AIAA JOURNAL
Vol. 44, No. 3, March 2006

Proper Estimation of Random Uncertainties in Steady-State Testing

Hugh W. Coleman* and David M. Lineberry†
University of Alabama in Huntsville, Huntsville, Alabama 35899

The estimation of the random uncertainty of an experimental result (such as specific impulse or drag coefficient) that is determined from multiple measured variables in a steady-state test is considered. Often in the actual engineering testing of complex systems, all conditions cannot be tightly controlled, and the variations in successive measurements of the variables are caused both by unsteadiness and by measurement errors. In such cases, the distributions of the measurements of different variables might not be independent of one another, that is, the variables "drift" in unison. The results of two test programs are used to show that calculation of the random standard uncertainty (the standard deviation) of a result using the traditional propagation equation can produce an estimate many times too large or too small because the correlations among the variations in the measured variables are not taken into account. In contrast, when the random standard uncertainty is calculated directly from a sample of multiple results determined within a single test, the effects of the correlations are accounted for implicitly. The direct approach has not commonly been used in single test situations because unfortunately it has been presented in the past with nomenclature that caused a mindset that experimental results from multiple tests were necessary for its use, and that is not the case.

I. Introduction

THIS paper considers the estimation of the random uncertainty of an experimental result that is determined in a steady-state test. When a steady-state test is conducted, it is intended to model a situation in which all experimental conditions are constant, and the experimental result is thus constant during the test period. In actual practice, of course, measured values of a variable at successive times are not the same: they vary with time as a result of the effects of measurement errors that are not constant and also because of the inherent unsteadiness of the process that is at "steady state." The standard deviation of the sample of measurements of a variable taken during the test period is called the *random standard uncertainty*, and usually that standard deviation is multiplied by a coverage factor (often taken as 2) to obtain the random uncertainty of the variable at some coverage or confidence level (often taken as 95%).[1–3]

In most tests an experimental result of interest is determined by combining in a data-reduction equation the measured values of several variables and calculating a value for the result. Consider as an example the testing of a rocket engine in a thrust stand to determine the specific impulse of the engine. (This is one of two tests whose results are discussed later in this paper.) The specific impulse I is determined by dividing the thrust by the total weight flow rate of oxidizer and fuel. In the test considered here, the thrust was measured independently by two load cells, and so the thrust is calculated as the average of the two measured thrust values F_1 and F_2. The data-reduction equation for specific impulse is then

$$I = \frac{0.5(F_1 + F_2)}{(\rho_{\text{ox}} Q_{\text{ox}} + \rho_f Q_f) g} \tag{1}$$

where the subscript ox refers to oxidizer and f refers to fuel, ρ is density, Q is volumetric flow rate, and g is gravitational acceleration. The densities and g were taken from reference data and were considered as constants (with only systematic uncertainty) for a given test.

Shown in Fig. 1 are measured values of thrust and volumetric flow rates during a steady-state period from one of the 17 rocket engine tests discussed in more detail later. In a test, the data were averaged over 1-s intervals and recorded throughout the test period. In the figure, the data have been normalized by dividing each recorded value by the test average, and they have also been shifted on the ordinate to show comparative trends more clearly. (The additional subscript n indicates a quantity that is normalized.)

In common engineering practice, an average value and a random standard uncertainty are computed for the given test period for each variable; the average values of the variables are inserted into Eq. (1) to calculate a test-average value of specific impulse, and the random standard uncertainties of the measured variables are inserted into a propagation equation to calculate a random standard uncertainty for the specific impulse. The propagation equation normally used assumes that the variations in the different variables are independent of one another. It is clear from the figure that this assumption is not a good one for this test. The variations in the variables appear to be caused primarily by process unsteadiness, and those variations are not independent of one another: they are correlated. The implications of this observation are demonstrated with results from two different test programs in a later section.

In another approach, a value of the result I can be calculated at each time the values of the measured variables are recorded, thus creating a sample of results over the test period. Figure 2 shows the measurements of the four variables presented earlier in Fig. 1 and additionally a calculated value of I at each time the measurements are recorded throughout the test. An average and standard random uncertainty for I can now be directly calculated from the sample of I results, and this standard deviation includes the effects of the correlated variations of the variables. (The additional subscript n indicates a quantity that is normalized.)

Unfortunately, this direct calculation approach has been presented in the past with nomenclature that has caused a mindset that experimental results from multiple tests were necessary for its use. In the 1985 version[4] of the American Society of Mechanical Engineers Standard PTC 19.1, the approach was presented with the heading "uncertainty—more than one test," and in the 1998 version[3] the heading used is "multiple test." In both the first (1989) and second (1999) editions of Coleman and Steele,[1] "direct determination. . . : multiple tests" is used. The heading "multiple tests" was also used in both the 1995 (Ref. 5) and 1999 (Ref. 6) versions of the AIAA

Received 30 October 2004; revision received 27 June 2005; accepted for publication 29 August 2005.

*Professor, Propulsion Research Center, Department of Mechanical and Aerospace Engineering. Associate Fellow AIAA.

†Graduate Research Assistant, Propulsion Research Center, Department of Mechanical and Aerospace Engineering. Student Member AIAA.

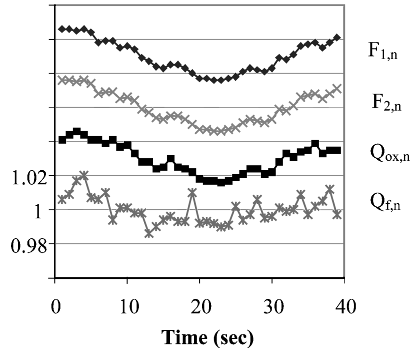

Fig. 1 Normalized measurements from a liquid-rocket-engine ground test.

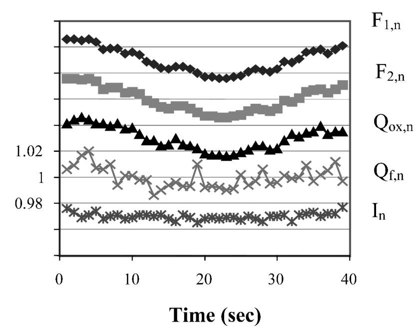

Fig. 2 Normalized variables and specific impulse from a liquid-rocket-engine ground test.

Standard S-071. As explained in the preceding paragraph, the direct calculation approach can be used in a single test situation.

In the next section, a brief review of experimental uncertainty analysis is presented as background for the discussion that follows of results from two test programs. The first set of results considered is from full-scale, hot-firing ground tests of a liquid rocket engine. Obviously in such a test facility there are more uncontrollable factors than in a bench-top experiment in a laboratory. In these tests, the performance of the engine is influenced by variations in the propellant mass flow rates, which affect the combustion performance and, consequently, the thrust. The average thrust and the specific impulse are the experimental results presented and discussed. The second set of results considered is from a test program using a laboratory-scale cold-flow ejector apparatus. The ejector experiment was carried out in a well-controlled laboratory environment, and the results considered are the secondary (induced) mass flow rate and the suction ratio.

In both types of tests, the random uncertainty determined using the traditional propagation approach (with no correlation assumed) differed from the random uncertainty determined using the direct calculation method.

II. Experimental Uncertainty Analysis

Uncertainties in Measurements of a Single Variable

The total error in a measurement of a variable is the difference between the measured value and the true value. This total error is considered to be the combination of errors from many elemental sources, and these elemental errors have traditionally been divided into two categories: systematic errors and random errors. Errors that are present in the same amount for each measurement are classified as systematic errors, and they do not contribute to scatter in the sample of measured values. Those errors that do contribute to the scatter in the measurements of a quantity at a steady state are classified as random errors, and these tend toward being normally distributed by virtue of the Central Limit Theorem.[1] As discussed in the preceding section, another factor that contributes to scatter in the measurements is unsteadiness in the postulated steady-state process. This corresponds to temporal variations of the true value of the quantity during the period measurements are being made, and these variations do not follow any particular distribution. The effects of both the random measurement errors and the nonrandom unsteadiness are generally lumped together into a random uncertainty as discussed next.

A systematic uncertainty for a measured variable is an estimate of the range in which the true systematic error lies. A standard systematic uncertainty estimate b is typically made such that the interval $\pm 2b$ will contain about 95% of the possible systematic errors that could be realized from a parent distribution for a particular error source.[1]

An indicator of the amount of scatter in a sample of multiple measurements of a variable is the sample standard deviation S. For a sample containing N measurements of a variable X, the sample standard deviation S_X is defined as

$$S_X = \left[\frac{1}{N-1} \sum_{i=1}^{N} (X_i - \bar{X})^2 \right]^{\frac{1}{2}} \tag{2}$$

where X_i is a single measurement and $\bar{X}$ is the mean of the sample calculated from

$$\bar{X} = \frac{1}{N} \sum_{i=1}^{N} X_i \tag{3}$$

The sample standard deviation is termed the *standard uncertainty* in the *International Organization for Standardization Guide.*[2] In common engineering practice in the United States,[1,3] S_X is termed the *random standard uncertainty*, although as pointed out in the preceding discussion the variations in the measurements might not be truly random. The random standard uncertainty of the mean $S_{\bar{X}}$ is given by

$$S_{\bar{X}} = S_X \big/ \sqrt{N} \tag{4}$$

Uncertainties in a Calculated Experimental Result

When a result r is calculated from several measured variables using a data-reduction equation [such as Eq. (1)], a general functional formulation is

$$r = r(X_1, X_2, \ldots, X_J) \tag{5}$$

where the X are the measured variables (such as F_1, Q_{ox}, etc). The uncertainties in the measured variables cause uncertainty in the result, and this is often modeled using a propagation equation based on a Taylor-series expansion. Details of the derivation and assumptions involved are given in Appendix B of Ref. 1.

Using the propagation equation approach, the systematic standard uncertainty b_r of the calculated result is given by

$$b_r^2 = \left(\frac{\partial r}{\partial X_1}\right)^2 b_{X_1}^2 + \left(\frac{\partial r}{\partial X_2}\right)^2 b_{X_2}^2 + \cdots + \left(\frac{\partial r}{\partial X_J}\right)^2 b_{X_J}^2 + 2\left(\frac{\partial r}{\partial X_i}\right)\left(\frac{\partial r}{\partial X_k}\right) b_{X_i X_k} + \cdots \tag{6}$$

where the b_X are the systematic standard uncertainties associated with the X variables and there is a correlated systematic uncertainty term containing a covariance factor b_{XiXk} for each pair of variables that share a common elemental systematic error source.[1] Correlated systematic uncertainties occur often in engineering testing (such as when multiple transducers are calibrated with the same standard), and it is important to include consideration of the correlation terms in the b_r propagation equation.

An analogous propagation equation for the random standard uncertainty S_r of the calculated result is

$$(S_r)_{\text{prop}}^2 = \left(\frac{\partial r}{\partial X_1}\right)^2 S_{X_1}^2 + \left(\frac{\partial r}{\partial X_2}\right)^2 S_{X_2}^2 + \cdots + \left(\frac{\partial r}{\partial X_J}\right)^2 S_{X_J}^2 + 2\left(\frac{\partial r}{\partial X_i}\right)\left(\frac{\partial r}{\partial X_k}\right) S_{X_i X_k} + \cdots \tag{7}$$

where there is a correlated random standard uncertainty term containing a covariance factor S_{XiXk} for each pair of measured variables whose random variations are not independent of one another.[1] These correlated random uncertainty terms have traditionally always been assumed to be zero, so that the propagation equation actually used is

$$(S_r)^2_{\text{prop}} = \left(\frac{\partial r}{\partial X_1}\right)^2 S^2_{X_1} + \left(\frac{\partial r}{\partial X_2}\right)^2 S^2_{X_2} + \cdots + \left(\frac{\partial r}{\partial X_J}\right)^2 S^2_{X_J} \quad (8)$$

If the X values used in Eq. (5) are single measured values, the standard uncertainties in Eq. (8) are the S_X values from Eq. (2). If the X values used in Eq. (5) are mean values $\bar{X}$, the standard uncertainties in Eq. (8) are the $S_{\bar{X}}$ values from Eq. (4), and the random uncertainty calculated from Eq. (8) is the random standard uncertainty of the mean result $(S_{\bar{r}})_{\text{prop}}$.

As discussed briefly in the introductory section, another way of estimating S_r in a steady-state test is by direct calculation using an equation analogous to Eq. (2). As an example, suppose a result r is a function $r = r(y, z)$ of two measured variables y and z, and both variables are measured at time $t = 1, 2, \ldots, 20$ s so that $N = 20$. Values of S_y and S_z can be calculated from Eq. (2) and then $(S_r)_{\text{prop}}$ determined using the propagation Eq. (8). But a value of r can also be calculated at $t = 1, 2, \ldots, 20$ s (as in Fig. 2) so that one has a sample of N determinations of r. Then a direct calculation of S_r can be made using an equation analogous to Eq. (2):

$$(S_r)_{\text{direct}} = \left[\frac{1}{N-1}\sum_{i=1}^{N}(r_i - \bar{r})^2\right]^{\frac{1}{2}} \quad (9)$$

where

$$\bar{r} = \frac{1}{N}\sum_{i=1}^{N} r_i \quad (10)$$

Because the sample of N values of r implicitly contains the effects of any correlated variations among the measured variables, so does the $(S_r)_{\text{direct}}$ calculated using Eq. (9). An $(S_{\bar{r}})_{\text{direct}}$ can be defined using an equation analogous to Eq. (4):

$$S_{\bar{r}} = S_r \big/ \sqrt{N} \quad (11)$$

In steady-state tests in which the variations of the measured variables are not correlated, $(S_r)_{\text{prop}}$ and $(S_r)_{\text{direct}}$ should be about equal (within the approximations inherent in the propagation equation). However, when the variations in different measured variables are not independent of one another, the values of S_r from the two equations can be significantly different. This was shown by Hudson et al.,[7] who published the results of a venturi calibration test in which the S_r values from the two approaches differed by more than an order of magnitude in some cases.

Examination of Eqs. (6) and (7) shows that each of the correlation terms contains a covariance factor (b_{XiXk} or S_{XiXk}) multiplied by the product of the partial derivatives of the result with respect to the two variables. Assume the covariance factor is positive. If the partial derivatives are of opposite sign, the correlation term is negative, and the effect of the correlation is to reduce the uncertainty. On the other hand, if the partial derivatives are of the same sign, the correlation term is positive, and the effect of the correlation is to increase the uncertainty. Both of these behaviors were observed in the experimental results discussed in the following section.

III. Results and Discussion

The first set of results considered is from full-scale, hot-firing ground tests of a liquid rocket engine. The average thrust and the specific impulse are the experimental results examined. The second set of results considered is from a test program using a laboratory-scale cold-flow ejector apparatus. The ejector experiment was carried out in a well-controlled laboratory environment, and the results considered are the secondary (induced) mass flow rate and the suction ratio.

The relative magnitudes of the random uncertainties determined using the propagation method and the direct method are of interest in this discussion, so that $(S_r)_{\text{prop}}$ and $(S_r)_{\text{direct}}$ as given in Eqs. (8) and (9) are compared in the following analysis. Obviously, the ratio would remain the same if both were divided by $\sqrt{N}$ to produce $(S_{\bar{r}})_{\text{prop}}$ and $(S_{\bar{r}})_{\text{direct}}$.

Rocket-Engine Tests

Results from 17 tests of eight "identical" liquid rocket engines are considered. Data were recorded as 1-s averaged values (as shown in Fig. 1), and the steady-state test period was taken as a 39-s interval after the initial startup transient. The thrust was measured independently with two load cells, and an average thrust F was computed using the data-reduction equation

$$F = 0.5(F_1 + F_2) \quad (12)$$

As described earlier, the specific impulse was determined from

$$I = \frac{0.5(F_1 + F_2)}{(\rho_{\text{ox}} Q_{\text{ox}} + \rho_f Q_f)g} \quad (1)$$

where the volumetric flow rates were measured using turbine meters.

For each test, S_F was determined using Eqs. (8) and (9), and the results for the 17 tests are compared in Fig. 3. In each test, $(S_F)_{\text{direct}}$ is greater than $(S_F)_{\text{prop}}$ by about 40%, as seen in Fig. 4, where the ratio is plotted.

Consider a propagation equation similar to Eq. (7) applied to the data-reduction equation for average thrust [Eq. (12)]:

$$(S_F)^2_{\text{prop}} = \left(\frac{\partial F}{\partial F_1}\right)^2 S^2_{F_1} + \left(\frac{\partial F}{\partial F_2}\right)^2 S^2_{F_2} + 2\left(\frac{\partial F}{\partial F_1}\right)\left(\frac{\partial F}{\partial F_2}\right) S_{F_1F_2} \quad (13)$$

Both of the partial derivatives are equal to $+\frac{1}{2}$, and $S_{F_1F_2}$ is positive because the outputs of the two load cells vary in unison. This means that the correlation term is positive, and when it is neglected as in Eq. (8) the random standard uncertainty $(S_F)_{\text{prop}}$ that is calculated is erroneously low.

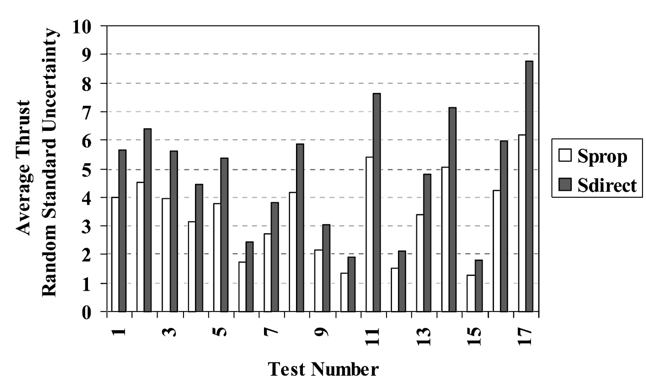

Fig. 3 Random standard uncertainties for average thrust.

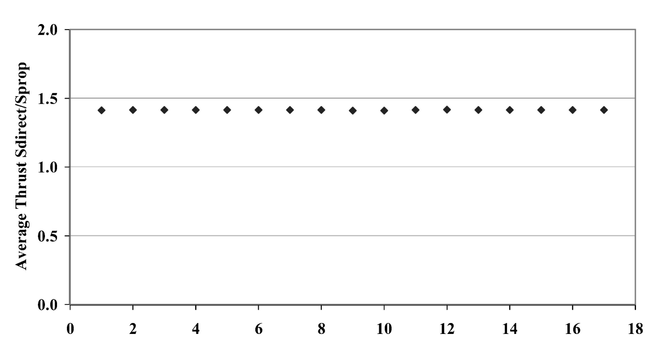

Fig. 4 Ratios of random standard uncertainties for average thrust.

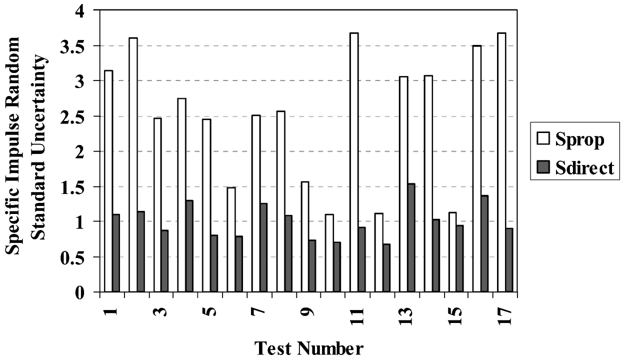

Fig. 5 Random standard uncertainties for specific impulse.

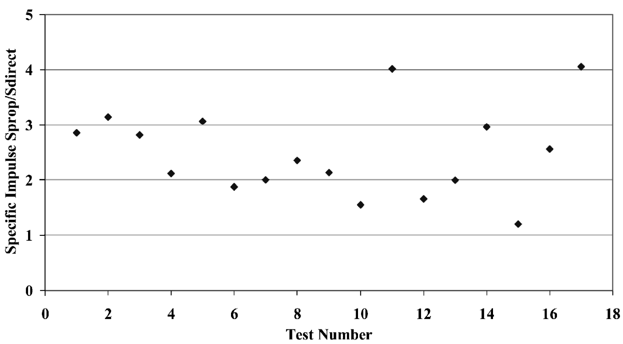

Fig. 6 Ratios of random standard uncertainties for specific impulse.

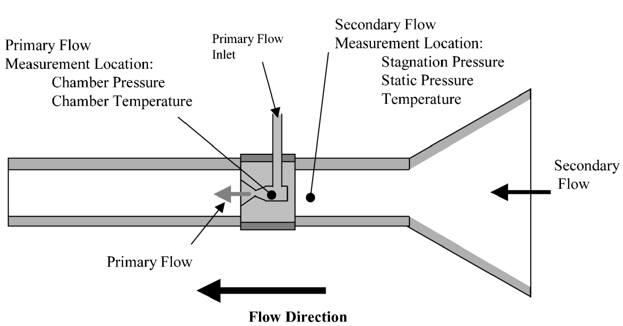

Fig. 7 Diagram of the cold-flow ejector.

The random standard uncertainties calculated for specific impulse in the 17 tests are shown in Figs. 5 and 6, and the behavior is the opposite of that observed for average thrust. For specific impulse, in each test $(S_I)_{\text{prop}}$ is greater than $(S_I)_{\text{direct}}$. Considering $(S_I)_{\text{direct}}$ to be the "correct" estimate, the random standard uncertainty $(S_I)_{\text{prop}}$ calculated using Eq. (8) is too large: by up to a factor of four in two of the tests. This means that the combination of the (neglected) correlation terms involving F_1, F_2, Q_{ox}, and Q_f is negative, leading to estimates that are too large when Eq. (8) is used.

Cold-Flow Ejector Tests

In the laboratory-scale cold-flow ejector experiment, compressed air is driven through a rocket nozzle embedded in a strut to produce a primary flow stream in an open ended duct as shown in Fig. 7. The action of the air exhausting from the nozzle into the duct induces a secondary flow through the upstream open end of the duct. Measurements of the secondary flow stagnation pressure and static pressure upstream of the strut, the inlet temperature of the secondary flow, and the rocket chamber pressure and temperature were made as indicated on the schematic. The experimental results discussed here are the mass flow rate of the induced secondary flow and the suction ratio.

The secondary mass flow rate $\dot{m}_s$ was determined assuming isentropic, incompressible, uniform flow of an ideal gas in the duct. The data-reduction equation was

$$\dot{m}_s = P_1 A_{\text{duct}} \left(P_{0_1}/P_1\right)^{(\gamma-1)/2\gamma} \times \sqrt{[2\gamma/RT_0(\gamma-1)]\left[\left(P_{0_1}/P_1\right)^{(\gamma-1)/\gamma} - 1\right]} \tag{14}$$

where stagnation pressure P_{01} and static pressure P_1 were measured upstream of the strut as indicated in Fig. 7, and the stagnation temperature T_0 was measured at ambient conditions outside the duct.

The suction ratio for the ejector is defined as the ratio of the secondary mass flow rate to the primary mass flow rate

$$\omega = \dot{m}_s/\dot{m}_p \tag{15}$$

and the primary mass flow rate $\dot{m}_p$ was determined from

$$\dot{m}_p = (P_c/T_c)A_t\sqrt{(\gamma/R)[2/(\gamma+1)]^{(\gamma+1)/(\gamma-1)}} \tag{16}$$

where P_c is the chamber pressure of the rocket nozzle, T_c is the chamber temperature, and A_t is the nozzle throat area. The suction ratio was thus determined from the data-reduction equation

$$\omega = P_1 A_{\text{duct}}\left(\frac{P_{0_1}}{P_1}\right)^{(\gamma-1)/2\gamma}\sqrt{\left[\frac{2\gamma}{RT_0(\gamma-1)}\right]\left[\left(\frac{P_{0_1}}{P_1}\right)^{(\gamma-1)/\gamma} - 1\right]} \Bigg/ \frac{P_c}{T_c}A_t\sqrt{\frac{\gamma}{R}\left(\frac{2}{\gamma+1}\right)^{(\gamma+1)/(\gamma-1)}} \tag{17}$$

The independent, controlled variable in the experiment was the chamber pressure P_c, and the inevitable variations with time in this controlled pressure caused corresponding variations in P_{01} and P_1 upstream of the strut. The variations in the three pressures were therefore correlated, and this effect is not accounted for when Eq. (8) is used to calculate values of $(S_{\dot{m}_s})_{\text{prop}}$ and $(S_\omega)_{\text{prop}}$.

Results from eight tests are considered here, with the set value of chamber pressure increasing with test number. The values of random standard uncertainty calculated using the propagation and direct approaches are shown in Fig. 8 for secondary mass flow rate and in Fig. 9 for suction ratio. In all tests, $(S_{\dot{m}_s})_{\text{prop}} > (S_{\dot{m}_s})_{\text{direct}}$

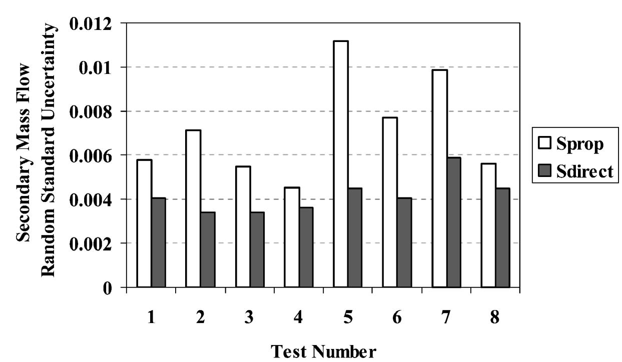

Fig. 8 Random standard uncertainties for secondary (induced) mass flow rate.

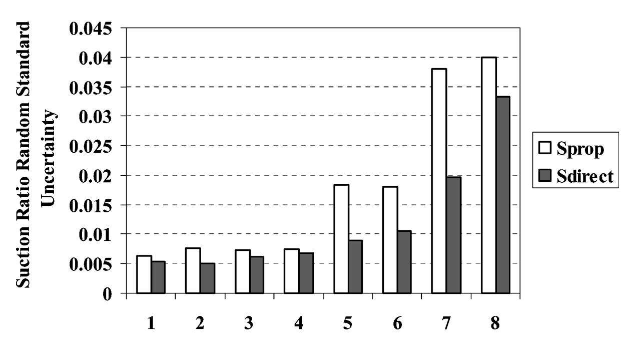

Fig. 9 Random standard uncertainties for suction ratio.

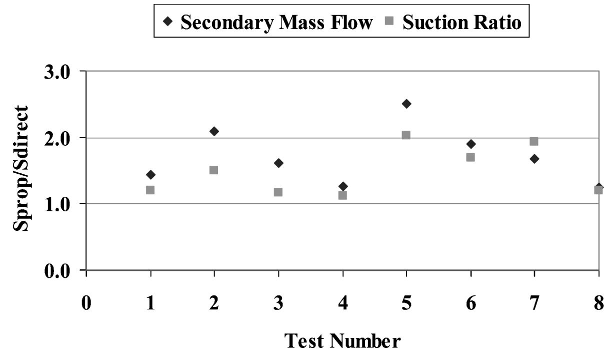

Fig. 10 Ratios of random standard uncertainties from ejector tests.

and $(S_\omega)_{\text{prop}} > (S_\omega)_{\text{direct}}$, indicating that if correlation terms were included in Eq. (8) their combination would result in a negative contribution that would decrease the $(S_{\dot{m}_s})_{\text{prop}}$ and $(S_\omega)_{\text{prop}}$ values calculated. As shown in Fig. 10, for these tests the estimates of random standard uncertainty could be as much as 2 to $2\frac{1}{2}$ times too large if the traditional propagation approach is used instead of the direct approach of Eq. (9).

IV. Conclusions

In steady-state tests conducted in a tightly controlled calibration-laboratory-type situation, variations in successive measurements of variables can be caused predominately by random measurement errors so that the distributions of the measurements of different variables are independent of one another. In much engineering testing of complex systems, however, conditions cannot be as tightly controlled, and the variations in successive measurements of the variables are caused both by unsteadiness and by measurement errors. In such cases, the distributions of the measurements of different variables are not independent of one another, that is, the variables drift in unison.

As illustrated in the results of the two test programs discussed in this paper, calculation of the random standard uncertainty (the standard deviation) of a result using the traditional propagation equation can produce an estimate many times too large or too small because the correlations among the variations in the measured variables are not taken into account. In contrast, when the random standard uncertainty is calculated directly from a sample of multiple results the effects of the correlations are accounted for implicitly.

In the direct method, if the variables are measured at N times during the test period a value of the result is calculated at each of those times. This produces a sample of N values of the result, from which a mean result and a random standard uncertainty (standard deviation) can be determined. In the traditional propagation method, for each variable a mean and standard deviation are calculated from its N values; the mean values of all of the variables are put into the data-reduction equation to produce a mean result; and the standard deviations of all of the variables are put into the propagation equation to produce a random standard uncertainty of the result.

The direct approach has not commonly been used in single test situations because unfortunately it has been presented in the past with nomenclature that caused a mindset that experimental results from multiple tests were necessary for its use. Clearly, the direct and the propagation approaches use exactly the same information and produce two estimates of the mean result and two estimates of the random standard uncertainty. It seems reasonable to calculate random standard uncertainty both ways in any steady-state test. A significant difference between the two estimates can be an indication of correlated variations among the measurements of the variables, and the direct estimate should be considered the more "correct" because it implicitly includes the correlation effects.

It is becoming increasingly common for uncertainty analyses to be performed using direct Monte Carlo simulations. In that approach, the random error in a measurement of a variable is generally taken from a normal distribution that is independent of the distributions of the random errors in the other measured variables. This does not model the effects of an unsteady (varying) steady state discussed in this paper. One way to model the unsteadiness effect is to consider successive iterations as the passage of time, specify an error source that varies with iteration number, and in a given iteration use that source to provide a single error value that is scaled appropriately for each variable's measurement. An equivalent way is to specify true values of the variables that vary with iteration in a correlated fashion.

References

[1]Coleman, H. W., and Steele, W. G., *Experimentation and Uncertainty Analysis for Engineers*, 2nd ed., Wiley, New York, 1999.

[2]*Guide to the Expression of Uncertainty in Measurement*, International Organization for Standardization, Geneva, 1993 [corrected and reprinted, 1995].

[3]"Test Uncertainty," American National Standards Inst./American Society of Mechanical Engineers, PTC 19.1-1998, New York, 1998.

[4]"Measurement Uncertainty," American National Standards Inst./ American Society of Mechanical Engineers, PTC 19.1-1985, New York, 1985.

[5]"Assessment of Wind Tunnel Data Uncertainty," AIAA Standard S-071-1995, AIAA, Washington, DC, 1995.

[6]"Assessment of Experimental Uncertainty with Application to Wind Tunnel Testing," AIAA Standard S-071A-1999, AIAA, Reston, VA, 1999.

[7]Hudson, S. T., Bordelon, W. J., and Coleman, H. W., "Effect of Correlated Precision Errors on the Uncertainty of a Subsonic Venturi Calibration," *AIAA Journal*, Vol. 34, No. 9, 1996, pp. 1862–1867.

J. Gore
Associate Editor