Flow Measurement Proficiency Testing for Key Comparisons of Flow Standards among National Measurement Institutes and for Establishing Traceability to National Flow Standards

G. E. Mattingly
Chairman, Working Group for Fluid Flow
Consultative Committee for Mass and Related Quantities (CCM)
Committee Internationale des Poids et Mesures (CIPM)
and
Deputy Chief,
Process Measurements Division
Chemical Science and Technology Laboratory
National Institute of Standards and Technology (NIST)
Gaithersburg, MD
USA

Abstract

To eliminate measurement-based trade barriers to international commerce, Key Comparisons (KCs) are now being planned in the newly formed Working Group for Fluid Flow (WGFF) in the CIPM to quantify the equivalency of the flow standards maintained in the world's National Measurement Institutes (NMIs). This equivalency will establish the "horizontal comparability" of flow standards across NMIs.

To validate the accuracies of flow calibration laboratories, flow traceability can and should be established and maintained to the flow standards maintained in the country's NMI. This "vertical traceability" quantifies the validity of flow measurement capabilities within the country's metrological pyramid with the NMI at the peak of the pyramid.

The purpose of this paper is to describe flow measurement proficiency testing techniques and their potential for comprising the highly precise KCs needed to quantify the equivalency of flow standards among NMIs as well as for disseminating flow measurement traceability from the country's NMI. Examples are given of recent proficiency testing programs that involved NIST for specific fluids and flow conditions. These demonstrate the unique levels of precision that can be attained with appropriate choices of test conditions, transfer standards, test protocols, and data analyses.

Conclusions are drawn that when proficiency tests have been used both to quantify the "horizontal comparability" i.e., the equivalency of flow standards among the world's NMIs and the "vertical traceability" to establish traceability to the country's NMI, a comprehensive hierarchy in flow measurement will be established. In it, the uncertainties quoted for flow measurements will be quantified, verified, understood, and therefore assured for all involved.

Introduction

Accurate flow measurements are needed for many reasons and for wide varieties of fluid and flow conditions, for example:

(1) the custody transfer of scarce natural resources and valuable fluid products,
(2) the measurement, control, and optimization of the productivity of continuous industrial processes to produce valuable products, and
(3) the monitoring of the environment for public health and safety.

Additionally, accurate results in these areas are needed not only within national boundaries but also across national borders. To assure the accuracy of flow meters, flow standards are established and maintained in National Measurement Institutes (NMIs). Traceability to these standards validates flow measurements within the nation's borders. To assure flow accuracies, internationally, the equivalency of the flow standards among the world's NMIs needs to be established and maintained. When this equivalency or "horizontal comparability" is produced among the NMIs and when the "vertical traceability" is produced to an NMI, measurement-based barriers to trade should be effectively reduced or eliminated.

The Committee Internationale des Poids et Mesures (CIPM) created a Working Group for Fluid Flow (WGFF) in its Consultative Committee for Mass and Related Quantities (CCM). The objective of the WGFF[1] is to conduct the Key Comparisons (KCs) needed to quantify the equivalency of the flow measurement and related standards in the NMIs to eliminate measurement-based trade barriers in six measurement areas: water flow, hydrocarbon liquid flow, low pressure air flow, high pressure gas flow, air speed, and volume. The WGFF is comprised of the NMI staff (or the delegate of the NMI) responsible for the national standards maintained in these areas by their respective nations.

Conventionally, national flow measurement reference standards are primary standards in that they are not calibrated using another flow standard; they are established, specified, and used according to physical and metrological principles [1-5][2] . These standards exist for appropriate conditions dictated by national needs; they cover wide ranges of fluid and flow conditions; and they serve wide varieties of purposes. Their capabilities are the basis from which domestic flow measurement traceability is disseminated in the country. This traceability can be established in several ways.

Unfortunately, a commonly practiced method of establishing traceability for a flow laboratory is done by establishing only the traceability of the basic component measurement systems that comprise the elements of the calibration facility with the assumption that these function as characterized in an assumed model for the whole calibration procedure. Invariably, this approach to traceability realization overlooks critical aspects of the process, and thus leads to deceptive descriptions of the quality of the measurements.

When flow traceability is established using proficiency testing techniques, the entire measurement process is involved: the component measurement systems, the routine operating conditions, the details associated with the meter installation, the personnel performing their normal procedures, and the software and calculation methods that generate the lab's final product which is the resulting calibration data. Since this method of establishing traceability involves the whole, dynamic calibration process as it is typically used, it leads to a more realistic description of the quality of the measurements produced. Similarly, when the qualities of standards maintained

[1] See: www.bipm.fr
[2] Square bracketed integers refer to References given below.

in National Measurement Institutes are compared using proficiency testing techniques, highly realistic descriptions of the capabilities of these standards can be obtained. It is these descriptions that are needed to quantify the equivalencies of national flow standards.

The purpose of this paper is to describe flow measurement proficiency testing techniques and their potential for quantifying the precision and accuracy of KCs that quantify the equivalency of flow standards among NMIs as well as for disseminating flow measurement traceability to laboratories from an NMI. Examples are given of proficiency testing programs that involved NIST for specific fluids and flow conditions. These demonstrate the unique levels of precision that can be attained via the use of proficiency testing techniques to compare fluid flow and related standards.

Flow Standards

Conventional designs for national flow measurement standards and for other primary flow meter calibration facilities usually consist of timed collections of the fluid flowing through the meter under test. These designs are made up of a number of component measurement systems. These include gravimetric or volumetric collection systems, timing techniques, temperature and pressure instrumentation, one or more databases to obtain needed physical properties for the fluids involved, and data acquisition and calculation procedures. Other critical features are the actual steadiness of the conditions involved, the influences of the personnel controlling the test, the testing environment in which the flow meter is being tested, etc. Therefore, the assessment and quantification of the complete flow measurement performance of such designs are intrinsically more involved than simply specifying the uncertainty of the basic assessment of the amount of fluid collected or stating the accuracy of the timing device.

For example, static gravimetric flow meter calibration capabilities used to determine the volumetric flow rate through a meter require accurate assessment of the true mass of the fluid being weighed. This entails accurate determination of the density of the fluid flowing through the meter under test and the buoyancy effects in the weighing environment. It also requires an accurate timing device and the characteristics of the diverter mechanisms with pertinent corrections needed to associate the proper time interval with the true mass of the collected fluid. Conventionally, this result is equated to the averaged mass rate of fluid flowing through the meter during the collection interval. This assumes that the mass of fluid in the connecting piping between the meter and collection tank did not change during the collection interval; this assumption can hinge critically on the decision by the facility operator that the flow had attained steady state conditions, dynamically and thermodynamically. Finally, the volumetric flow rate through the meter can be computed from the collected fluid mass by determining the fluid density through the meter during the collection interval. This density measurement requires pertinent temperature and pressure instrumentation so that a satisfactory density can be computed using an appropriate constitutive relationship.

Throughout the above procedure, the assumption is made that the meter under test experiences the same environmental conditions of temperature, pressure, inlet pipe flow profile, pipe vibration, etc. that will prevail where the meter is to be used or where it will be next tested. When all of these assumptions are valid, the calibration results can give a meter user an accurate characterization of meter performance that can be expected where it will be used. Similarly, where the meter is being used as a transfer standard, the results obtained can be expected to serve as a basis for comparing measurement performance among laboratories. It therefore seems apparent that assessing laboratory flow measurement performance involves considerably more than just the accuracy of the weigh system and the timing device. The following example will illustrate.

In Figure 1, a flow meter calibration arrangement is sketched with components labeled to facilitate using conventional control volume techniques to analyze its performance. This analysis will use conservation of mass principles applied to an arbitrary, stationary control volume, V, which contains all of the pipe and the flow determination system downstream of the meter under test. The control surface, S, surrounds this control volume. The conservation of mass equation is written:

$$0 = \frac{\partial}{\partial t}\int_V \rho\, dV \; - \int_S \rho \vec{v} \cdot \vec{n}\, dS \;\; , \qquad (1)$$

where ρ is the density and ∂/∂t is the partial derivative with time. The quantity $\vec{v}$ is the vector velocity of the fluid and $\vec{n} dS$ is the vectorial surface element of the control surface with positive direction taken inward and normal to the surface element of *S*. Applying Equation (1) to the facility shown in Figure 1 where letters refer to stationary sub-elements of the control volume and integers refer to sub-elements of the control surface gives:

$$\dot{M} = \int \rho_1 \upsilon_{1n} dS_1 = \frac{\partial m_c}{\partial t} + \int_{S_4} \rho_4 \upsilon_{4n} dS_4 + \int_{V_a} \frac{\partial \rho_a}{\partial t} dV_a + \int_{V_b} \frac{\partial \rho_b}{\partial t} dV_b \;\; , \qquad (2)$$

The mass flow rate, $\dot{M}$, through the 1 surface is that into the meter and $\partial m_c / \partial t$ is the rate of fluid mass collected in volume c. Additional subscripts n refer to velocity components normal to the numbered surfaces. The term $\int_{S_4} \rho_4 \upsilon_{4n} dS_4$ the rate of mass flow into volume 4; this term could represent leakage in a system in which case υ_{4n} would be negative. The two volume integrals completing the right side of the equation represent unsteady mass accumulation effects in the respective volumes and are generally assumed to be zero because calibrations are done under steady state conditions. When Figure 1 represents a timed-volumetric system, $\partial m_c / \partial t$, would represent any unsteadiness in the mass in the connecting piping or chamber, c, and the term $\int_{S_4} \rho_4 \upsilon_{4n} dS_4$ would be the timed-volumetric result.

For ideal, timed-gravimetric systems, the three integrals on the right hand side of equation (2) are zero. The performance levels for such a system can be assessed by analyzing and quantifying the different types of uncertainties involved. For example, a static gravimetric facility for calibrating a volumetric flow meter can be assessed by simplifying equation (2) and writing:

$$\dot{V} = \frac{M_n}{\rho t} \qquad (3)$$

where, in compatible units, $\dot{V}$ is the volumetric flow rate through the meter being calibrated, M_n is the net mass collected, i.e., the difference between the gross mass collected and the tare mass of the collection tank; ρ, is the density of the fluid in the meter during the collection time; and, *t,* is the collection time. Based on this model, the statistical uncertainty that can be expected in $\dot{V}$ can be assessed from the uncertainties expected in the quantities M_n, t, and ρ. Using the root-sum-of-squares (RSS) technique to combine the uncertainties of these independent component measurements, yields

$$\frac{\Delta \dot{V}}{\dot{V}} \cong \left[(\frac{\Delta M_n}{M_n})^2 + (\frac{\Delta t}{t})^2 + (\frac{\Delta \rho}{\rho})^2 \right]^{1/2} \qquad (4)$$

where the numerators can be considered as the imprecisions, i.e., standard deviations of replicated measurements and the denominators as the respective means of these measurements, [18]. Consequently, these fractions can, when multiplied by 100, be considered as "percentages of readings". Equation (4) can be used to guide calibration procedures regarding the amount of fluid, M_n, to collect or the size of the collection interval, t, to obtain satisfactory levels of imprecision for $\dot{V}$, as follows. The bound on the uncertainty in $\dot{V}$ expressed in Equation (4) can, within the limits of the specific facility, be partially controlled by the selection of the amount of fluid collected and the collection time. Using large values for both of these, the first two terms on the right side of Equation (4) can, depending upon the imprecision of these measurements, be arranged to be satisfactorily small. Therefore, taken together with the density term, which cannot be manipulated in the same way, the three terms can produce reliable estimates of uncertainty for $\dot{V}$.

These implications obtained from Equations (3) and (4) are based on a number of important assumptions. The first is that Equation (3) is the proper model for the respective calibration facility and the test conditions. Secondly, an adequate database exists to form and quantify the component uncertainties. Thirdly, no other factors are involved in the calibration process.

The assessment of measurement uncertainties at NIST is performed according to NIST-IR 1297, [5]. For flow rate measurements, the standard error of the mean of sets of measurements replicated to quantify the reproducibility of the process, expressed as a percentage of the mean, is defined as a Type A uncertainty; i.e., the type determined by statistical procedures. Type B uncertainties are defined as those determined by means other than statistical; it is this type of uncertainty that proficiency testing, as described below, can determine, experimentally. At NIST, Type A and Type B uncertainties at the 68% confidence level, u_A and u_B, respectively, are combined using the RSS method to produce the combined uncertainty, U_C, with a 68% confidence level. This combined uncertainty is multiplied by a coverage factor, *k,* to produce the expanded uncertainty, *U*. In equation form:

$$U = kU_C = k\sqrt{u_A^2 + u_B^2} \qquad (5)$$

where *k* is taken to be 2, for a 95% confidence level. Proficiency testing among the NMIs can generate the data base needed to validate estimates for u_B and quantify the equivalence or the "horizontal comparability" of these flow standards. Proficiency testing among flow laboratories and the NMI can produce estimates for u_B which then quantify the lab's "vertical traceability" to the NMI.

Traceability

Traceability is a concept that can be used to validate the estimate of u_B. Traceability [6 and 18] is the property of a measurement, either physical or chemical, or the value of a standard whereby it can be related to stated references, usually national or international standards, through an unbroken chain of comparisons all having stated uncertainties. It is noted that traceability only exists when scientifically rigorous evidence is produced on a continuing basis to show that the whole measurement process is producing documented results for which the total measurement uncertainty is quantified.

Several strategies exist for establishing traceability for laboratory flow measurement results. For dynamic measurements such as fluid flow rate which are comprised of a number of component measurements, traceability establishment is frequently attempted by establishing traceability for each of the component measurement systems separately. Such attempts can fall far short of their intended goal because all of the components of the measurement process may not be included. As a consequence of this traceability-through-the-components technique, low levels of uncertainty can result which give the deceptive conclusion that the dynamic composite measurement is perhaps better than the national standard for the same measurement. Establishing traceability using proficiency testing techniques can remedy this situation.

Proficiency testing which uses highly reliable flow transfer standards in the laboratory's typical, dynamic, measurement procedures can assess and quantify laboratory flow rate measurement performance relative to referenced standards. The results, i.e., "scientifically rigorous evidence" produced on "a continuing basis" i.e., on appropriate schedules, "to show the whole measurement process" produce data "for which the total uncertainty is quantified", relative to the referenced standard. As such, proficiency testing can fulfill the above definition's four critical elements.

When equivalency is needed among groups of laboratories, i.e., the National Metrology Institutes (NMIs), proficiency testing can be used to quantify this characteristic. The Key Comparisons that are being planned within the Consultative Committees of the CIPM to establish these equivalencies should be based on these techniques. When these are done, the uncertainties quoted by NMIs will provide the basis for the acceptance of flow measurement results across national borders.

The NIST flow measurement capabilities have uncertainties quoted as tabulated:

NIST Fluid Flow Rate Measurement Uncertainties

Fluids	**Flow**	**Type**	**Type**	**Expanded**	**Notes**
Liquids	**Range**	**A**	**B**		
a) Water	0-38 cmm[1]	±0.03%	±0.05%	±0.12 %	
b) Hydrocarbon Liquids	0-1.5 cmm[1]	±0.03%	±0.05%	±0.12%	
Gases					
a) Dry Air	0-85 scmm[2]	±0.05%	±0.08%	±0.19%	**See [11]**
b) Nitrogen	0-1 slm[3]	±0.02%	±0.05%	±0.10%	**See [12]**
c) Simulated Auto Exhaust	0-3 scmm[2]	±0.14%	±0.48%	±1.00%	**See [13]**

[1] cubic meters per minute
[2] standard cubic meters per minute; standard means 0 degree C and atmospheric pressure.
[3] standard liter per minute; standard means 0 degree C and atmospheric pressure

The above listed values of Type B uncertainty are estimated using a range of non-statistical techniques. These include using different techniques to make the same measurements or using similar techniques but different systems. The resulting estimates have been made using limited international comparisons

When these CIPM KCs are performed for wider ranges of fluid and flow conditions using highly precise proficiency testing techniques, as described below, confidence in the equivalencies will be established for the participant NMIs. These will be prominently stated in appropriate databases and publicly disseminated to produce global acceptance of national flow measurement standards.

Proficiency Testing

Proficiency testing can be arranged using a variety of transfer standards and ranges of fluid and flow conditions. These standards and conditions should be selected to span pertinent capabilities so that the proficiency testing is efficient and effective and so that high confidence can be placed in the results. Specifications of fluid and flow conditions using dimensionless parameters, such as Reynolds numbers with close tolerances applied to the values tested, further improves the precision of this type of testing by closely reproducing the same parametric effects in the different dimensional fluid and flow conditions.

Transfer standards can be selected to consist of single flow meters; fluid and flow conditions can be chosen to attempt to produce high levels of measurement reproducibility, [2]. However, improved reproducibility and higher levels of assured stability can be obtained using transfer standards composed of tandem meter arrangements, [14].

Figure 2 sketches proficiency testing using a transfer standard composed of tandem meters and how these meters can be tested to generate multiple sets of highly precise, statistically independent data that assesses the proficiency of the facility performance. Figure 2 also indicates that an in-line flow monitoring and diagnostic device can be installed upstream of the tandem meters. This device, here termed an "AUFM" i.e., an Advanced Ultrasonic Flow Meter, could be a multi-chord ultrasonic flow meter which can non-intrusively assess not only any temporal variations in the test flow but also the distribution of the pipe flow entering the units being tested. Additionally, the Superior Flow Conditioner (SFC) unit placed between the tandem meters is intended to isolate the downstream meter from any anomalies that are present in the pipe flow entering the upstream meter from the laboratory's pipe arrangement. As well, the SFC isolates the downstream meter from any flow effects produced by the upstream meter. The resulting sets of data should then be able to quantify the proficiency of the facility through its ability to determine the fluid flow rate and to calibrate the selected types of flow meters used in the selected conditions. Additionally, the data can be used to diagnose inlet pipe profile characteristics and whether these include temporal features that can influence the measurement performance of the selected type of meter.

As shown in Figure 2, the test procedure consists of testing the two meters in each of the arrangements shown and using the flow determination system to calibrate each of these meters at the selected flow condition. This test procedure should be replicated so that an adequate description of the reproducibility of these measurements is obtained; this set of configurations should be done at least twice*. These meters can be the same type or they can be different. However, before each of these tests begins, the output signals from each flow meter can be compared, for example by their ratio as shown in Figure 2. This ratio of meter outputs, displayed or computed, can be used to assess flow meter integrity and thus verify that these flow meters are performing as found in previous testing. When this ratio falls within a small, pre-selected tolerance of the expected value obtained from previous testing, high confidence can be placed in the assumption that these meters have not changed since the previous testing. If this ratio should fall beyond this tolerance, several alternatives are feasible. For example, if the ratio changes monotonically in time in steady flow conditions, it would seem apparent that one of the meters is continuously deteriorating and the test can be stopped. If the ratio falls beyond the tolerance but is steady in time, the interpretation can be either that a meter has changed or that the inlet flow profile to the upstream meter is producing a variation in meter performance in excess of that observed in previous testing. Depending upon the size of the ratio deviation from the expected value, different decisions can be made. If the deviation is small, it may be concluded this is due to an inlet flow effect on the upstream meter and testing can be continued. If the deviations large, it can be concluded that one of the meters has changed or deteriorated significantly, but the testing can be continued so that it can be determined which meter has changed. If this can be determined, it may be feasible to continue the tests and results based only on the other meter until the damaged meter can be repaired or replaced and the complete test re-run.

Tandem meter testing also enables correlation of meter outputs to categorize the components of the variance in the results obtained. For example, after repeated determinations of the meter factors are performed for a specific configuration in the tandem meters and for a specific flow rate, these can be statistically analyzed. Results are obtained both for the meters and for the flow facility, as follows. The averages, standard deviations, and the

* Private Communication, Drs. J.J. Filliben, S. D. Leigh, and M. Vangel, NIST Statisticians, Aug. 2000.

correlation coefficient for the meter factors for each meter for each configuration and for each flow rate can be written:

$$\bar{k}_1 = \frac{1}{n}\sum_{i=1}^{n} k_{1i} \quad , \tag{6}$$

$$\bar{k}_2 = \frac{1}{n}\sum_{i=1}^{n} k_{2i} \tag{7}$$

$$S_1 = \left[\frac{1}{n-1}\sum_{i=1}^{n}\left(k_{1i} - \bar{k}_1\right)^2\right]^{1/2} = \sqrt{\sigma_1} \quad , \quad \text{and} \tag{8}$$

$$S_2 = \left[\frac{1}{n-1}\sum_{i=1}^{n}\left(k_{2i} - \bar{k}_2\right)^2\right]^{1/2} = \sqrt{\sigma_2} \ . \tag{9}$$

where $\bar{k}_1$ and $\bar{k}_2$ are the average of the *n* individual determinations of meter factors k_{1i} and k_{2i} of meters 1 and 2, respectively, S_1 and S_2 are the standard deviations of the meter factors and σ_1 and σ_2 are the variances of these respective values. The correlation coefficient, r_{12}, of the *n* ordered pairs of the meter factors, can be written:

$$r_{12} = \frac{\sum_{i=1}^{n}\left(k_{1i} - \bar{k}_1\right)\left(k_{2i} - \bar{k}_2\right)}{\left[\left[\sum_{i=1}^{n}\left(k_{1i} - \bar{k}_1\right)^2\right]\left[\sum_{i=1}^{n}\left(k_{2i} - \bar{k}_2\right)^2\right]\right]^{1/2}} \ . \tag{10}$$

The correlation coefficient, i.e., its absolute value, can then be interpreted as that fraction of the variance of the respective meter factors that can be attributed to that component of the measurement that is common to both meters, i.e., the flow determination system:

$$\left|r_{12}\right| = \frac{\sigma_{1s}}{\sigma_1} \quad , \tag{11}$$

where σ_{1s} is that portion of the total variance for meter 1 that can be attributed to the flow determination system. The remaining fraction of the variance can be attributed to that component of each measurement that is not common, i.e., the variation of individual meters:

$$1 - \left|r_{12}\right| = \frac{\sigma_{1m}}{\sigma_1} \quad , \tag{12}$$

where σ_{1m} is that portion of the total variance, σ_1 that can be attributed to meter 1; similar values can be computed for meter 2. As a further check on meter well-being, these individual meter variances can be assessed to apply further confidence in tandem meter testing.

The data generated using the proficiency testing arrangements of the tandem meters in each of the configurations for each specified flow as sketched in Figure 2 consists of two statistically independent sets of data: one for meters tested in the upstream position, the other for the downstream. These data can be in the form of meter factors or other flow meter indicators that are pertinent to the goals of the proficiency testing. In each of these configurations, statistical independence is based on the assumption that, because the two meters were tested at different times in the same position, each could not have affected the other. Consequently, the Youden graphical analysis of variance techniques can be applied to each of these data, [15]; this application of the Youden techniques is the same for the case where the meters are the same or for cases where the meters are different[*]. These Youden results can be used to assess and to quantify not only the performance of the participating laboratory's flow determination system but also the flow profile effects that the piping arrangement delivers to the meter test section.

Figure 3 sketches Youden plots of proficiency testing results. These data should represent the normal, routine performance of the labs calibrating these meters. Plotted results therefore should be the reproducible averages of the selected meter factors for the specified conditions or they can be selected ratios of flows determined by the facility to that measured with each meter. Each participant laboratory will have a single point on this graph; by pre-test agreement the identity of the laboratory can be listed or the results can be plotted anonymously. The analysis described below is ideally applied to data sets consisting of sufficient numbers of participants.

The analyses applied to proficiency testing results vary depending upon whether the testing is that of a Key Comparison among NMIs or that of a testing program to establish flow measurement laboratory traceability to an NMI. To analyze KC data, the approach shown here will be to weight all laboratory results equally and to assume that a consensus of all the participant laboratories should be the basis of the analysis. In this case and as sketched in Figure 3(a), horizontal and vertical lines are drawn through the median data points plotted for the 2 meters. By consensus, the intersection point of these median lines can be considered, on the basis of the data available, to produce the best estimates of the true values for the conditions tested. To analyze data intended to establish the traceability of flow measurement laboratory results to an NMI, the approach taken is to reference the NMI data by drawing the horizontal and vertical lines through the NMI results, as sketched in Figure 3(b).

In either case, these horizontal and vertical lines divide the plot into 4 quadrants and the data in these can be interpreted to assess flow laboratory performance in the following ways. Laboratories having points which lie in the northeast quadrant can conclude that their facilities produced both measurement results larger than that considered as the true value, although at this point causes for this are not clear. Similarly, laboratories having points in the southwest quadrant can conclude that their processes produced both results smaller than those considered true. Laboratories finding their points in the northwest or southeast quadrants can conclude that their results were somewhat random relative to the values considered true in that one result was higher but the other was lower than the reference values.

Youden analyses can be continued to quantify not only the random uncertainty of the specific proficiency testing procedures and results but also the systematic or Type B uncertainties associated with the participating facilities, as follows. As shown in Figures 3(a) & (b), a line of slope + 1 is drawn through the intersection point and the data is then projected perpendicular and parallel to this diagonal line. The respective projections are then used to produce standard deviation type quantities for the *m* participants, i.e.,

$$\sigma_r = \left[\frac{1}{m-1} \sum_{i=1}^{m} N_i^2 \right]^{1/2} \quad , \text{ and} \qquad (13)$$

---

[*] Private Communication, Drs. J.J. Filliben, S. D. Leigh, and M. Vangel, NIST Statisticians, Aug. 2000.

$$\sigma_s = \left[\frac{1}{m-1} \sum_{i=1}^{m} P_i^2 \right]^{1/2} \quad , \tag{14}$$

where σ_r and σ_s can be interpreted, respectively, as the random and systematic deviations of the available data, and N_i and P_i are the respective normal and parallel components of the data projected on and to the slope + 1 line. The ratio of these quantities gives the degree of circularity of the data:

$$c = {}^{\sigma_s}\!/_{\sigma_r} \tag{15}$$

When this ratio is larger than unity, the interpretation is that the elliptical data pattern shows that systematic type uncertainties exist among the measurement processes involved. The levels of these systematic-type uncertainties are indicated by the location of the data along the major axis of the elliptical data pattern.

The preferred result would be to find that the circularity of the data pattern is unity. In this case, the interpretation could be made that systematic and random aspects of the data are essentially of the same magnitude. Therefore, the participant facilities have testing proficiencies that can be quantified in terms of the radius of the circle, which can enclose the data about the reference point. The participant laboratories, in this case, could interpret the results as indicative that their measurement uncertainties are comparable to the reference within the radius of the circle.

The minor axis of the elliptical data pattern or the radius of circular data plots is essentially the randomness of the whole proficiency testing program, i.e., a "noise" level within which bias-type uncertainties or "signals" will be hidden. Consequently, there is a need for highly precise transfer standards and measurement condition controls when one considers the low levels of uncertainties claimed for NMI standards to establish “vertical traceability” to the NMI.

Typical Results

NIST has conducted several proficiency testing programs for the establishment of flow measurement traceability to the reference flow standards provided by NIST. These programs use selected types and sizes of tandem arrangements of flow meters as transfer standards; they span limited ranges of flow rate; the fluids used are specific fluids, i.e., water, a hydrocarbon liquid, and a gas, i.e., nitrogen.

Typical results obtained using a hydrocarbon liquid and turbine flow meters are shown in Figures 4 and 5. Test details can be found elsewhere, see [16]. Results plotted are the averages of the ratios of meter factor determined in each participant lab to that determined at NIST. Figures 4 (a) and (b) present results from some 19 facilities in 14 laboratories; this data was taken over a period of about 3 years. In Figures 4(a) and (b), the test conditions are for the lower of the two flows tested; the diametral Reynolds number is 2750 in these 12.5 mm diameter meters. In Figure 5, the conditions are the high flow rate for which the Reynolds number is 8250. Horizontal and vertical lines are drawn through the NIST points designated by the letter X.

Although the quantification of the random uncertainty of the results shown in Figures 4 and 5 is not done here as indicated in equations (13) and (14), these results show graphically these random uncertainities. Results shown in Figures 4 and 5 indicate that for this transfer standard, these test conditions, and the laboratory procedures that the random uncertainties of these participants are bounded at about ±0.2%, or better, for the upstream meters. For the downstream meters, the random or Type A uncertainties are about ±0.1%, or better. These values can be considered as the reproducibility for these tests for these conditions; repeatability levels for these tests and conditions in single day testing at NIST show levels of ± 0.05% for upstream meters and ±0.03% for downstream. In Figure 5, systematic or Type B uncertainty levels relative to NIST range as high as about ±0.4%. These uncertainties for specific participants are reduced when downstream results are compared to upstream results. The conclusion can be drawn that perhaps pipe flow profile effects are responsible. The higher flow rate results shown in Figure 5 indicate that reduced random uncertainties occur in the downstream position, i.e., the random uncertainties are about ± 0.05%. That reduced random uncertainty occurs at higher flow conditions is typical performance for this type of flow meter. However, the ability to quantify the reproducibility obtained over such a range of tests, labs, and time is in large measure attributable to the tandem arrangement of testing these meters and to the high levels of control that can be implemented.

Typical results obtained using a transfer standard composed of tandem critical nozzles flowing nitrogen gas at low Reynolds numbers are shown in Figures 6 and 7. Test details can be found elsewhere, see [17]. The results plotted show averaged values of flow rate ratio, i.e., that determined by the participant facility to that measured by the respective critical nozzle. These results include some 55 tests in 23 participant laboratories. The data is normalized relative to the NIST value. These results indicate random or Type A uncertainties for this type of transfer standard in these conditions are about ± 0.3 %, or less. Systematic or Type B uncertainties in this group of laboratories range up to ±8.0%. Such results can effectively guide search and repair efforts to improve laboratory flow measurement performance and traceability.

Typical results obtained using water and turbine meters are shown in Figures 8 and 9. These data were obtained for 100 mm diameter meters at diametral Reynolds numbers of 1.3×10^5 and 6.4×10^5, respectively, see [18]. Results plotted are normalized with respect to meter factors obtained in NIST testing. The results in Figure 8 indicate the random uncertainties for these conditions are about ± 0.1%; systematic uncertainties range up to about ±0.2%. Analogous results at the higher flow rate in Figure 9 show improved precision levels especially when the meter was tested in the downstream position.

Conclusions

The primary conclusion of this paper is that proficiency testing techniques should be used to conduct the Key Comparison tests that reflect the total performance of the flow measurement capabilities of the participant NMIs and establish the “horizontal comparability” or equivalence of flow standards in the NMIs.

Key Comparison tests need to be designed and used for widely ranging fluid and flow conditions to satisfactorily quantify the equivalence of the NMI flow standards. Such testing should have very high levels of reproducibility to successfully quantify the low levels of uncertainty claimed, and should be performed on appropriate schedules to maintain the equivalences achieved.

It is also concluded that proficiency testing should be performed periodically to establish and maintain flow measurement traceability for flow measurement laboratories to the pertinent national flow reference standards. Ideally, such testing also needs to be performed in a manner that quantifies the uncertainties claimed and spans the wide ranges of conditions required.

To attain the desired measurement uncertainty in proficiency testing, the design and use of transfer standards consisting of tandem flow meter arrangements and testing each meter in each position at each flow to enable the Youden graphical analysis of variance techniques to be applied, has been shown to hold high potential for Key Comparison testing. Selection of meter type, size, etc. is concluded to depend upon the requirements of the proficiency tests. It can be performed to conform to routine types of meters tested in the participant laboratories, and it can be done for typical fluids and flow conditions. Extremely low levels of random uncertainty can be obtained to assess flow laboratory claims.

It is concluded from specific proficiency tests conducted at NIST that ranges of reproducibility can be achieved depending upon the conditions involved. It is shown that improved random uncertainties occurred at the higher flows tested and for the meters tested in the downstream position; this is due to the fact that the upstream meter tends to isolate the downstream meter from the pipe flow profile variations that enter the upstream meter. In the examples given, it is concluded that when the results from a participant laboratory are acceptable for downstream meter results, but unacceptable for the upstream, the diagnosis can be made to assess the inlet pipe flow profile to the meter test section.

Finally, it is concluded that when proficiency testing has successfully been used to compare national flow reference standards, and when proficiency testing has succeeded in establishing and maintaining flow measurement traceability to pertinent national flow standards, the international flow communities will have an improved basis upon which to assess and accept flow measurement results.

References

[1] Marshall J. L. (Editor), NIST Calibration Services Users Guide, NIST Special Publication 250, Jan. 1998.

[2] Mattingly, G.E., Fluid Flowrate Metrology, in ISA Practical Guide for Flow Measurement, edited by D.W. Spitzer, Published by ISA, Research Triangle Park, NC, 1991.

[3] Mattingly, G.E., Gas Flowrate Metrology, NCSL Newsletter Vol. 29, No. 1, Jan. 1989, pp. 9-16.

[4] Benson, K.R., et al, NBS Primary Calibration Facilities for Air Flow Rate, Air Speed, and Slurry Flow, Procs. of the American Gas Association's Symposium on Flow Measurement, Crystal City, VA, 1986.

[5] Mattingly, G.E., Primary Flow Calibrations, Reference and Transfer Standards in Developments in Flow Measurement - I, edited by R.W.W. Scott, Published by Applied Science, Inc., Englewood Cliffs, NJ, 1982.

[6] Belanger, B.C., Traceability - An Evolving Concept, ASTM Standardization News, Vol. 8, No. 1, Jan. 1980, pp. 22-27.

[7] Mattingly, G.E., Dynamic Traceability of Flow Measurements, Invited Lecture, IMEKO Tokyo Flow Symposium, 1979, Published by the Society of Instrument and Control Engineers, Tokyo, Japan.

[8] Garner, E.L. and Rasberry, S.D., What's new in Traceability, Journ. of Testing, and Evaluation, Vol. 21, No. 6, Nov. 1993, pp. 505-509.

[9] ISO International Guide to the Uncertainty of Measurement, Published by the International Standards Organization, P.O. Box 56, Geneva, Switz, 1984.

[10] Taylor, B.N. and Kuyatt, C.E., Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results, NIST Tech. Note 1297, Sept. 1994.

[11] Wright, J.D. and Mattingly, G. E. NIST Calibration Services for Gas Flow Meters: Piston Prover and Bell Prover Gas Flow Facilities, NIST Special Publication 250-49, Aug. 1998.

[12] Tison, S.A., Development of a Laminar Flow Transfer Standard, Procs. Measurement Science Conference, 1996.

[13] Wright, J.D., et al, NIST's New Variable Composition, Variable Temperature Gas Flow Calibration Facility, in preparation.

[14] Mattingly, G.E. and Spencer, E.H., Steps Toward an Ideal Flow Transfer Standard, in the Procs. of the 1978 FLOMEKO Symposium , convened in Groningen, The Netherlands, 1978, published by North Holland Publ. Co., Amsterdam, NL 1979.

[15] Youden, W.J., Graphical Diagnosis of Interlaboratory Test Results, Journal of Industrial Quality Control, Vol. 15, No. 11, pp. 133-37, May 1959.

[16] Mattingly, G.E., A Round Robin Flow Measurement Testing, Program Using Hydrocarbon Liquids: Results for First Round Testing, NISTIR 88-4013, Sept. 1988.

[17] Mattingly, G.E., Round Robin Mass Flow Controller Gas Flow Calibration Test Results Using the NIST-SEMATECH-Mark 1 Gas Flow Measurement Artifact: Final Report, Tech. Transfer No. 93051631 A-Eng. SEMATECH, Austin, TX, May 1993.

[18] Mattingly, G.E., An Interlaboratory Round Robin Flowmeter Test Using Turbine Meters Flowing Water, in preparation.

Figure Captions

Figure 1 Typical Flow Meter Calibration Facility

Figure 2 Sketch or Proficiency Tests Using Tandem Meter Configurations

Figure 3 Sketches of Typical Youden Plots of Proficiency Testing Results

(a) Youden Plot Sketching Analysis for International Comparison Test Data; Dashed Lines are Drawn Through Median Values.

(b) Youden Plot Sketching Analysis of Domestic Testing to Quantify Traceability to a National Standard; Dashed Lines are Drawn Through the Values for the Referenced Standard.

Figure 4 Proficiency Testing Results for Turbine Meters Flowing Hydrocarbon Liquid for Reynolds Number of 2750: (a) Upstream Meters, and (b) Downstream Meters.

Figure 5 Proficiency Testing Results for Turbine Meters Flowing Hydrocarbon Liquid for Reynolds Number of 8250: (a) Upstream Meters, and (b) Downstream Meters.

Figure 6 Proficiency Testing Results for Critical Nozzles Flowing Nitrogen for a Tube Reynolds Number of 125: (a) Upstream Nozzles, and (b) Downstream Nozzles.

Figure 7 Proficiency Testing Results for Critical Nozzles Flowing Nitrogen for a Tube Reynolds Number of 300: (a) Upstream Nozzles, and (b) Downstream Nozzles.

Figure 8 Proficiency Testing Results for Turbine Meters- Flowing Water for Diametral Reynolds Number of 1.3×10^5: (a) Upstream Meters, and (b) Downstream Meters.

Figure 9 Proficiency Testing Results for Turbine Meters Flowing Water for Diametral Reynolds Number of 6.4 x 10^5: (a) Upstream Meters, and (b) Downstream Meters.

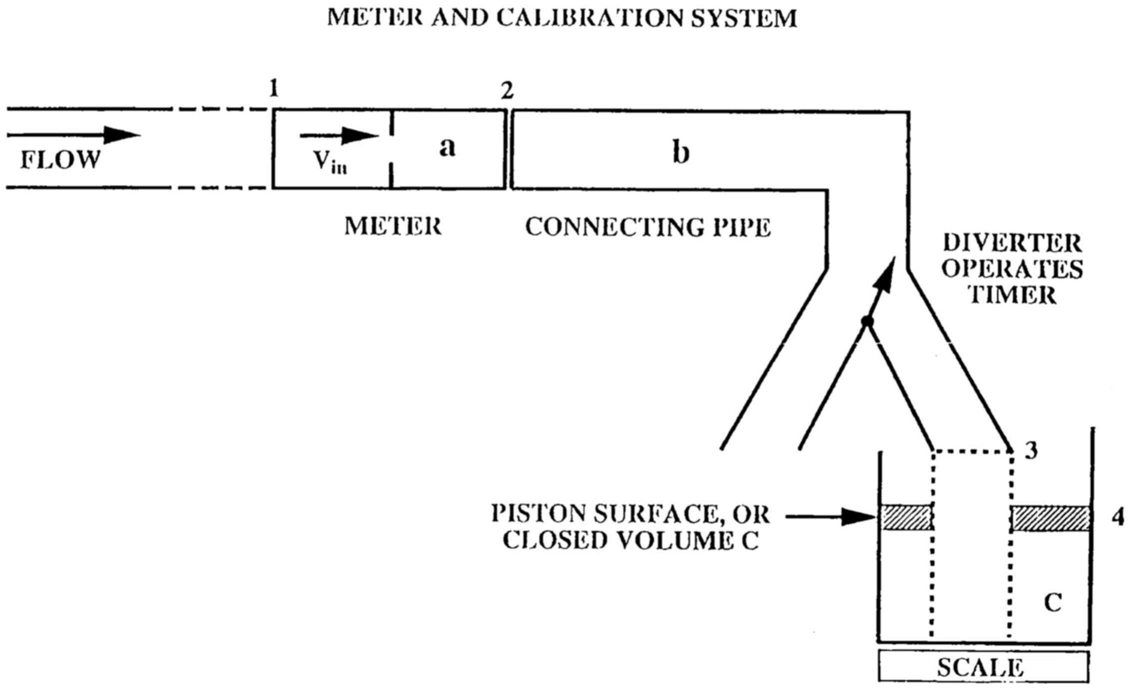


Figure 1 Typical Flow Meter Calibration Facility

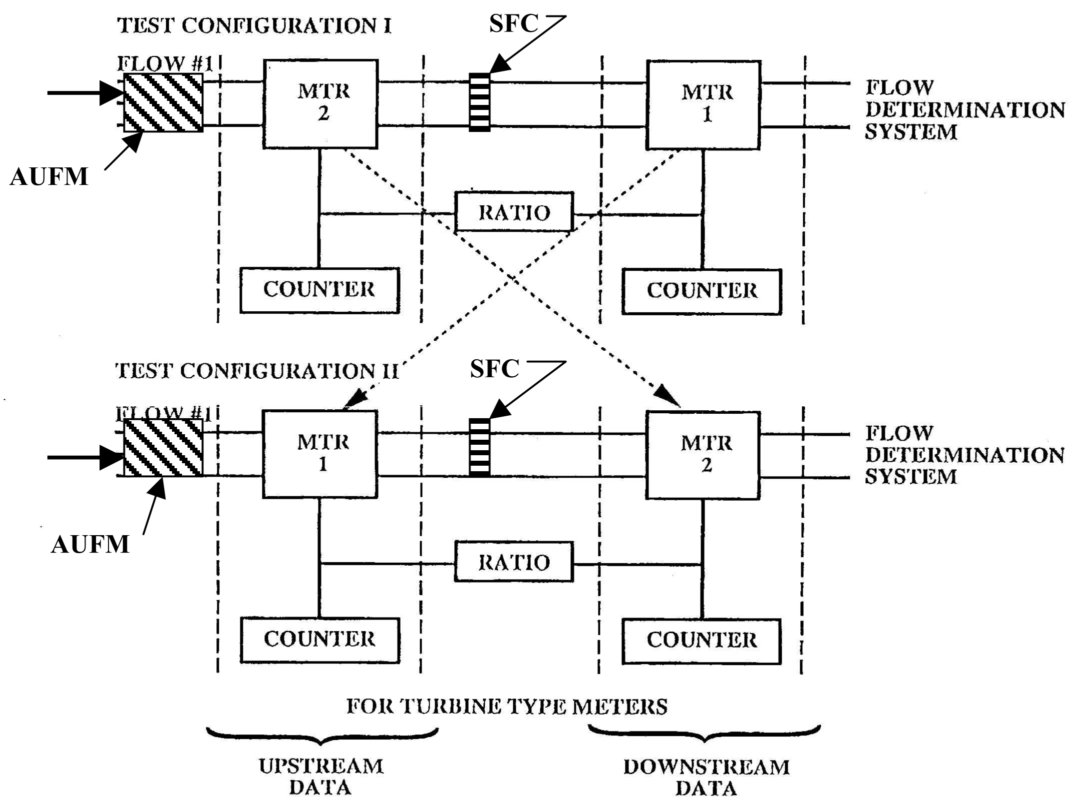
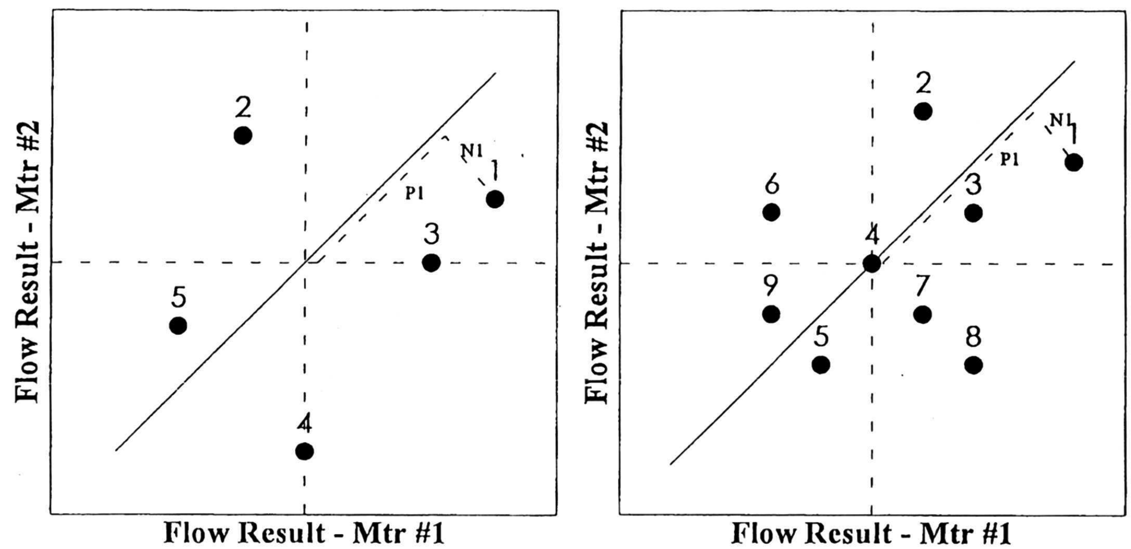


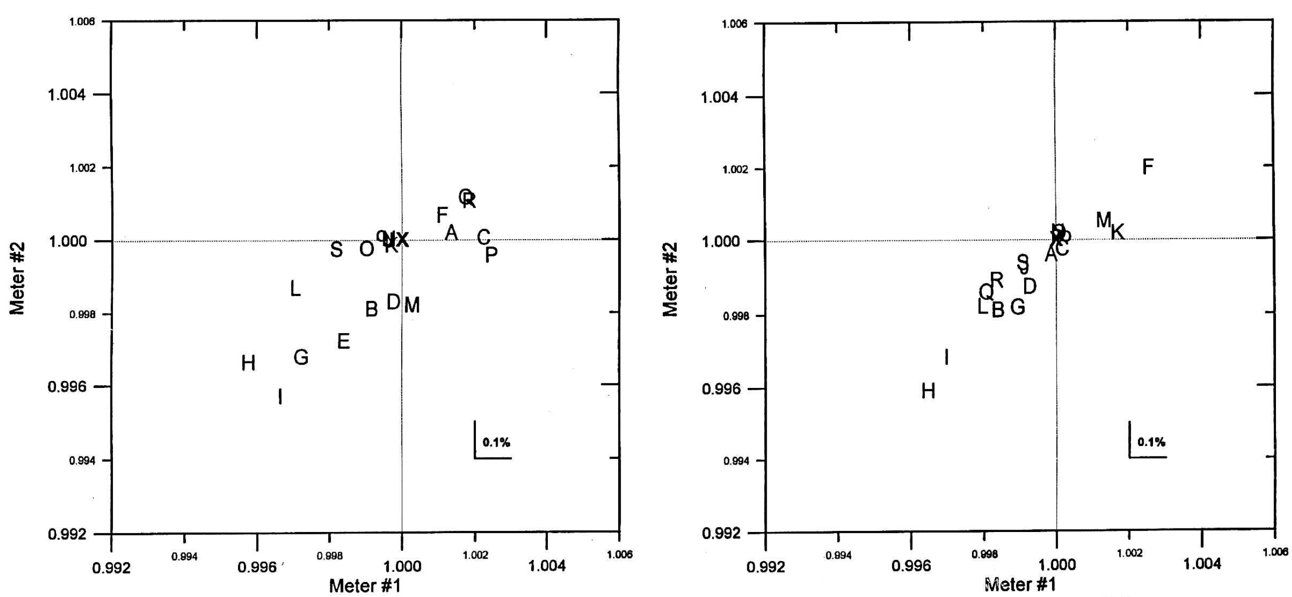
Figure 2 Sketch of Proficiency Tests Using Tandem Meter Configurations



(a) Youden Plot Sketching Analysis for International Intercomparison Data; dashed lines are drawn through median values.

(b) Youden Plot Sketching Analysis for Traceability to a National Standard; dashed lines are drawn through the values for the referenced standard.

Figure 3 Sketches of Typical Youden Plots of Proficiency Testing Results



(a) Upstream Meters

(b) Downstream Meters

Figure 4 Proficiency Testing Results for Turbine Meters Flowing Hydrocarbon Liquid for Reynolds Number of 2750: (a) Upstream Meters, and (b) Downstream Meters.

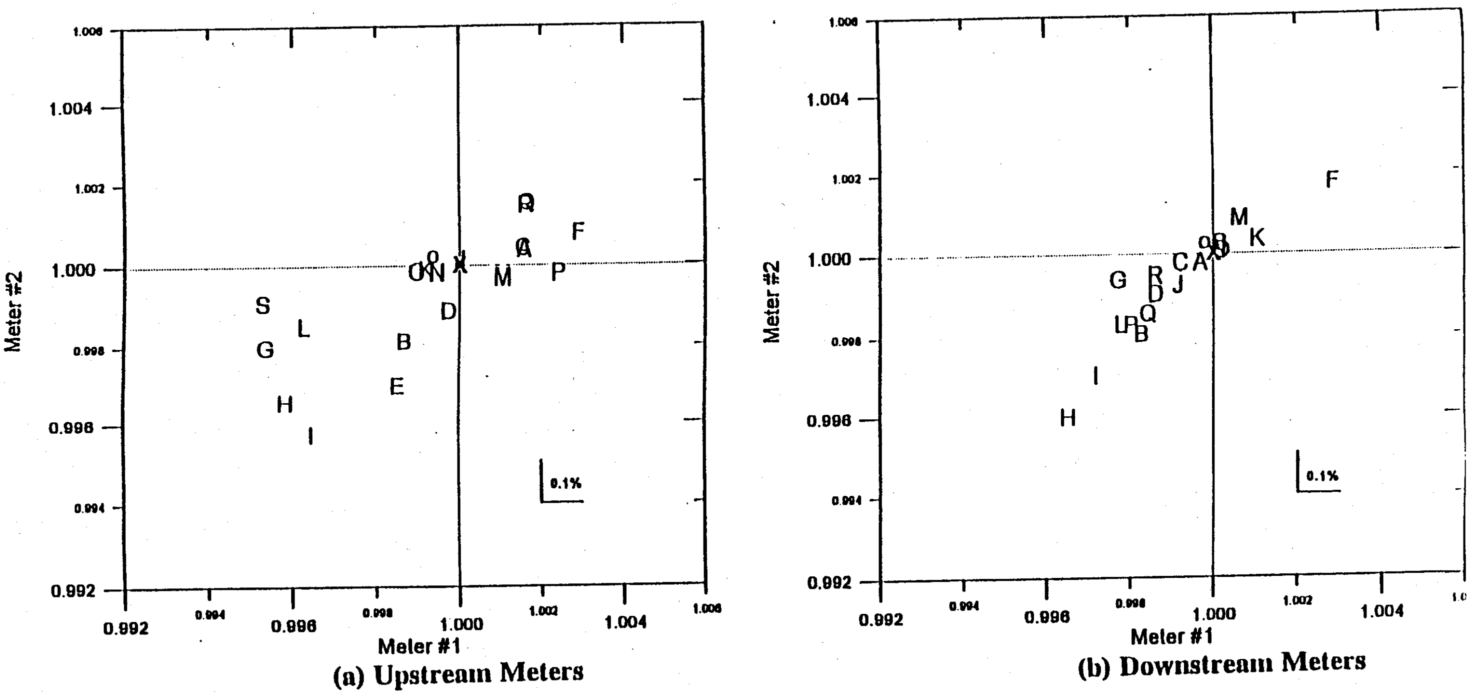


Figure 5 Proficiency Testing Results for Turbine Meters Flowing Hydrocarbon Liquid for Reynolds Number of 8250: (a) Upstream Meters, and (b) Downstream Meters.

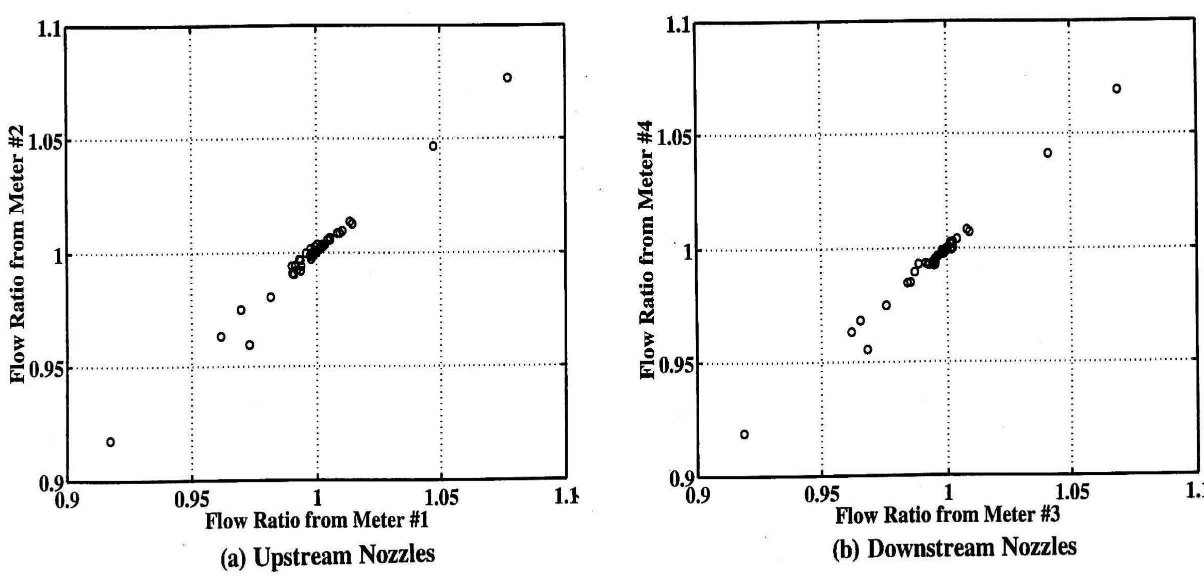
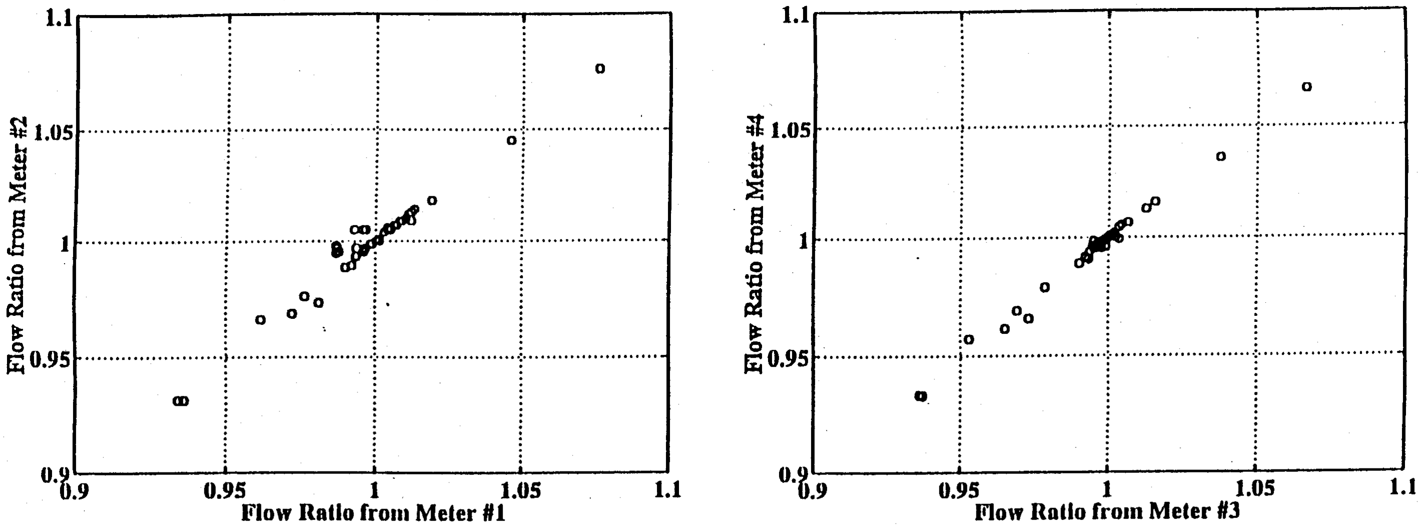


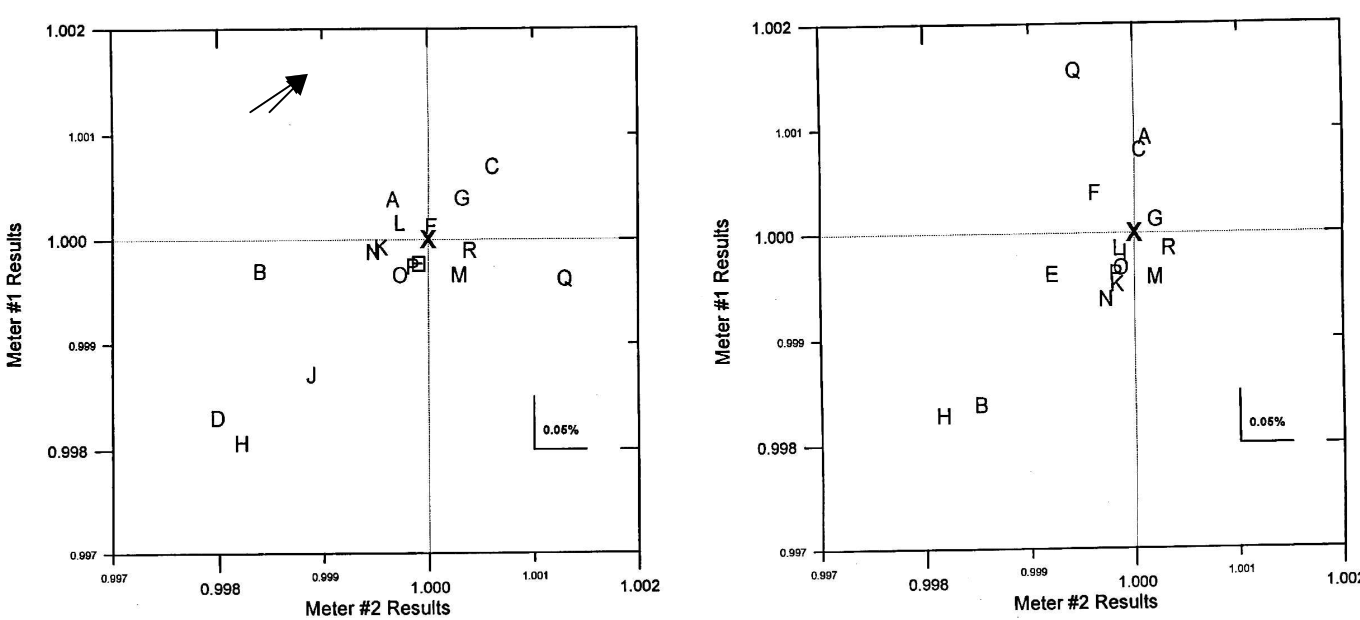
Figure 6 Proficiency Testing Results for Critical Nozzles Flowing Nitrogen for a Tube Reynolds Number of 125: (a) Upstream Nozzles, and (b) Downstream Nozzles.



(a) Upstream Nozzles

(b) Downstream Nozzles

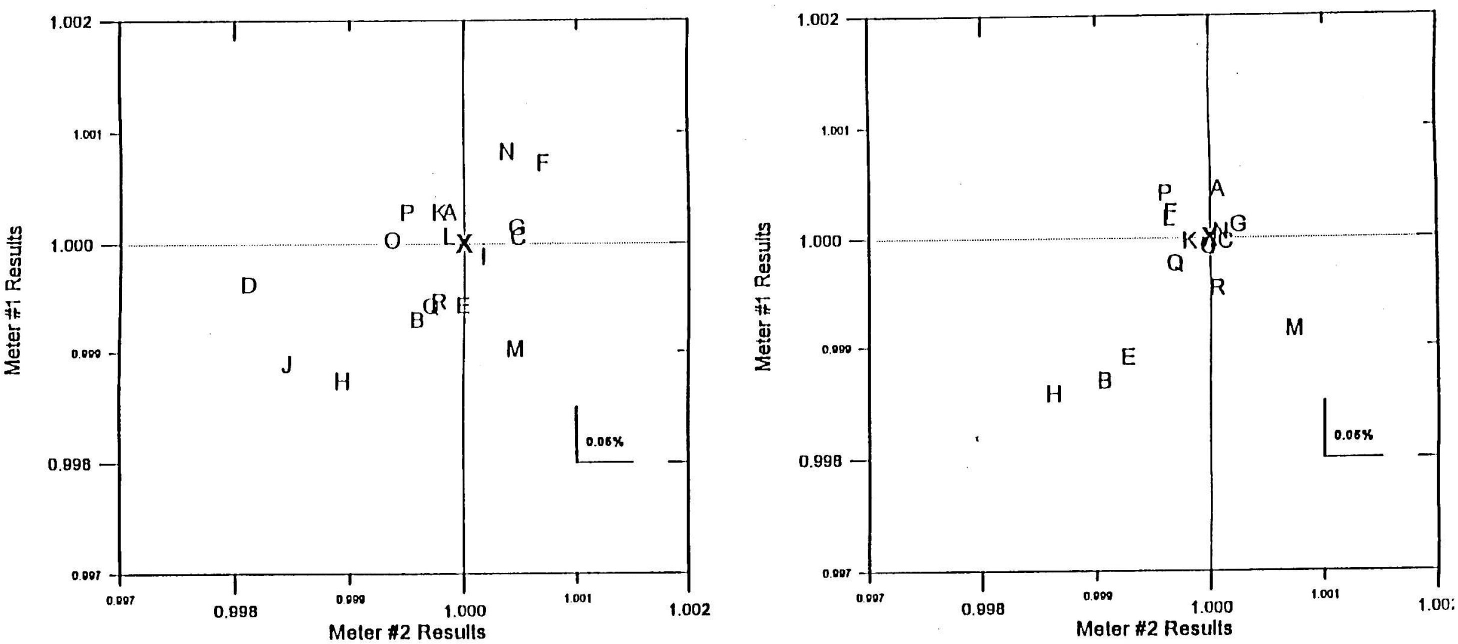
Figure 7 Proficiency Testing Results for Critical Nozzles Flowing Nitrogen for a Tube Reynolds Number of 300: (a) Upstream Nozzles, and (b) Downstream Nozzles.



(a) Upstream Meters

(b) Downstream Meters

Figure 8 Proficiency Testing Results for Turbine Meters Flowing Water for Diametral Reynolds Number of 1.3 x 10^5 : (a) Upstream Meters, and (b) Downstream Meters.



(a) Upstream Meters

(b) Downstream Meters

Figure 9 Proficiency Testing Results for Turbine Meters Flowing Water for Diametral Reynolds Number of 6.4 x 10^5 : (a) Upstream Meters, and (b) Downstream Meters.