widely in the duct system, so making an assumption about the pressures often provides a poor estimate of the actual leakage
(Cummings et al. 2000; Francisco and Palmiter 2000). In addition, sealing of registers and grilles can be a very time-consuming process, especially if there are many registers and/or they are difficult to reach. Separating the supply and return sides from each other can also be very time-consuming.

Other problems can arise from inexact zeroing of the pressure between the ducts and the house when there is large leakage to inside (Pigg and Francisco 2008), or if there are significant pressure non-uniformities within the ducts.

**Pressure Pan**

Another method that is common in weatherization involves depressurizing the house and then covering the registers one at a time with a box and measuring the pressure across the box. This method, called the “pressure pan” test, is commonly used in weatherization due to its simplicity. The larger the pressure reading, the more well-connected the duct is to outside, implying leakage.

This method, however, is purely qualitative, requires the ability to access each register sufficiently well to completely cover them, and is prone to misinterpretation. A hole in the duct of a certain size will produce a larger pressure pan result the closer it is to the register, despite the fact that under normal conditions the operating pressures are lower at the register end of the system. Therefore, it rewards most the sealing of leaks that matter the least, since the holes near the register where the pressures are low will provide a stronger signal than holes further away where pressures are higher. The primary advantage of the pressure pan test is that it can identify a specific duct run that has a particularly large problem.

**Delta-Q Test Development**

The Delta-Q test makes use of the change in house pressure due to duct leakage to outside. By doing blower door tests over a range of house pressures, both with and without the air handler operating, a set of “delta-Q’s” can be obtained (Q being a common variable representing flow). Figure 1 shows an example. The resulting delta-Qs and their corresponding house pressures can be combined with a mathematical model to estimate the leakage.

The original Delta-Q test used a “stations” sampling methodology, wherein data was collected at each of a number of discrete house pressures, or stations. Figure 1 is an example of the stations methodology. Initially, the stations were between 5 and 25 Pa, both pressurization and depressurization, in 5 Pa increments. Later the pressure range was increased to have a maximum of 50 Pa, still in 5 Pa increments.

The equations for the Delta-Q test analysis also require estimates of a characteristic leakage pressure. Initially this was simply assumed to be either the plenum pressure or half of the plenum pressure. Following additional development these pressures were allowed to vary, and were included in a non-linear least-squares regression of the data as
additional variables. This was the version that was initially implemented into the ASTM Standard on duct leakage in 2003.

Figure 1. Flow vs. pressure from Delta-Q test (left) and resulting “delta-Q” vs. pressure (right)

After more research was done on the Delta-Q test the fitting of the leakage pressures was modified in two critical ways. First, the possible range of leakage pressures was constrained. This basic principle was originally espoused in 2004 (Walker et al. 2004). The limits that have eventually been settled on were a low of twice the minimum house pressure measured and a high of 100 Pa. This limitation prevented the analysis from diverging and producing spurious results. The second change was to implement a “scanning” technique. In this technique, rather than fitting the leakage pressures as a part of the regression, the pressures were incrementally changed through the range of possibilities and the regression was done with each set of assumptions. The regression that produced the lowest root-mean-square error was the one chosen as the resulting leakage estimates.

A later development was the “ramping” technique for data collection, which slowly increased and decreased the blower door speed while continuously recording data rather than collecting data at a set of discrete levels of house pressurization and depressurization (Walker and Dickerhoff 2008). The data are then collected into “bins” of a specified width, with 5 Pa having been settled on as a reasonable bin size. Figure 2 shows an example of the ramping methodology.

In addition to these changes to the testing methodology, two correction factors were developed to account for biases in the underlying mathematical model (Dickerhoff et al. 2004). Both of these are included in the ASTM Standard but implementation is not required. One of these, which accounts for the change in house pressure due to unbalanced leakage, is straightforward and is included in the software used for the Delta-Q testing. The other correction requires an iterative procedure as well as knowledge of the air handler flow, and is not a part of the software.
Figure 2. Example of the ramping methodology.

By 2006 the ramping methodology had become the focus of most Delta-Q efforts. Additionally, another method of assessing the leakage pressures had been developed, called non-negative least-squares (NNLS). This method had the benefit of allowing multiple leakage pressures for a single test result. Francisco (2006) focused on the ramping technique only, and compared NNLS to scanning. The conclusion was that NNLS produced sufficiently good results compared to the scanning technique; however, since then this approach has been abandoned in large part because it prevents negative leakage estimates. While negative leakage is clearly not physically possible it still can show up in measured data due to measurement uncertainty, especially when the leakage is low. Preventing such results means that a positive leakage bias could be introduced to any sample of homes.

By the time of the Wisconsin study (Pigg and Francisco 2008) the stations methodology had garnered renewed interest and was evaluated alongside the ramping technique. For control tests (in which the air handler is not turned on and the leakage estimates should be zero for both supply and return) the stations methodology appeared to have a slightly larger positive bias, though there were few homes on which both sampling methods were performed and the sample size was small enough to prohibit firm conclusions. Both methods performed similarly for normal tests where the air handler was operated. The project team did develop a subjective preference for the ramping technique, which was also perceived to be faster and provided clearer feedback about changes in wind.

Another subject that was raised by Francisco was the possibility of doing Delta-Q tests with only one ring. In general, to get the house pressures range specified by the testing, two rings are required, although the stations protocol occasionally can be met with a single ring (the ramping protocol specifies reaching a high of 60 Pa and then using a second ring to get data at low pressures).
Finally, a theme that has run through many Delta-Q projects is the need for an uncertainty estimate to be applied to the measured data. Attempts to develop such an estimate has been a major focus of work over the last two years.

**The Nulling Test**

An additional duct leakage diagnostic test, called the nulling test, was also performed in this study. As with the Delta-Q test, the nulling test is predicated on the idea that duct leakage causes changes in house pressure. With the nulling test, however, the estimate of leakage is determined by using a calibrated fan to counter, or “null”, the change in pressure. This has to be done twice, once with the air handler running normally (which provides an estimate of unbalanced duct leakage) and again with the return side of the duct system isolated and with the return air instead being provided by a surrogate airtight duct (which provides an estimate of supply leakage). Taking the difference of the two tests provides an estimate of the return leakage. The original version of this test was developed in 2001 (Francisco and Palmiter 2001). It was later modified for extended data collection and automation so as to provide more reliable results (Pigg and Francisco 2008).

Because the pressures generated in the nulling test are very small it is sensitive to wind. In addition, it is exceedingly difficult to exactly counter the pressure change caused by duct leakage. As a result, measurements over a range of pressures are required and sampling can take an extended period of time to reduce uncertainty due to wind, typically in excess of an hour for each part of the test. As a result it is not a practical test for broad use but can be useful in a research context. Because it is based on a fundamental principle and has limited underlying assumptions it is considered to be a good estimate of actual leakage, though it is still not exact. One source of potential error that has been identified occurs when the change in house pressure is a large fraction of the leakage pressure, as might occur when there is large leakage at register ends due to disconnected ducts.

**Methodology**

**Site Characteristics/Recruitment**

In this project testing was done on 14 homes, with one home being tested a second time for a total of 15 test periods. Testing took two days at each house. All homes had at least some ducting outside of the conditioned space. Two homes (Sites 1 and 5) were also included in the 2006 study. One home (Site 6) was located near Springfield, IL, and was recruited because of evidence of duct leakage. All other homes were within a 30 minute drive of Champaign, IL, and were recruited by word of mouth. No screening for duct leakage was done on these remaining homes to ensure that there was substantial leakage; however, the results show that most homes did have elevated duct leakage levels.

Table 1 shows several pertinent characteristics of the test homes. Site 15 was the repeated testing of Site 10. This retest was done because a very large leak at the plenum
had been identified in the initial testing, and a contractor had come out to the house to seal it. The repeated testing was performed to assess the success of this repair and to evaluate the ability of the Delta-Q test to quantify the improvement.

The average floor area of the homes was 1816 ft$^2$ and the average leakage rate was 10.6 ACH50. Most homes had crawl spaces. Homes with basements were multi-story, and also typically had a small section of the home over a crawl space and/or ducts in exterior walls or an attic. Of the 10 homes with a crawl space, only two had the air handler located within the conditioned space. Most had a downflow air handler in the garage, with the supply ducts in the crawl space and the return ducts in the attic. One home had all of the ducts in the attic except for the short sections required to get to the air handler in the garage. Two systems were entirely in the crawl space, including the air handler.

Table 1. Site Characteristics

<table>
<thead>
<tr>
<th>ID</th>
<th>Found. Location</th>
<th>Supply Location</th>
<th>Return Location</th>
<th>AH Location</th>
<th>Area (ft$^2$)</th>
<th>Volume (ft$^3$)</th>
<th>Stories</th>
<th>ACH50</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Base.</td>
<td>Base., Int. walls</td>
<td>Base., Ext. walls</td>
<td>Base.</td>
<td>1360</td>
<td>10880</td>
<td>2</td>
<td>12.2</td>
</tr>
<tr>
<td>2</td>
<td>Crawl</td>
<td>Attic</td>
<td>Attic</td>
<td>Garage</td>
<td>1980</td>
<td>15840</td>
<td>1</td>
<td>7.9</td>
</tr>
<tr>
<td>3</td>
<td>Crawl</td>
<td>Crawl</td>
<td>Attic</td>
<td>Util. rm.</td>
<td>1800</td>
<td>14400</td>
<td>1</td>
<td>8.0</td>
</tr>
<tr>
<td>4</td>
<td>Crawl</td>
<td>Crawl</td>
<td>Attic</td>
<td>Garage</td>
<td>1144</td>
<td>9152</td>
<td>1</td>
<td>7.6</td>
</tr>
<tr>
<td>5</td>
<td>Crawl</td>
<td>Crawl</td>
<td>Crawl</td>
<td>Garage</td>
<td>1800</td>
<td>14400</td>
<td>1</td>
<td>11.4</td>
</tr>
<tr>
<td>6</td>
<td>Crawl</td>
<td>Crawl, Attic</td>
<td>Attic</td>
<td>Closet</td>
<td>1200</td>
<td>9600</td>
<td>1</td>
<td>14.4</td>
</tr>
<tr>
<td>7</td>
<td>Crawl</td>
<td>Crawl</td>
<td>Attic</td>
<td>Garage</td>
<td>1960</td>
<td>16408</td>
<td>1</td>
<td>6.7</td>
</tr>
<tr>
<td>8</td>
<td>Crawl</td>
<td>Crawl</td>
<td>Attic</td>
<td>Garage</td>
<td>2200</td>
<td>17600</td>
<td>1</td>
<td>10.4</td>
</tr>
<tr>
<td>9</td>
<td>Crawl</td>
<td>Crawl</td>
<td>Attic</td>
<td>Garage</td>
<td>1600</td>
<td>12800</td>
<td>1</td>
<td>11.7</td>
</tr>
<tr>
<td>10</td>
<td>Crawl</td>
<td>Crawl</td>
<td>Attic</td>
<td>Garage</td>
<td>1360</td>
<td>10880</td>
<td>1</td>
<td>11.7</td>
</tr>
<tr>
<td>11</td>
<td>Crawl</td>
<td>Crawl</td>
<td>Crawl</td>
<td>Crawl</td>
<td>2787</td>
<td>23690</td>
<td>1</td>
<td>9.5</td>
</tr>
<tr>
<td>12</td>
<td>Base.</td>
<td>Garage, Int. walls</td>
<td>Garage</td>
<td>Garage</td>
<td>1203</td>
<td>9624</td>
<td>1.5$^2$</td>
<td>15.6</td>
</tr>
<tr>
<td>13</td>
<td>Base.$^1$</td>
<td>Base., Ext. walls</td>
<td>Base, Int. walls</td>
<td>Base.</td>
<td>2156</td>
<td>17248</td>
<td>2</td>
<td>11.0</td>
</tr>
<tr>
<td>14</td>
<td>Base.$^1$</td>
<td>Base., Attic</td>
<td>Base., Attic</td>
<td>Base.</td>
<td>3332</td>
<td>26656</td>
<td>2</td>
<td>8.3</td>
</tr>
<tr>
<td>15</td>
<td>Crawl</td>
<td>Crawl</td>
<td>Attic</td>
<td>Garage</td>
<td>1360</td>
<td>10880</td>
<td>1</td>
<td>12.8</td>
</tr>
</tbody>
</table>

Ave. 1816 14671 10.6

1. Sites 13 and 14 had a small area over a crawl space through which there was ducting.
2. Site 12 was a split-level house.

**Duct Leakage Testing**

A variety of Delta-Q tests were done at each house. At least 5 repeats of “normal” tests were done for each of the ramping and stations protocols; at least one ramping test was done with only one ring used but with the ramping being done twice for each step (so that the total sampling time was the same as for the normal tests); and in most cases one set of
“control” tests for each of the ramping and stations protocols was done, where the air handler was not operated.

In addition, pressure pan tests were done at all homes except Site 1. Duct pressurization and nulling tests were done whenever possible. Circumstances that prevented these tests from being possible included time limitations, very high wind (making nulling impractical), physical limitations (e.g. Site 11, which had the crawl space access inside the house and the air handler in the crawl space), and systems that had a long time delay for air handler fan operation that could not be overridden (making nulling impractical).

Table 2 shows the tests that were done at each house. Control tests were done at all but 4 houses. More than 5 sets of normal Delta-Q tests were done at 2 houses. Nulling tests were done at 11 homes. Nulling and duct pressurization were not done at Site 15 (the repeat of Site 10) due to an instrumentation problem.

Table 2. Leakage Tests done at Each Site

<table>
<thead>
<tr>
<th>ID</th>
<th>Delta-Q Tests</th>
<th>Nulling</th>
<th>Duct Press.</th>
<th>Pressure Pan</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Normal</td>
<td>One-ring</td>
<td>Control</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>√</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>√</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>√</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>√</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>√</td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td>3</td>
<td>1</td>
<td>√</td>
</tr>
<tr>
<td>7</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>√</td>
</tr>
<tr>
<td>8</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>√</td>
</tr>
<tr>
<td>9</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>√</td>
</tr>
<tr>
<td>10</td>
<td>5</td>
<td>3</td>
<td>1</td>
<td>√</td>
</tr>
<tr>
<td>11</td>
<td>7</td>
<td>3</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>6</td>
<td>4</td>
<td>0</td>
<td>√</td>
</tr>
<tr>
<td>13</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>√</td>
</tr>
<tr>
<td>14</td>
<td>5</td>
<td>3</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>√</td>
</tr>
</tbody>
</table>

In most cases all of the Delta-Q tests were done on one day, with perhaps the pressure pan tests also being done on the same day. In a few cases Delta-Q tests were done on both days, especially at those houses where more than 5 sets of normal tests were done. In general, the ramping tests and stations tests were interleaved, with ramping first followed by stations followed by ramping, etc. The one-ring test was usually done after the first two pairs of ramping/stations normal tests, and the control tests were usually done after the fourth pair of normal tests.

While the Delta-Q tests were being done there were additional pressures measured via a datalogger. These included the envelope pressure on the blower door side of the house; a “manifold” pressure created by connecting pressure hoses from all four sides of the house.
to a single pressure tap; the supply plenum; the return plenum; a supply register; a return grille; and the attic and crawl space pressures if applicable. Wind speed using a 3-cup anemometer and indoor and outdoor temperatures were also measured.

The nulling tests were done using automated software that controlled up to two calibrated fans as well as the air handler fan. To get the unbalanced leakage the computer cycled the air handler fan and a Duct Blaster® that was installed in place of the blower door. The fan in the door was used to provide a range of flows that would produce pressures that would bracket the change in house pressure resulting from the duct leakage. A flapper provided by Collin Olson that opened when the envelope fan was on and closed when the envelope fan was off enabled the envelope fan to be closed off whenever the test was in the fan-off mode.

The fans were cycled on and off at 1-minute cycles, 30 seconds off and 30 seconds on. This was done for eight different envelope fan flows, including one with the envelope fan off. Several sweeps through this set of flows were done at each house, typically resulting in at least one hour of testing. An example of the cycling is shown in Figure 3. In this graph the blue data are the supply plenum pressure indicating when the air handler is on, the red data are the house pressure, and the black data are the envelope fan pressure measurements.

Figure 3. Example of nulling test cycling for measuring unbalanced duct leakage.
When doing the nulling test for supply leakage only, an airtight barrier was placed in the filter slot to isolate the return side of the system and a second Duct Blaster® was attached to the air handler to assist the air handler in providing the return flow. In cases where the air handler was outside the conditioned space the second Duct Blaster® was installed in a doorway closest to the air handler (usually the door to the garage) and connected to the air handler via a 14-inch diameter flexible duct. This allowed all of the return flow to be drawn from the house. Once set up, the air handler was turned on and the assist Duct Blaster® was adjusted via the speed controller to provide the same supply plenum pressure as had been measured during normal operation. The speed controller was left at this setting and connected to a cable that was connected to the computer for cycling.

**Analysis**

There are a few details regarding the Delta-Q analysis that warrant discussion. One of the major issues of interest in this project was the evaluation of different envelope pressure options to minimize the impact of wind on the leakage results. The envelope pressure that was used for the primary reference during the testing was placed on the leeward side of the house, though for the pressurization portion of the Delta-Q tests the blower door fan pressure reference needed to be to the blower door side of the house. The additional pressures measured on the blower door side of the house and via the manifold were also of interest as potential references, but were measured using a different device and at different time intervals.

The data files from the Delta-Q test and the additional logging were combined and ordered chronologically. There were many more data points from the Delta-Q file. To synchronize the data to the same time stamps, surrogate data for the leeward side were created at the same time stamps as for the blower door side/manifold pressures by interpolating between the adjacent points. The Delta-Q data were chosen for interpolation instead of the additional data to minimize the amount of interpolation required. Once the combined datasets were created the Delta-Q analysis was performed for each of the outdoor pressure options. The analysis was performed on the interpolated leeward data in order to provide a fair comparison across the different outdoor pressures, with each option having the same number of data points available for analysis.

Another focus of the project was the potential use of only one ring for the Delta-Q tests. In addition to the tests done with double ramps but only one ring, a set of single-ring tests was created using the normal tests. At each house the range of house pressures achieved by each ring was assessed, and one of the rings was chosen to be the preferred ring for a single-ring version (this ring was used for the double-ramp tests). Copies of the data files for each normal Delta-Q test were made with only the selected ring included, and these data files were also analyzed for leakage estimates. This allowed the number of single-ring tests to be greatly expanded.
Results

The results section is organized sequentially to evaluate different issues with regard to Delta-Q. For example, the first discussion compares ramping and stations normal tests to assess which of those two testing methodologies performed better. The second discussion evaluates different outdoor pressure reference options, restricted to the testing methodology found to be superior in the first discussion.

The bulk of this section is focused on the Delta-Q test, with comparisons to the nulling test where appropriate. This includes evaluation of uncertainty estimates. Following the Delta-Q results there is a comparison to the duct pressurization and pressure pan tests.

Ramping vs. Stations

Several strategies were employed to compare the ramping and stations protocols. The first was a direct comparison of the two methods. Second was a comparison of the normal test results to the results when the house pressure data was “reduced”, meaning that the data that was interpolated to be synchronized with the other measured outdoor pressures was used. This is analogous to comparing results with a fast sampling rate vs. a slow one, indicating how sensitive the test is to greatly differing sampling rates. Third, comparisons to the nulling test results were evaluated.

Figures 4 and 5 show the comparison of the normal test results for the supply and return leakage, respectively. The leakage values are the averages of the pertinent tests for each house. These graphs show that the two methods produce similar results across houses, with the stations testing producing a statistically-insignificant higher result, on average. Linear regressions of the ramping method results on the stations method results had coefficients of 1.0 and $R^2$ values greater than 0.98, indicating that the two methods are virtually identical except for an offset. The offsets indicate that the stations tests produce leakage levels about 6 cfm higher on the supply side and 11 cfm higher on the return side, which are well within the standard deviations of the results of the two methods.

A problem with the stations tests shows up when comparing the normal test results to those using the reduced sampling data, as shown in Figures 6 and 7. Figure 6 shows this comparison for the ramping tests, which do not show a substantial change in results from the reduced sampling rate. Figure 7 shows that most cases using stations testing display little sensitivity to sampling rate, but there are a few cases where the results change by a more substantial amount. This is almost certainly due to the reduced sampling rate including more of the wind-induced excursions during data collection for the pressure station of interest, whereas the ramping technique is more able to lessen the impact of excursions because data are collected on both the up- and down-portions of the ramp.

This pattern is reinforced by Figure 8, which shows box plots of the differences between normal test results and reduced sampling rate results. In these plots, the line within the boxes is the median; the boxes represent the 25th through 75th percentiles (otherwise known as the first and third quartiles); the whiskers are at the most extreme value that is
within 1.5 times the interquartile distance (IQD, i.e. size of the box) of the box; and the symbols are outliers.

Figure 8 shows that the ramping tests have smaller IQDs, meaning less variation; the ramping tests have medians that are closer to zero change from normal tests; and the ramping tests have less extreme outliers.

Figure 4. Stations vs. ramping supply leakage results.

Figure 5. Stations vs. ramping return leakage results.
Figure 6. Comparison of reduced sampling rate results to normal test results, ramping.

Figure 7. Comparison of reduced sampling rate results to normal test results, stations.
Figure 8. Distributions of the changes in supply and return leakage estimates due to reduced sampling rate, ramping and stations tests.

Figures 9 and 10 compare the ramping tests to the nulling tests on the supply and return sides, respectively. The line represents perfect agreement with the nulling test results. The capped spike range around the line represents the 95% confidence interval of the nulling test results. The symbol represents the average normal ramping Delta-Q test result for each house, and the box around the symbol represents the 95% confidence interval of the Delta-Q test results. It should be noted that the 95% confidence intervals for the two tests do not mean the same thing. For the nulling test it is the 95% confidence interval for the mean, signifying that there is 95% confidence that the actual leakage estimate falls within that range. Since the nulling test is the method used as the best available comparative metric for operating duct leakage it is the confidence interval for the actual answer that is of interest. For the Delta-Q tests the 95% confidence interval represents the range in which it is expected that 95% of individual tests will fall. This is a wider range than if what was being considered was the 95% confidence that the sample mean was representative of the population mean. This way of representing the 95% confidence interval was chosen for the Delta-Q tests since in practice only a single Delta-Q test is likely to be done at a house and we want to know what range those results are likely to fall in.

Figures 9 and 10 have two common themes. The first is that there are three (for supply leakage) or four (for return leakage) out of eleven cases where the 95% confidence intervals do not overlap and seven or eight where they do. The second is that the Delta-Q test tends to always be higher than the nulling test. This apparent positive bias, which is usually of questionable statistical significance on individual houses, is partly due to not applying the second correction described in the discussion on the development of the Delta-Q test. However, given the difficulty of implementing this correction and the need
to also measure air handler flow, this correction was considered to be too complex for routine application and some positive bias needs to be accepted and accounted for in general use.

Overall, these results provide multiple reasons to prefer the ramping technique to the stations technique. The stations method appears to be more sensitive to sampling rate. The stations tests appear to have a small (though not statistically significant) positive bias relative to the ramping tests, which themselves display a positive bias relative to the nulling tests. Combined with the fact that the required time to do the ramping method of the Delta-Q test is more reliable and often shorter than the stations method (especially when there is noticeable wind), the ramping technique comes out slightly ahead of the stations protocol. The stations testing often took longer because the sampling protocol collected data until either a small standard error had been achieved on the house pressure or 30 seconds had elapsed, and only started collecting data after the house pressure became relatively stable. Under windy conditions it is common to collect data for the full 30 seconds for each station, which is 21 minutes just for data collection (including baseline measurement), plus time to turn the blower door around for pressurization and time to wait for the house pressure to stabilize before each station.

Figure 9. Comparison of normal test ramping Delta-Q supply leakage estimates to nulling test results.
Figure 10. Comparison of normal test ramping Delta-Q return leakage estimates to nulling test results.

**Choice of Outdoor Reference**

The Delta-Q test can be sensitive to wind noise. This is especially true in leakier houses, where a change in pressure corresponds to a greater amount of flow. It is therefore important to understand where best to put the outdoor reference pressure tap to minimize the impact of wind on the results. It is clearly easiest to use the pressure tap that is on the blower door side since only one outdoor reference pressure is then required, but if another strategy can greatly improve the results then it may be worthwhile setting up a second outdoor pressure tap.

Two strategies that have been considered for reducing noise due to wind are to place the outdoor reference tap on the leeward side of the house, regardless of where the blower door itself is located and to set up a pressure “manifold”. The manifold consists of pressure hoses run to each of the four sides of the house and then joined into a single hose that relays the pressure measurement to the datalogger (see Figure 11). The pressure from the manifold is the average of the pressures on the four sides of the house.

In this project three different outdoor pressure references were measured and evaluated: the leeward side, the manifold, and the blower door side. The leeward side pressure was measured by the digital manometer used for the Delta-Q test while the other two were measured by the pressure datalogger. In some cases the blower door was on the leeward side and so two pressure taps were installed on that side of the house. Regardless of what side was the leeward side the blower door was also referenced to the outside using the pressure tap on the blower door side of the house when pressurizing since that is the reference pressure pertinent to the fan pressure measurement.
All of the following discussion is focused on the ramping protocol only since that was the sampling methodology deemed preferable in the previous section. The hypothesis of this evaluation was that the pressure tap on the blower door side would be the worst of the three options, with the leeward and manifold pressures being better options. One of the questions regarding the manifold pressure is whether the noise reduction resulting from the averaging around the entire house was significant enough to warrant the hassle of setting it up. In this discussion, Figures 12-25 are at the end of the section in order to avoid breaking up the discussion.

Figures 12 and 13 show the mean leakage estimates for each house using each of the three outdoor pressure options compared to the estimate for the nulling test. In these graphs the 95% confidence intervals are not shown in order to avoid overloading the graphs.

Neither of these two graphs shows a significant difference among the three pressure options. The few cases where the one pressure option is slightly different are well within the 95% confidence intervals of the estimates, and they are not at the windiest sites.

Figures 14-16 show each of the three outdoor pressure options compared to the nulling test results for the supply side including the 95% confidence intervals, and Figures 17-19 show the corresponding results for the return side. All of these show similar characteristics, with 4-5 sites not having the Delta-Q and nulling 95% confidence intervals overlap. There are individual sites where each outdoor pressure option performs better than other options. Overall there is no apparent trend in these graphs to clearly prefer one option over another. Surprisingly, this includes the measurement on the blower door side of the house even if it is not the leeward side.
Table 3 shows the average leakage estimates and 95% confidence intervals (CI) for supply leakage ($Q_s$) and return leakage ($Q_r$), for each of the three outdoor pressure options. Table 3 is broken into two sections, one for all 15 sites and one for the 11 sites with nulling tests so that the results can be compared to those from the nulling tests. The confidence intervals are expressed as half of the width of the interval, such that the high and low values of the confidence interval are the mean ± the value shown in Table 3.

Table 3. Average Leakage Estimates and 95% Confidence Intervals for Sample Houses

<table>
<thead>
<tr>
<th></th>
<th>$Q_s$, Leakage</th>
<th>95% CI</th>
<th>$Q_r$, Leakage</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>All tests (n=15)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leeward</td>
<td>216</td>
<td>70.8</td>
<td>183</td>
<td>67.0</td>
</tr>
<tr>
<td>Manifold</td>
<td>195</td>
<td>97.0</td>
<td>169</td>
<td>64.0</td>
</tr>
<tr>
<td>Blower Door</td>
<td>220</td>
<td>54.9</td>
<td>193</td>
<td>60.4</td>
</tr>
<tr>
<td>Nulling sites (n=11)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nulling</td>
<td>168</td>
<td>23.8</td>
<td>106</td>
<td>31.4</td>
</tr>
<tr>
<td>Leeward</td>
<td>238</td>
<td>38.5</td>
<td>194</td>
<td>46.7</td>
</tr>
<tr>
<td>Manifold</td>
<td>235</td>
<td>29.2</td>
<td>190</td>
<td>31.5</td>
</tr>
<tr>
<td>Blower Door</td>
<td>243</td>
<td>40.2</td>
<td>201</td>
<td>44.3</td>
</tr>
</tbody>
</table>

Table 3 shows that, for the sites that include the nulling tests, the leakage estimates for the three outdoor pressure locations are all very similar and that there is an apparent reduction in noise using the manifold pressure. However, when expanded to all sites, it is the blower door side pressure that has the smallest amount of noise. Given that the windiest houses were among those for which nulling tests were not done, and with the purpose of using these alternate pressure locations being to reduce wind noise, these results do not encourage the use of either a leeward pressure or a manifold.

For the sites with nulling tests, Table 3 shows that the 95% CIs for the various Delta-Q leakage estimates do not overlap the 95% CIs for the nulling test estimates. This is largely due to a few cases. As shown in Figures 14-19 the majority of cases have 95% CIs that overlap, although there is a clear trend of the Delta-Q estimates being higher than the nulling test estimates. For the blower door side outdoor pressure cases, the median difference between the Delta-Q leakage estimate to the leakage from the nulling test is about 50 cfm, corresponding to about 5% of the average air handler flow of 958 cfm. Relative to the leakage estimate from the nulling test the Delta-Q estimates using the blower door side reference tap had a median of about one-third greater leakage for homes with at least 50 cfm of leakage. (Homes with lower leakage were screened out of this calculation because the low leakage means that any overestimate can be a substantial percentage even if the actual difference in flow is small).

Figures 20-23 show the distributions of the standard deviations across houses for the Delta-Q tests using the three pressure location options. (The standard deviations are effectively half of the 95% CIs.) Figures 20 and 21 show the results for the supply side and Figures 22 and 23 show the results for the return side, with the second of each pair of graphs omitting the standard deviations for Site 14. This site was essentially unsheltered.
and wind speeds were high during the test period, resulting in very large standard deviations for the leakage estimates, and viewing the graphs without this site allows for a better understanding of the distribution of standard deviations among the rest of the sites.

Figures 20 and 22 show that the manifold pressure had the worst performance of the three pressure options at Site 14, indicating that it was the least able to address wind noise at the windiest site. Figures 21 and 23 show that the leeward side pressure was the worst of the three at reducing wind noise across the sample with Site 14 excluded. The manifold pressure was better than the blower door side pressure on the return side (Figure 23), and produced lower standard deviations than the blower door side pressure for the supply side when the standard deviations were low. However, for the higher standard deviations on the supply side the blower door side performed better than the manifold pressure, and the overall supply-side distribution of standard deviations was more clustered using the blower door side pressure. Though these results do not provide overwhelming evidence of superior performance by the blower door side pressure than other options, neither do they provide a compelling argument that other options are sufficiently better to warrant the additional effort of running these extra pressure hoses.

There are a number of possible reasons why the blower door side pressure does not produce worse leakage results than options intended to reduce wind noise, including some amount of luck of the draw due to where in the data collection the wind noise causes the largest excursions. However, one contributing factor is that the blower door flow is impacted by the pressure on that side of the house, and when the leeward side is on another side of the house the envelope pressure using the leeward side can change opposite to the impact on the flow. For example, in the case where the leeward side is opposite the blower door side, the wind will pressurize the blower door side and will depressurize the leeward side.

Figures 24 and 25 show an example of this behavior, using one of the ramping tests from Site 11. Figure 24 shows the flow vs. pressure during the Delta-Q test using the leeward side pressure for the envelope pressure difference measurement, while Figure 25 shows the corresponding graph using the blower door side pressure for the envelope pressure difference measurement. Though both graphs show a lot of noise in the data due to wind, the use of the leeward pressure produces substantially more data where the flow-pressure changes are approximately perpendicular to the overall leakage curve. When the data are binned this will result in more data being included in each bin that are not reflective of the actual flow that should correspond to that pressure bin.

The conclusion of this evaluation is that, despite the apparent potential of implementing measures to reduce the noise due to wind on the Delta-Q leakage estimates, the benefits are minor and in some cases the attempt actually is counter-productive. This result does have some benefits to the user of the test. Only one outdoor pressure tap is required, which can be used in conjunction with a “T”-connector to provide an outdoor reference for both the house and the fan when doing the pressurization portion of the Delta-Q test.
Figure 12. Delta-Q supply leakage estimates from 3 outdoor pressure options.

Figure 13. Delta-Q return leakage estimates from 3 outdoor pressure options.
Figure 14. Supply leakage estimates using leeward outdoor pressure.

Figure 15. Supply leakage estimates using manifold outdoor pressure.
Figure 16. Supply leakage estimates using blower door side outdoor pressure.

Figure 17. Return leakage estimates using leeward outdoor pressure.
Figure 18. Return leakage estimates using manifold outdoor pressure.

Figure 19. Return leakage estimates using blower door side outdoor pressure.
Figure 20. Distribution of supply leakage standard deviations for three outdoor pressure options.

Figure 21. Distribution of supply leakage standard deviations for three outdoor pressure options with Site 14 excluded.
Figure 22. Distribution of return leakage standard deviations for three outdoor pressure options.

Figure 23. Distribution of return leakage standard deviations for three outdoor pressure options with Site 14 excluded.
Figure 24. Flow vs. pressure for Delta-Q test using leeward side pressure, site U11, 4th ramping test.

Figure 25. Flow vs. pressure for Delta-Q test using blower door side pressure, site U11, 4th ramping test.
**One-Ring Test vs. Two-Ring Test**

Two of the biggest logistical problems with the Delta-Q test are the time it takes to do the test and the need to change rings. This is especially an issue when pressurizing, since the rings are outside. As a result, if a protocol for a suitable one-ring Delta-Q test could be developed and if the resulting errors were acceptable, the utility of the test in a production-type program such as weatherization would be greatly enhanced. Even if doing two ramps per step was required to produce reliable results with a single ring there would be time savings because there would be no need to spend time changing rings, which also has the benefit of eliminating changes that could be done incorrectly. If a single ring could be used successfully and only one ramp per step was required then the testing time could be reduced from about 20 minutes to about 10 minutes, with the associated benefit of also producing a blower door house airtightness measurement that is already required by most programs.

Two techniques were implemented to evaluate the possibility of a single-ring test. One was to do tests that were intended to only use one ring but in which two ramps were done for each step, and the second was to take the two-ring tests and use only the ring that seemed most suitable for standing on its own. In practice, the ring that was used for the single-ring test was the smallest ring that could depressurize the house by approximately 30 Pa. This usually resulted in depressurization of the house by 40 Pa or more, though there were a couple of cases where this was not the case.

In all but 5 homes, only one test was done that was designed to use only a single ring with two ramps. As a result most houses do not have sufficient data of these tests to determine standard deviations among tests. However, standard deviations can be determined from the tests that were done as single-ring extractions from the normal tests.

In order to assess the representativeness of the extracted single-ring tests relative to the dual-ramp one-ring tests the estimates from the dual-ramp tests were compared to the average and 95% confidence intervals of the extracted tests. The results are shown in Figures 26 and 27 for supply and return leakage, respectively. These results use the blower door side outdoor reference.

These graphs show that the dual-ramp tests all fall within the 95% confidence interval of the extracted tests with one exception on the supply side, which was at Site 13. The dual-ramp tests are also essentially unbiased relative to the extracted tests. Linear regressions of the dual-ramp results on the extracted results produce offsets of about -10 cfm for the supply side and about 14 cfm for the return side. Overall, these results suggest that the results from the extracted tests are a suitable surrogate for intentional single-ring tests.

Figures 28 and 29 show the comparison of the extracted one-ring test average per-house results to the average per-house results from the normal two-ring tests for supply and return leakage, respectively. The numbers next to the symbols are the site ID numbers. These figures show that, with the exception of Site 14, there is little difference between the average per-house estimates regardless of whether one or two rings are used. Site 14
is the home at which there was high unsheltered wind and for which 95% confidence intervals are extremely wide (shown for extracted tests in Figures 26 and 27), and so the difference between one- and two-ring tests for Site 14 is far from statistically significant.

Figure 26. Comparison of supply leakage estimates from dual-ramp single-ring tests with extracted single-ring tests

Figure 27. Comparison of return leakage estimates from dual-ramp single-ring tests with extracted single-ring tests
Figure 28. Comparison of extracted one-ring supply leakage estimates to those from normal two-ring tests.

Figure 29. Comparison of extracted one-ring return leakage estimates to those from normal two-ring tests.
Further comparisons of the extracted one-ring tests and normal two-ring tests reinforce the conclusion that there is no statistical difference between the two. Including Site 14 in the analysis, linear regressions of the supply and return estimates produce slopes of about 1.04 and offsets of under 7 cfm, with \( R^2 \) values around 0.98. Excluding Site 14 reduces the slopes to about 1.03 with offsets of about 13 cfm and \( R^2 \) values over 0.99. Paired t-tests of the data indicate that the two sets of estimates are statistically identical.

Having established that there is no statistical difference between one-ring and two-ring tests, it is of interest to evaluate the extent to which the one-ring tests are less precise than the two-ring tests by analyzing the standard deviations of the estimates. Two-ring tests may be more precise than the extracted one-ring tests for a variety of reasons: they are more able to fully define the shape of the Delta-Q curve; twice as much data are used in the two ring tests; and the two-ring tests do not include the same time lags between steps of the test as do the extracted tests. This last issue is a characteristic of the one-ring data extraction process, and would not occur in a test intentionally using one ring. The second issue, the amount of data used in the analysis, can be addressed for one-ring tests by collecting data for two ramps at each step.

Figure 30 shows box plots of the standard deviations for the 15 homes for supply and return leakage estimates for both the two-ring tests and the extracted one-ring tests. Figure 31 shows the same data with Site 14 removed. Table 4 also provides a summary of pertinent statistical metrics of these standard deviations. Site 14 is also removed in Table 4 since it is clear that this test is anomalous and can greatly bias the summary.

The graphs and the table show that the medians of the sample-wide standard deviations are similar regardless of whether the test uses one or two rings, with the one-ring tests having median standard deviations of 3\% and 7\% higher than those for the two-ring tests on the supply and return sides, respectively. However, on the high end there is more of a difference, resulting in the interquartile distances (size of the boxes) for the one-ring tests being about twice as large as for the two-ring tests and the average sample-wide standard deviation being about 20\% higher for the one-ring tests.

Table 4. Summary of Sample-Wide Standard Deviations for Two-Ring and One-Ring Delta-Q Tests

<table>
<thead>
<tr>
<th></th>
<th>Two-ring tests</th>
<th>One-ring tests</th>
<th>Ratio, one-ring vs. two-ring</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Supply Return</td>
<td>Supply Return</td>
<td>Supply Return</td>
</tr>
<tr>
<td>Mean</td>
<td>20.8 21.4</td>
<td>25.0 25.2</td>
<td>1.20 1.18</td>
</tr>
<tr>
<td>Min</td>
<td>6.7 5.5</td>
<td>9.3 5.8</td>
<td>1.39 1.05</td>
</tr>
<tr>
<td>Median</td>
<td>18.6 18.7</td>
<td>19.1 20.1</td>
<td>1.03 1.07</td>
</tr>
<tr>
<td>Max</td>
<td>53.9 64.4</td>
<td>68.4 63.3</td>
<td>1.27 0.98</td>
</tr>
<tr>
<td>IQD\textsuperscript{1}</td>
<td>7.7 9.2</td>
<td>16.2 17.8</td>
<td>2.10 1.93</td>
</tr>
</tbody>
</table>

\textsuperscript{1} IQD is the interquartile distance, which is the difference between the 75\textsuperscript{th} and 25\textsuperscript{th} percentiles of the data. It is the height of the boxes in Figure 31, not including the whiskers.
Figure 30. Distribution of sample-wide standard deviations for two-ring and one-ring ramping tests, all sites.

Figure 31. Distribution of sample-wide standard deviations for two-ring and one-ring ramping tests, Site 14 excluded.
The conclusion from this analysis is that the one-ring tests are comparable to the two-ring tests but are more prone to greater error when the data are noisier due to effects such as wind. This strongly suggests that the benefit from the two-ring test is largely in the collection of more data, not in the greater range of house pressure data. Therefore, when conditions are calm, a single-ramp, single-ring test can be expected to produce results similar to a two-ring test, and when conditions are noisier a second ramp may be appropriate for a single-ring test. Under the calm conditions the testing time could be reduced to about 10 minutes and include producing a house blower door test result. Under the noisier conditions, if a second ring was used then the testing time would remain at close to 20 minutes but the hassle of changing rings would be eliminated.

**Error Estimation**

One problem with the Delta-Q test is that the uncertainty in the estimate, especially in the presence of substantial wind, can be large enough to throw into doubt whether the test is correctly indicating the existence of a duct leakage problem big enough to warrant repair. While all duct leakage diagnostic tests are prone to providing inaccurate feedback in some cases, the usefulness of the Delta-Q test would be greatly enhanced by having an uncertainty estimate for each house tested. Such an error estimate would provide guidance to the user regarding whether or not there was sufficient reliability in the result to make a retrofit decision.

Figure 32 shows the ratio of the supply and return leakage standard deviations, for both two-ring and extracted one-ring tests, and using the blower door side outdoor reference pressure. Consistent with Dickerhoff and Walker (2008), 1% of house leakage at 50 Pa is
a reasonable “typical” value for uncertainty. However, the spread around 1% shows that this value will often not be suitable for a specific house. The house leakage should have an impact, since a change in pressure across the building envelope will cause a larger change in flow in a leakier house than in a tighter house. However, the result on a calm day should also be a lot less uncertain than the result at the same house on a windy day.

In order to develop a more site-specific metric, the strategies employed by Olson (2008) were applied to the data in this project. Both strategies involve calculating the difference between each individual Delta-Q test leakage result and the average of all leakage results in a consistent set (e.g. all ramping tests using the same outdoor reference pressure and the same number of rings and ramps). These differences are then divided by the test-specific standard error of the leakage estimate as calculated by the Delta-Q test software. This calculation results in a multiplier for each supply and return leakage estimate for each test at each house that would be multiplied by the standard error from the software to provide an uncertainty estimate for that test.

In the “counting” strategy, the multiplier that would be used in general is the value that, when multiplied by the standard error to give an uncertainty estimate, results in at least 68% of the individual test leakage differences from the appropriate average estimate falling within the uncertainty estimate. The choice of 68% assumes a single standard deviation and is consistent with Olson. To generate a 95% confidence interval uncertainty the multiplier would be doubled from the one obtained for the 68% level.

In the “standardized error” strategy, the standard deviation of the multipliers from the individual tests is calculated, and this value is the multiplier used to provide the uncertainty at the 68% level. Again, this value would be doubled to provide uncertainty at the 95% confidence level.

Two sets of data were analyzed to determine potential multipliers. The first was the repeated normal tests, both two-ring and extracted one-ring versions, using the ramping protocol and the blower door side pressure. For this data the average leakage test result was used in calculating test-specific differences as described above. The second dataset was the control tests that were done at 11 houses. In the control tests the air handler was never turned on, so the correct leakage result should be zero and any leakage estimate other than zero is the error. For the control tests the leakage estimate itself was used as the test-specific difference.

Table 5 shows the results of this analysis. For the normal two-ring tests the multipliers all range from 1.9 to 2.3. For the one-ring tests the multipliers ranged from 2.7 to 3.2. For the control tests, which were all two-ring tests, the multipliers ranged from 3.2 to 3.7. The standardized error approach always produced multipliers slightly higher than the counting approach, consistent with Olson.

This analysis shows that the one-ring tests produce multipliers about 40% higher than the two-ring tests. For the counting method on the return side the one-ring value in Table 5 is about 70% higher than for the two-ring test, but an analysis of the individual site data
found that the average increase in return-side multiplier using the counting method was about 40% for the one-ring tests, but that a larger increase was required to meet the threshold of 68% of all tests having a difference from the set mean within the resulting uncertainty.

Table 5. Standard error multipliers from counting and standardized error approaches.

<table>
<thead>
<tr>
<th></th>
<th>Two-ring</th>
<th>One-ring</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Supply</td>
<td>Return</td>
<td>Supply</td>
</tr>
<tr>
<td>Counting</td>
<td>2.0</td>
<td>1.9</td>
<td>2.7</td>
</tr>
<tr>
<td>Standardized Errors</td>
<td>2.1</td>
<td>2.3</td>
<td>2.8</td>
</tr>
</tbody>
</table>

The following example illustrates how these multipliers would be implemented. Consider a Delta-Q test result that indicated 200 cfm of supply leakage with a standard error of 20 cfm. If we use the average of the two-ring standardized error multipliers, 2.2, then the uncertainty estimate would be 10 cfm*2.2 or 22 cfm. If we instead did a one-ring test and increased the multiplier to 3.0 then the uncertainty estimate would become 10*3.0 or 30 cfm, meaning that the leakage was between 170 and 230 cfm using a one standard deviation criterion.

The fact that the control tests produced a much higher set of multipliers suggests that the multipliers may scale with the leakage estimate. This makes intuitive sense since a wind-induced change in building envelope flow should not make as much difference to the leakage estimates if there is large duct leakage, since the change in building envelope flow would be a smaller fraction of the Delta-Q flows (difference between air-handler-off and air-handler-on flows). However, an analysis of the data did not show a statistically significant correlation of multiplier with duct leakage. Given the sample size, range of duct leakage levels, range of house leakage levels, and range of outdoor conditions, it is not surprising that no statistical correlation could be found, but the issue does warrant additional investigation.

It should be noted that all of the potential uncertainty estimates discussed focus on the precision (i.e. scatter), not the accuracy (i.e. bias). Comparisons with the nulling test done in earlier sections suggest a positive bias of the Delta-Q test in addition to the scatter. As stated previously, one of the correction factors that has been developed for the Delta-Q test (and optional in the ASTM standard) was not applied due to its complexity and additional measurements required. Its application would move the Delta-Q results into better agreement with the nulling test, but due the complexity it is probably going to be more efficient for production to approximate this correction factor separately, perhaps by discounting the Delta-Q result by 20% or so.

**Comparison to Duct Pressurization and Pressure Pan**

Of particular interest to production programs is whether the Delta-Q test provides feedback as to whether or not duct repairs are warranted that is at least as reliable as the methods currently used, duct pressurization and pressure pan.
Figures 33 and 34 compare the Delta-Q test result (two-ring using blower door side pressure) and the duct pressurization results to the nulling leakage estimates for the supply and return sides, respectively. The duct pressurization estimates presented were from separate supply and return leakage tests and done in conjunction with a blower door so as to provide leakage to outside only.

While the duct pressurization estimates do tend to increase with increasing leakage from nulling and Delta-Q, there are cases where the duct pressurization test results in significantly different feedback, especially on the supply side. On the supply side there are two cases where both the nulling and Delta-Q tests indicated low leakage under normal operation but the duct pressurization test indicated large leakage. This indicates that the leaks were at low pressures and that 25 Pa greatly overestimated the leakage. There is one other supply-side case where the duct pressurization strongly indicated large leakage (greater than 300 cfm @ 25 Pa) while the nulling and Delta-Q tests were less certain that the leakage was large (about 160 cfm). For other cases on the supply side the tests gave similar feedback regarding the merits of air sealing, even when the actual estimates differed substantially.

On the return side shown in Figure 34 the comparison of Delta-Q and duct pressurization is more mixed. There are two cases where the duct pressurization indicates that the leakage is not particularly large consistent with the nulling test but the Delta-Q is more indicative of larger leakage. There are also two cases where the Delta-Q test more strongly agrees with the nulling test regarding the merits of air sealing, one in which the merits are questionable (Delta-Q about 140 cfm, duct pressurization about 300 cfm @ 25 Pa) and one in which the leakage appears to merit sealing (the right-most point).

Figure 33. Duct Pressurization and Delta-Q estimates compared to nulling, supply.
Figure 34. Duct Pressurization and Delta-Q estimates compared to nulling, return.

Given the relatively greater importance of supply leakage on energy losses, the tendency of the Delta-Q test to better assess the merits of sealing the supply ducts than the duct pressurization test argues for preferring the Delta-Q test over the duct pressurization test.

Figures 35 and 36 compare the average and median register pressure pan results for each house to the Delta-Q test estimate for the supply and return sides, respectively. In Figure 36 the cases where the two symbols lie atop each other have only one return.

On the supply side there is a better correlation between the median pressure pan measurement and the Delta-Q leakage estimate, suggesting that the median value would be a better metric to indicate excessive leakage than either the mean or any individual register measurement. There are no cases where the average or median supply pressure pan is above 4 Pa and the Delta-Q estimate is above about 150 cfm. Note that this does not mean that there are no individual registers that have pressure pan reading above 4 Pa, but it does indicate that for this set of homes an average pressure pan reading greater than 4 Pa was necessary before the Delta-Q test (or nulling, since nulling results were lower than Delta-Q estimates) indicated that there was leakage above 150 cfm. There was also one case where the average pressure pan reading was over 7 Pa yet the Delta-Q test indicated that the leakage was only about 120 cfm. (The nulling result for this case was about 110 cfm, and the median pressure pan was below 3 Pa.) There was one other case where the average pressure pan reading was just under 7 Pa and both the Delta-Q and nulling tests indicated a supply leakage of about 170 cfm, which may be a marginal rate to warrant air sealing for some programs. The median pressure pan reading was just short of 6 Pa for this house.
At higher Delta-Q supply estimates the pressure pan was in agreement that air sealing was warranted with no exceptions. In other words, there were no Delta-Q “false positives” where the Delta-Q test indicated that there was leakage worth sealing but the ducts were actually tight. However, the Delta-Q test did catch several pressure pan “false positives” where air sealing would have been suggested by the pressure pan when it was likely not a high priority.

Figure 35. Pressure pan results compared to Delta-Q estimates, supply.

Figure 36. Pressure pan results compared to Delta-Q estimates, return.
Figure 36 shows similar results for the return side. Again there were no false positives by the Delta-Q test, but there were cases where the pressure pan reading was greater than 5 (usually interpreted as warranting air sealing) though the Delta-Q test (and therefore nulling test) indicated leakage of 150 cfm or below.

Taken together, the Delta-Q test shows substantially superior performance compared to the pressure pan test, and focuses resources more appropriately.

Aside from the comparison of the duct pressurization test to the Delta-Q test, other issues with the duct pressurization test were noted. Two assumptions of the duct pressurization test is that the entire duct system is pressurized to the same pressure and that the operation of the blower door causes the entire duct system to have a pressure of zero relative to the house. To the extent that these assumptions prove untrue the leakage estimates from the duct pressurization test are misleading. Some of the measured leakage may be at very different pressures than 25 Pa, and if some of the ducts are not at zero relative to the house then some leakage that is said to be to outside may be to inside.

Two examples of these problems follow. The first is for a supply-side only test, with the return isolated as was done for the results in the previous discussion. The second is for a full duct system test, including both supply and return. This method is often used to avoid the need to isolate the return from the supply side and to do two separate tests.

For the first example, done at Site 5, the supply duct pressure used to zero out the pressure between the house and the ducts was at the supply plenum. When the house was pressurized by 42 Pa (as high as could be reached in the ducts) and the Duct Blaster® was turned up until the supply plenum was zero relative to the house, the pressure at the supply register was about -8 Pa relative to the house. When the house and supply plenum were both at 25 Pa (zero between the supply plenum and the house) the supply register was about -5 Pa relative to the house. This indicates that there was house air entering the duct system at the register end even though the pressure difference between the ducts and the house was nominally zero and the assumption was that there was no transfer of air between the house and the ducts.

For the second example, done at Site 10, the Duct Blaster® was installed at a return grille and the entire duct system was tested. The nominal target duct pressure was measured at a different return grille, and pressures were also measured at each plenum and at the supply register. Without the blower door running (i.e. total leakage including to inside) and with the return grille at 50 Pa, the return plenum was at 33 Pa and the supply plenum and register were each at about 10 Pa. All of these values were about halved when the return grille was at 25 Pa. When the blower door was operated to get leakage to outside, and again using the return grille as the location at which the nominal (zero) pressure between the house and ducts was measured, the return plenum was at -14 Pa and the supply side at -30 Pa relative to the house when the return grille and house were at 50 Pa, and about halved when the return grille and house were each at 25 Pa.
These measurements show that the leakage nominally at 25 Pa or 50 Pa was actually only that at one location but very different in other places. For the Site 10 example, if the return plenum had been used as the reference location that resulting nominal leakage at 25 and 50 Pa would have been substantially higher since more flow through the Duct Blaster® would have been required to achieve the target pressures.

**Conclusions**

A number of valuable conclusions have resulted from this project.

- The ramping sampling protocol is preferable to the stations sampling protocol.

- There is no substantial benefit to placing the outdoor reference tap in a location other than the blower door side of the house, and in some cases is detrimental. Therefore the blower door side of the house is the recommended location for this pressure tap.

- The one-ring version of the Delta-Q test, with the ring chosen to be the smallest that can reach at least 30 Pa house depressurization, produces results comparable to the two-ring test. Under calm conditions the test should produce reasonable results using a single ramp, but a second ramp is recommended when there is noticeable wind.

- The one-ring test with one ramp has approximately 40% greater uncertainty than the two-ring test. It may be that this increased uncertainty can be removed with a dual-ramp test, and this possibility warrants further investigation.

- To provide uncertainty estimates for the Delta-Q leakage estimates, the results from these homes suggest that the standard error from the software be multiplied by about 2.2 for the two-ring tests or about 3.0 for the single-ring, single-ramp tests.

- The Delta-Q test performs better than both the duct pressurization and pressure pan tests at indicating the need to repair the ducts. This is especially true for the pressure pan tests using typical interpretations.

- The median value of the pressure pan tests had much better correlation with the Delta-Q and nulling leakage estimates, suggesting that if the pressure pan is to be used that the median value would be a preferable metric to use when determining whether to perform air sealing.

- The duct pressurization test is subject to errors in leakage estimates due to non-uniformities in the pressures within the ducts, which increases the uncertainty about the reliability of the estimates.
Implementation Recommendations

This report has focused on an objective analysis of the Delta-Q test, both on how to best do the test and how it compares to other tests. This section provides the author’s views on how and when the Delta-Q test should be considered for use, based on this project and previous work done by the author and others. The recommendations that follow do not preclude the user from walking around the house during blower door operation and feeling for air leaks, including from registers, but rather provide guidance on quantifying duct leakage problems.

In existing home programs such as Home Weatherization Assistance Programs, the Delta-Q test is an attractive option unless wind noise is high. With suitable uncertainty estimates it can provide at least a quick screening of whether there is substantial duct leakage. In 10-20 minutes it can produce a leakage estimate (as well as a blower door result) without the need to access each register and grille. The advantage of not having to access registers and grilles, which in occupied houses can be logistically troublesome and which are frequently of odd shapes, is often significant. In many cases the Delta-Q test will indicate, within the uncertainty estimate, that there is either too little leakage to worry about under the context of the program or that there is clearly enough leakage to warrant repairs. In those cases where, given the uncertainty, the leakage levels are borderline with regard to retrofit it may be warranted to do additional duct testing but overall it seems highly likely that the cumulative testing time across houses would be reduced by doing the Delta-Q test first except under conditions of high winds or in homes with few, easily accessed and standard shape registers.

In houses with few, easily accessible registers of standard shape, it may be that the pressure pan test is just as fast and useful for indicating whether or not duct retrofits are required. In such cases where the pressure pan test may be warranted instead of Delta-Q (or when wind noise is high enough to make the uncertainty estimate for the Delta-Q results too large for reliable conclusions) it is the median pressure pan reading that appears most suitable for drawing conclusions.

For programs with a tight specification on leakage, as is often the case in new construction, the Delta-Q test may be unclear about whether the ducts “pass” or “fail” because the uncertainty is often of a similar size to the specification. However, with a suitable uncertainty estimate applied to the test results the Delta-Q has more potential utility in this application as well. Given that the primary alternative for tight duct systems in many of these programs is the duct pressurization test, which can take a long time to set up and perform, doing the Delta-Q test as part of the blower door test for envelope tightness may be a reasonable screen, and if the result is uncertain due to the error bars on the Delta-Q results then additional testing could be pursued. As with the existing home programs this would save time in many homes.

An additional note for new construction is that the ducts should be tested at rough-in so that the ducts are still accessible for easy repair if warranted.
Acknowledgments

This material is based upon work supported by the Ohio Department of Development and the U.S. Department of Energy under Award No. CMDD 08-003. Tom Andrews was the Project Manager.

Collin Olson of The Energy Conservatory wrote the software, including amending it as issues arose, and provided technical support and feedback on data collection protocols and data analysis. Gary Nelson of The Energy Conservatory provided the support of the company to assist in the project.

Scott Pigg of Energy Center of Wisconsin provided input and guidance from his experiences with the Wisconsin project, especially with regard to the nulling test.

Special thanks go to all of the residents who allowed the use of their homes in this project. None of this would have been possible without them.

Disclaimer

Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the author and do not necessarily reflect the views of the Ohio Department of Development or the U.S. Department of Energy.

References


