FLOMEKO 2003
11th IMEKO TC9 Conference on Flow Measurement
Groningen, NETHERLANDS, 12 - 14 May 2003

"QUADRATIC WEIGHING", APPLIED IN HARMONIZATION, A TOOL TO COMPARE THE PERFORMANCE OF TEST-FACILITIES

UNCERTAINTY AS A YARD-STICK *) FOR THE QUALITY OF MEASUREMENT

R. van den Brink, M.P. van der Beek and I.J. Landheer - NMi VSL-Flow, Dordrecht, The Netherlands

SUMMARY
Measurement-data of laboratories participating in comparisons are often treated equivalently with respect to operational conditions and statistical impact. When analysing the comparison, this may distort some conclusions because no attention is paid to deviations caused by differences in operational conditions (e.g. pressure). A better way to perform a comparison is to transpose the sets of flow-rate data to Reynolds numbers, to make it easier to see whether or not the points are in the same Reynolds-region and thus allow for a valid comparison.
Next, not all participating laboratories contribute equally to establish a Mean (Average) Value, because the laboratories with a low uncertainty must have a greater impact on the final result than laboratories of a lesser quality. So the data-sets must be weighed and to do so, weighing-factors are needed.
Finally, a common way to analyse the comparison-results is by using Youden-plots.
This paper shows the "flow-rate to Reynolds-transformation" to compare test-data, and demonstrates which weighing-factors must be used to get the "Best Known Mean Value" with the smallest uncertainty. In the end, this method is also applied in Youden-plots.

INTRODUCTION
During the last decade, the reproducibility and stability in performance of e.g. turbine-meters have improved and with today's technology, variations in reference values at the end of the traceability chain have become detectable to the ordinary user. Normally this phenomenon would not create any interest of importance. But, gas measurements under operational conditions as used as custody transfer measurements are the base for 'traceable invoicing'. An accountant does not readily accepts uncertainty in an invoice, let alone sudden changes in the invoice-level.
It is therefore of vital importance that high-pressure gas test-facilities are aiming at reference levels that do not differ from each other. Generally there should not be any demonstrable (significant) difference between National reference values [1].

A problem at comparing high-pressure gas-flow test-facilities is that the operating conditions for flow, pressure, gas-temperature and gas-composition are not quite similar, and that the basic uncertainties expressed as CMC **) differ [2].
To overcome these disadvantages, NMi proposes a procedure in which a calculation method is given to overcome these drawbacks and pitfalls [3]. E.g. usually every participating laboratory is treated in the same way, i.e. the generated data-sets are assumed to be of equal importance or quality. However, following the standard procedure for the mean value, a considerable bias can be introduced. This bias affects conclusions with respect to the magnitude in which laboratories are differing from each other. Calculating a mean value by weighing the participating values instead of treating them identically, the value of data-analysis is enhanced.

In this paper a method of "quadratic weighing" of data-sets is introduced together with a method to make different operating conditions comparable. Methods are compared and illustrated with "Residual-plots" and "Youden-plots".

*) "Yard-stick", homage to the imperial system, turning into metric, inch by inch.
**) CMC, the highest level of calibration or measurement normally offered to clients, expressed in terms of a confidence level of 95%, sometimes referred to as Best Measurement Capability.

COMPARING GAS-FLOW MEASUREMENTS OF DIFFERENT OPERATIONAL CONDITIONS

Between the National laboratories, regular inter-comparisons are organized to investigate the relations between the observed reference values and to conclude that those reference values are (still) a valid representation of the Unit under investigation (with differences within permissible limits). For HP gas-flow measurements a practical problem of inter-comparison programmes are the operating conditions of the participating laboratories. For most laboratories they are different. Nevertheless, measurement-data of comparisons are treated equivalently with respect to those operating conditions. Key question is now, whether or not the observed measurement-data can be compared at all.
Are the measurement-data of the used transfer-meters comparable, i.e. are the meters behaving comparably at different conditions ? And, could flow-data be transposed to the Reynolds-domain to overcome this inconvenience ?
Comparing test-results in high-pressure gas-flow measurements with (proven) turbine-meters via application of Reynolds numbers is a well-established procedure. Reynolds-number is expressed as

$$Re = \frac{\rho \cdot c \cdot D}{\eta} \qquad (1)$$

with

Re	= Reynolds-number		[1]
ρ	= gas-density		[kg/m^3]
c	= gas-velocity		[m/s]
η	= dynamic gas-viscosity		[N·s/m^2]
D	= diameter of transfer-meter		[m]

Consider test-data of three different laboratories from the same two transfer-meters at different operational conditions (obviously the pressure). In figure 1, the data to be compared are plotted against the flow. Figure 2, shows the same data after the transposition to Reynolds numbers. It is the first step to have a fair comparison between the data-sets of three laboratories. Another advantage of using the Reynolds-domain is that meter-behaviour can be made visible in a more reliable way.

The transposition of flow-data from Flow-Rate to Reynolds-number is illustrated in figures 1 and 2

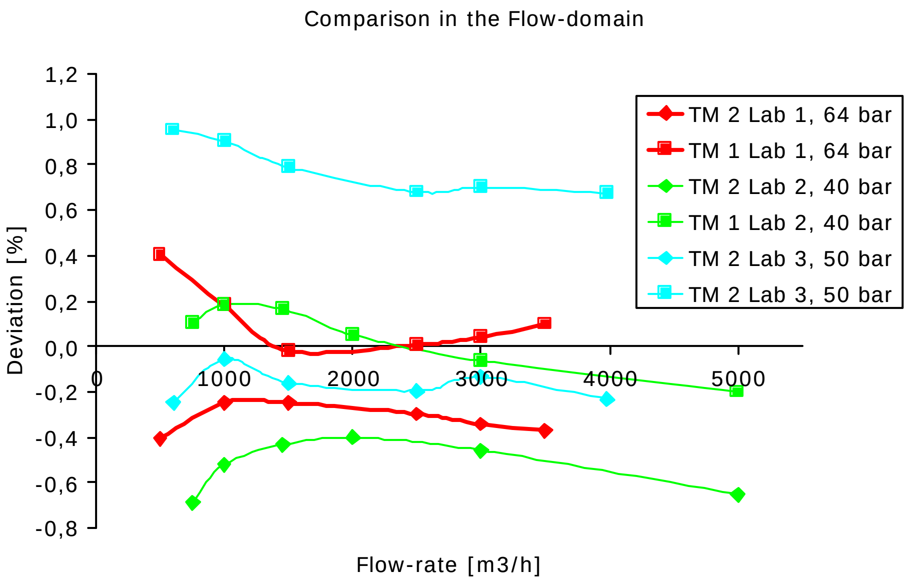

Figure 1, Comparison between three different laboratories under different operational conditions, with two turbine meters in tandem. The x-axis represents the Flow-Rate.

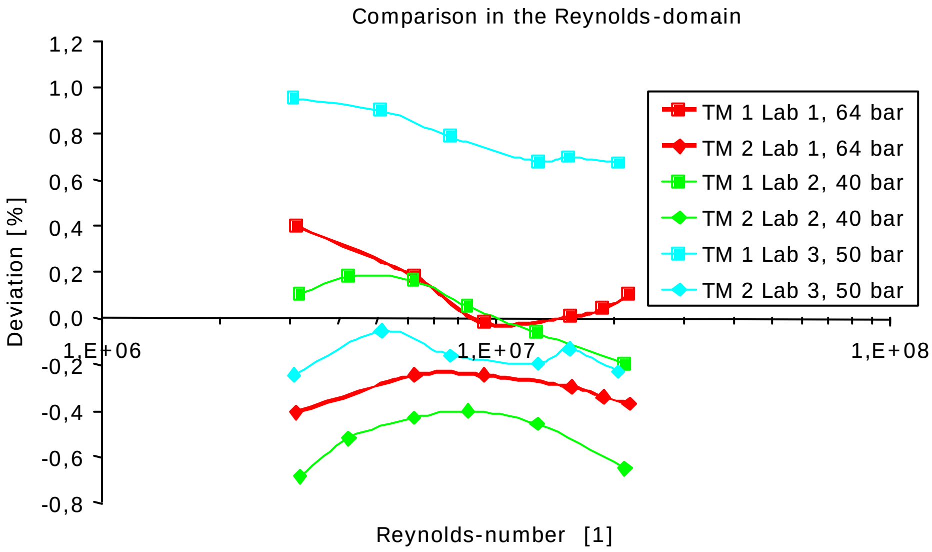

Figure 2, Comparison like in Figure 1, but with the same data in the Reynolds-domain.

COMPARING DATA WITH DIFFERENT UNCERTAINTIES

After making the different data-sets comparable, the next step is to search for the "best known" reference values. When comparing data-sets of different laboratories, the goal is to get a valid relation between the laboratories with respect to a common mean (or average) value. By just treating all measurements in the same way and calculating a traditional average, this will give an out-of-balance picture of the relations between the laboratories. No attention has been paid to the different qualities of the measurements, as expressed in the accompanying stated uncertainties of the measurements of the participating laboratories. A laboratory with the smallest CMC, i.e. the smallest lack of information, should of course have the biggest impact on the position of the "average value". So, the better approach to the mean is to give the laboratory with the smallest uncertainty the biggest impact. The looked for "best known value" is called "Weighed Mean Value" or x_{WMV}. Its determination starts with

$$x_{WMV} = w_1 \cdot x_1 + w_2 \cdot x_2 + \ldots + w_n \cdot x_n \tag{2}$$

x_i are the results (data-sets) of the (n) participating laboratories. For the weighing-factors w_i a standard condition to normalize the weighed expression exists

$$w_1 + w_2 + \ldots + w_n = 1 \tag{3}$$

The weighed uncertainty $u_{weighed}$ of (2) can be derived using the rules for Propagation of Uncertainty

$$u^2_{weighed} = \sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i} \right)^2 u^2(x_i) + 2 \sum_{i=1}^{n} \sum_{j=i+1}^{n} \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} u(x_i, x_j) \tag{4}$$

Because the case of interest is weighed comparison, we apply (4) to the Weighed Mean Value (2), by substituting function f by x_{WMV}.

Next, the assumption is made that the participating laboratories are all independent (no correlation between the data-sets), and thus the covariance-terms will be equal to zero. For comparisons of National metrology institutes this independency is likely, but not necessary the case [4].
With correlated data-sets, e.g. when participating laboratories share the same sets of reference values, the approach has to be different.

Assuming strict independency and thus neglecting covariance, expression (5) describes the uncertainty belonging to the weighed mean value x_{WMV}:

$$u_{weighed}^2 = w_1^2 \cdot u_1^2 + w_2^2 \cdot u_2^2 + \ldots + w_n^2 \cdot u_n^2 \tag{5}$$

Continuing the search for best weighing, (4) has to be minimized, i.e. is to find weighing-factors that give the smallest uncertainty to our Weighed Mean Value. The weighing-factors are a function of the variance $Var(x_i) = u_i^2$, which can be expected instinctively because the variance u_i^2 is an indication of the quality of value x_i.
To find extremes, we apply a partial differentiation of (5). Keep in mind that, by using (3), not all weighing-factors are independent, so that one factor must be expressed in all others to find a satisfactory solution at the end.

$$w_1 = 1 - w_2 - w_3 - \ldots - w_n \tag{6}$$

The partial derivatives will be

$$\frac{\partial u_{weighed}^2}{\partial w_i} = -2\,(1 - w_2 - w_3 - \ldots - w_n)\,u_1^2 + 2\,w_i\,u_i^2 = 0 \tag{7}$$

and can be expressed in (n-1) independent linear equations.

$$\begin{aligned} & w_2\left(1+\left(\frac{u_2}{u_1}\right)^2\right) + w_3 + w_4 + \ldots + w_n = 1 \\ & w_2 + w_3\left(1+\left(\frac{u_3}{u_1}\right)^2\right) + w_4 + \ldots + w_n = 1 \\ & \qquad \vdots \\ & w_1 + w_2 + w_3 + \ldots + w_n\left(1+\left(\frac{u_n}{u_1}\right)^2\right) = 1 \end{aligned} \tag{8}$$

One of the mathematical tools to solve this system of linear equations is the Cauchy Elimination Method, in which (8) is written as a matrix.

$$\begin{bmatrix} 1+\left(\frac{u_2}{u_1}\right)^2 & 1 & \cdots & 1 & 1 \\ 1 & 1+\left(\frac{u_3}{u_1}\right)^2 & \cdots & 1 & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 1 & \cdots & 1+\left(\frac{u_n}{u_1}\right)^2 & 1 \end{bmatrix} \tag{9}$$

Solving matrix (9) is a simple computational exercise, finally minimizing the matrix to one single row.

$$\begin{bmatrix} 1+\left(\frac{u_2}{u_1}\right)^2 + \left(\frac{u_2}{u_3}\right)^2 + \left(\frac{u_2}{u_4}\right)^2 + \ldots + \left(\frac{u_2}{u_n}\right)^2 & 0 & \cdots & 0 & 1 \end{bmatrix} \tag{10}$$

Now the first solution for the weighing-factors has been found

$$w_2 = \frac{1}{1 + \left(\frac{u_2}{u_1}\right)^2 + \left(\frac{u_2}{u_3}\right)^2 + \left(\frac{u_2}{u_4}\right)^2 + \ldots + \left(\frac{u_2}{u_n}\right)^2} \qquad (11)$$

Considering the geometry of the matrix, the general expression for the weighing-factors w_i is given by

$$w_i = \frac{1}{\left(\frac{u_i}{u_1}\right)^2 + \left(\frac{u_i}{u_2}\right)^2 + \ldots + \left(\frac{u_i}{u_n}\right)^2} \qquad (12)$$

For completeness and clarity, we divide by u_i^2 to get

$$w_i = \frac{\frac{1}{u_i^2}}{\frac{1}{u_1^2} + \frac{1}{u_2^2} + \ldots + \frac{1}{u_n^2}} \qquad (13)$$

or, written more compactly,

$$w_i = \frac{\frac{1}{u_i^2}}{\sum_{j=1}^{n} \frac{1}{u_j^2}} \qquad (14)$$

Expression (13) and (14) demonstrate the expected : the weighing-factor of value x_i is inversely proportional to the variance u_i^2, and is normalized by the sum of all these terms.
For now, this ends the search for weighing-factors that will give the weighed mean the smallest uncertainty and that can be expressed as "best known" weighed reference value in comparisons.

The necessary next step is to look for the weighed uncertainty $u_{weighed}$, belonging to x_{WMV}.
To this end, use result obtained in (13), in (5) and get

$$u_C^2 = \frac{\left(\frac{1}{u_1^2}\right)^2 u_1^2 + \left(\frac{1}{u_2^2}\right)^2 u_2^2 + \ldots + \left(\frac{1}{u_n^2}\right)^2 u_n^2}{\left(\frac{1}{u_1^2} + \frac{1}{u_2^2} + \ldots + \frac{1}{u_n^2}\right)^2} \qquad (15)$$

and write it as the simple expression

$$u_{weighed}^2 = \frac{1}{\frac{1}{u_1^2} + \frac{1}{u_2^2} + \ldots + \frac{1}{u_n^2}} \qquad (16)$$

This completes the expressions of the weighing-factors, the Weighed Mean Value with its corresponding weighed uncertainty.

When dealing with comparisons, the Weighed Mean is not the only term of interest in the evaluation of a participating laboratory (to obtain a delta for each partricipants). Also $u_{weighed}$ is important. It is used to determine if the established delta is significant or not.
The observed data-sets are usually determined by a large number of transfer standards over extended flow-ranges. For necessary overlap, two or even more different meter sizes are involved in the same Reynolds-region, showing variations in results. For the difference (delta) of a laboratory to the Weighed Mean Value, all results are taken into account.
Also results taken at different operational conditions (e.g. at different pressure) will show variations. So in the delta of a laboratory with the weighed mean value, the values contributing to the delta will give an additional uncertainty $u_{comparison}$. Even if laboratories are compared with a shared uncertainty, $u_{weighed}$, vis à vis the weighed mean, they differ in their delta-term $u_{comparison}$ as a quantification of the individual quality, on top of their CMC. As the Weighed Mean Value is composed by the results (deviations and uncertainties) of all laboratories, the delta of one laboratory with respect to the Weighed Mean is influenced by the results of all other laboratories and vice versa.

When the comparison term is added to (16) the expression is obtained that is used in the Quadratic Weighing procedure, applied in the harmonization between PTB and NMi VSL [3]. The harmonization can be regarded as an "ultimate comparison".

$$u_{harmonized,i} = \sqrt{u^2_{weighed} + u^2_{comparison,i}} \tag{17}$$

INTERPRETATION OF WEIGHED MEAN VALUE AND ITS UNCERTAINTY

When in a comparison procedure different laboratories are evaluated and a weighed reference value is determined, one property of that common reference is clear : the weighed uncertainty is smaller than any of the uncertainties of the participating laboratories. Look at (18).

$$u^2_{weighed} = \frac{1}{\frac{1}{u_1^2} + \frac{1}{u_2^2} + \ldots + \frac{1}{u_n^2}} < \frac{1}{\frac{1}{u^2_{min}}} \tag{18}$$

So, $u^2_{weighed} < u^2_{min}$ (19)

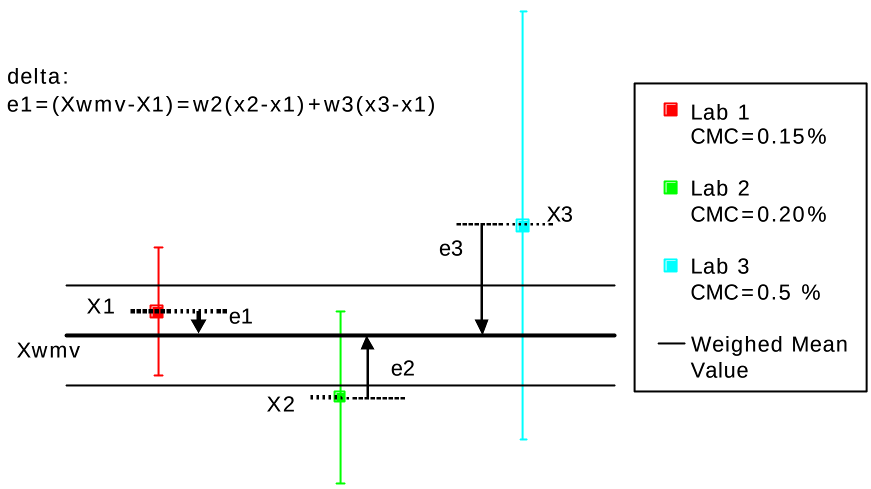

Figure 3 , Example of weighing between laboratories with different uncertainties

The delta between a value x_i (with uncertainty$_i$) and the Weighed Mean Value x_{WMV} (with weighed uncertainty) is called e. Delta's, expressed as e, are shown in figure 3. Also the significance of individual laboratories in the comparison needs to be established. Should the contributed data-set be used to calculate the weighed mean or should it be excluded ? A common expression is applied.

$$e_i \langle \sqrt{u_i^2 + u_{weighed}^2} \qquad (20)$$

In figure 4 and 5, illustrations of the weighing procedures are shown for the data-sets of the same three laboratories as in figures 1 and 2. Residues are plotted in relation to the "average" in figure4. All values have the same impact and data-points are distributed around, y = 0. When the data-sets are plotted with their residues in relation to the weighed mean, figure 5 is obtained. The "centre" of the data-points is no longer in the middle of the data-sets, but is shifted to the data-set with the largest impact. For analysis of results of comparisons a procedure with a weighed mean is advocated.

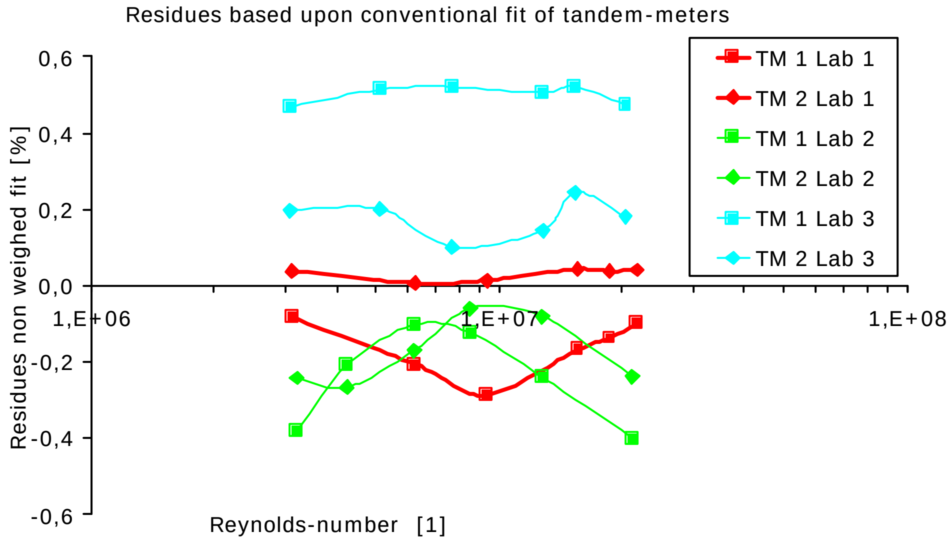

Figure 4, Residues of the three laboratories compared to a conventional "Average"

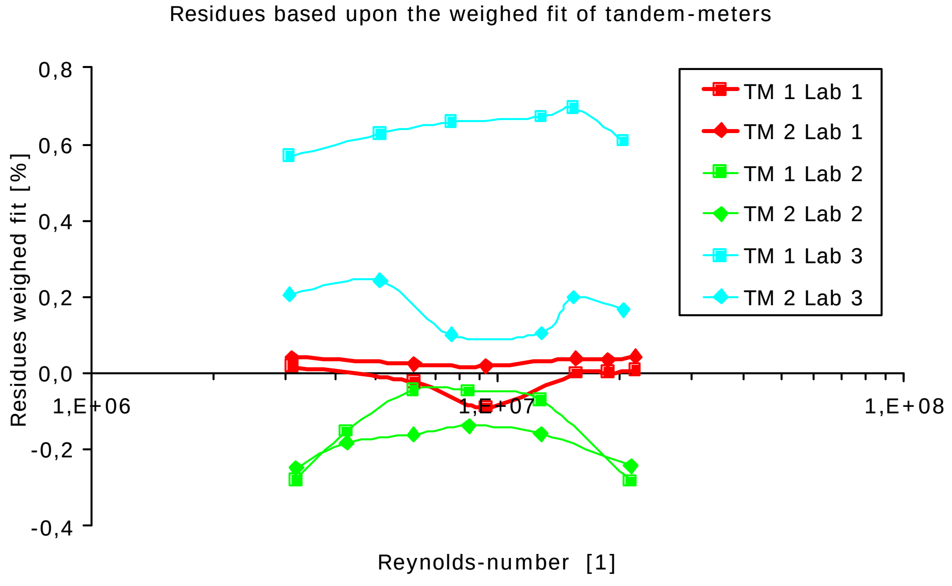

Figure 5, Residues compared to the Weighed Mean Value

USE OF YOUDEN-PLOTS IN COMPARISONS

Youden-plots are often used to evaluate comparisons. It is a method to evaluate both reference values and reproducibility of laboratories and is useful to compare the performance of different laboratories at comparable operational conditions. But analysis of comparisons, based on Youden-plots, could become more valuable when a weighed mean is used instead of an un-weighed one. Figures 6 en 7 show the differences between the two approaches, applied on the same data-sets as before.
It illustrates that assessments of e.g. inter-laboratory reproducibility, could produce different results depending on which method would be followed.

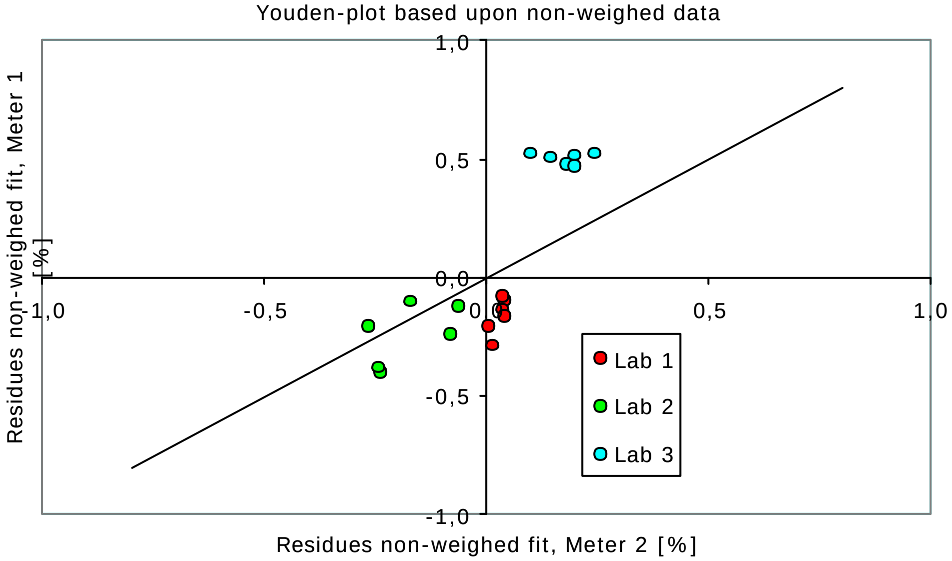

Figure 6 , Youden-plot with the performance of the three laboratories based on an "Average"

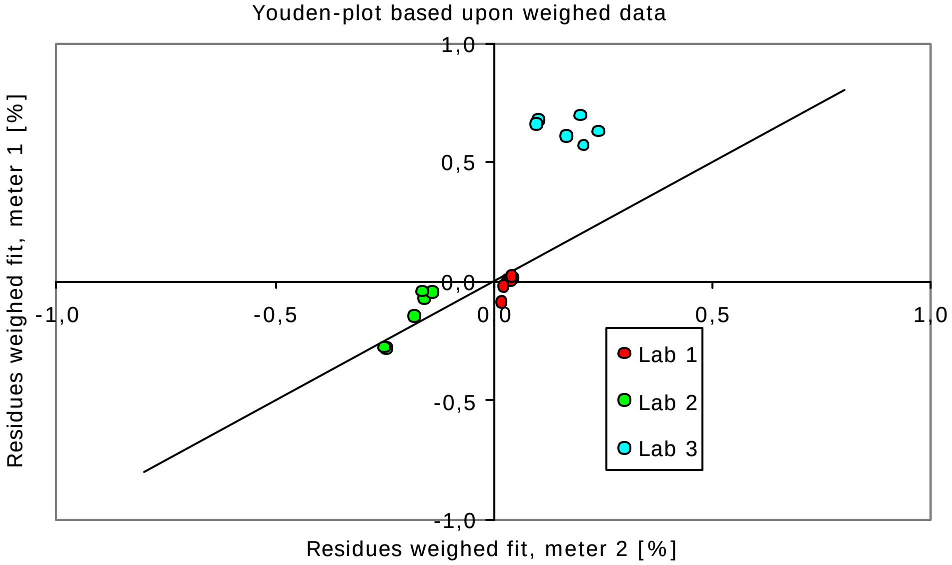

Figure 7 , Youden-plot showing the performance as in figure 6 with residues related to a Weighed Mean Value.

A lot of literature exists explaining the use of Youden-plots, but some points have to be kept in mind if used in a "weighed" comparison. The correct way to apply Youden-plots for the evaluation of different laboratories is to make sure that the observed data have the same process point. In fact, the data of figures 6 and 7 has to be included in six independent Youden-plots instead of just one. This is a disadvantage while more separate graphs need to be evaluated for different Reynolds-number at the same operational conditions. Another disadvantage is the behaviour of the individual, used transfer meters that is not visible anymore, so plots like in figure 2 and 5 remain essential.

CONCLUSION

Before a comparison is organized, one has to assure that the data-sets of the participating laboratories are comparable. For a proper use of comparison results a realistic model of process conditions is needed. The dimensionless Reynolds-number is a useful tool. Now, data within the same Reynolds-region can be compared and a weighed mean value established. It is demonstrated that the most likely value for "best known" reference values is obtained if weighing-factors are applied which minimize the corresponding uncertainty. These weighing-factors are inversely proportional to the variance u^2, and will lead to the smallest uncertainty of the Weighed Mean Value.

Weighed values could improve the assessment of the results of comparison of laboratories. As all theory discussed here is readily available and is not that hard to implement, we recommend including the available uncertainty information in weighed means and Youden-plots.
Including uncertainty information in comparisons means inclusion of quantified quality of the participating organizations.

References

[1] The Acceptance of Variations in Reference Values for high-pressure Gas-Flow measurements, M.P. van der Beek and I.J. Landheer, Flomeko 1998, LUND, Sweden

[2] Developments in methodology to compare the performance of high-pressure gas-flow measurement facilities, M.P. van der Beek and I.J. Landheer, FCRI - Conference on Flow Metering & Control 2000, PELAKKAD, India

[3] NMi VSL and PTB establish common Reference Values for the unit of Volume for high pressure Gas-Flow measurements, T. Kurschat, G. Wendt, M.P. van der Beek and I.J. Landheer, Proceedings of North Sea Flow Measurement Workshop 1999, OSLO, Norway

[4] Exclusive statistics : Simple treatment of the unavoidable correlations from key comparison reference values, A.G. Steele, B.M. Wood and R.J. Douglas, Metrologia, **38**, 483 (2001)