

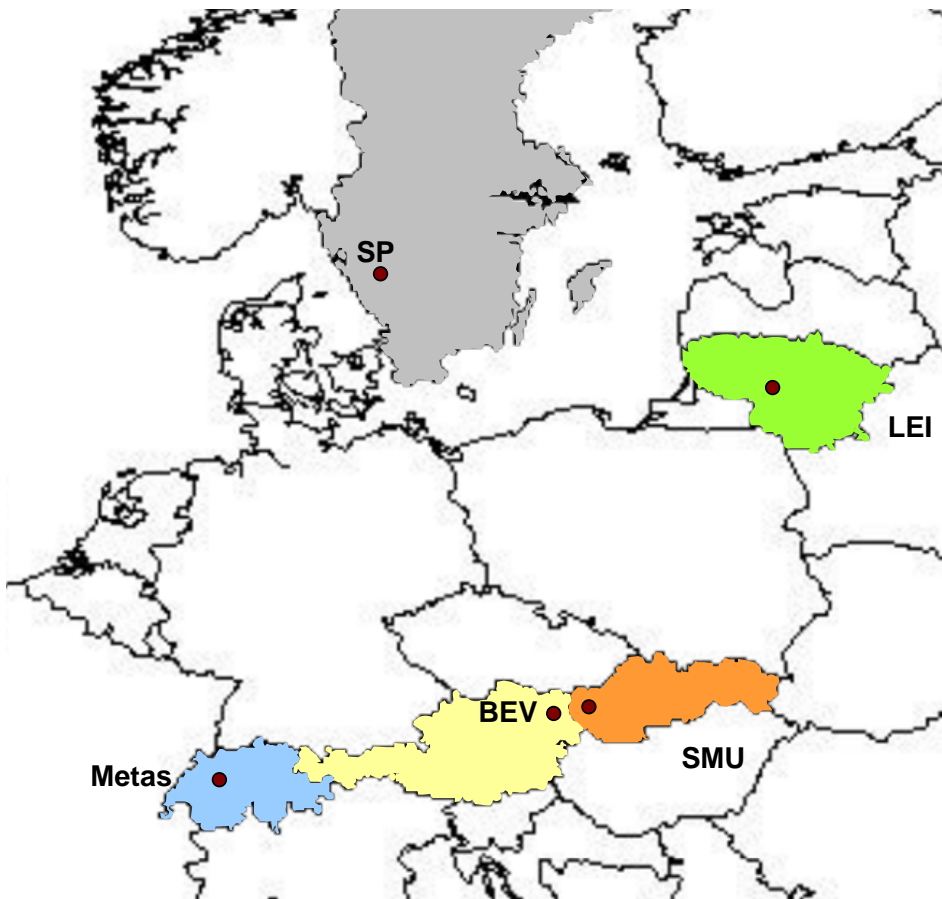
Intercomparison on water/heat meter calibration at 50 °C, 6 - 25 L/h

Krister Stolt, Peter Lau

SP Technical Research Institute of Sweden



Euramet project 877



Intercomparison on water/heat meter calibration at 50 °C, 6 - 25 L/h

Krister Stolt, Peter Lau

Abstract

Intercomparison on water/heat meter calibration at 50 °C, 6 - 25 L/h

A flow meter inter-comparison on water/heat meters at a temperature of 50 °C was performed involving 5 European national flow laboratories. SP as the pilot laboratory selected a transfer package of two stable flow meters, a coriolis mass flow meter and a magnetic inductive meter. The calibrations were performed with water at 50 °C and at three flow rates (6, 12 and 25 L/h). The inter-comparison was held open for more than the five final participants, but was closed after two years. The meters were transported by the project leader to the participating laboratories at convenient occasions. To simulate realistic everyday calibration work each laboratory had only two days to perform the job. Three signals from two meters were evaluated. The degree of equivalence (DoE) in these calibrations are reported along with the En-values.

The maximum difference between the participants with respect to the two transfer meters at all three flow rates was 0,6 % and 0,45 % for the magnetic inductive and coriolis meter respectively. Of the 42 partial results (2 meters, 3 signals, 3 flow rates, 5 participating laboratories) 3 results fell outside a $\pm 0,2$ % DoE-interval from the relevant reference values, being characterized as outliers.

Key words: flow meter, inter-comparison, transfer standards, degree of equivalence, En-value, low flow rate, hot water

SP Sveriges Tekniska Forskningsinstitut
SP Technical Research Institute of Sweden

SP Report 2009:49
ISBN 435-237 8869
ISSN 0284-5172
Borås 2009

Contents

Abstract	3
Innehållsförteckning / Contents	4
Förord / Preface	5
Sammanfattning / Summary	6
1 Scope	7
1.1 Participating laboratories	7
2 Experimental design	7
2.1 Transport	7
2.2 Measurement protocol	8
2.3 The meters	8
2.4 Installation and measurement conditions	8
2.5 Calibration information	9
2.5.1 Flow signal	9
2.5.2 Flow rates	10
2.5.3 Temperature	10
2.5.4 Time schedule	10
2.5.5 Meter zeroing	10
3 Results	10
3.1 Comparison results in numeric form	10
3.1.1 Danfoss mass signal	10
3.1.2 Danfoss volume signal	11
3.1.3 Endress and Hauser volume signal	12
3.2 Determining the comparison reference value, CRV	12
3.2.1 Arithmetic mean	13
3.2.2 Consistency test and weighted mean	13
3.2.3 Average median and Monte Carlo Simulation	14
3.3 Comparison results in graphical form	15
3.3.1 Danfoss mass signal	15
3.3.2 Danfoss volume signal	16
3.3.3 Endress and Hauser volume signal	17
3.3.4	17
3.4 Stability of the meters	17
3.4.1 Danfoss mass- and volume signal	17
3.4.2 Endress and Hauser volume signal	18
3.5 Comparison of results	19
3.5.1 Degree of equivalence	19
3.5.2 En-values	19
3.5.3 Tabulated results for DoE and En-value	20
3.6 Youden Plots	20
4 Interpretation of results and conclusions	22
5 References	23

Appendix: Applicationform

Preface

The service, repair, test, calibration and exchange of heat meters is mostly organized by accredited inspection bodies belonging to a private company or being part of a municipal organization. Those inspection bodies can base their traceability either on traceable mass and density measurements or they can receive traceability from their national flow laboratory.

The best way to verify uncertainty claims is for these and national flow laboratories alike to take part in suitable inter-comparisons.

The measurement of low flow rates of water at elevated temperatures is quite difficult. A number of effects can influence the outcome. Being a dynamic quantity, systematic errors can easily occur in water flow measurement and are difficult to detect. Inter-comparison have become the most appropriate way of evaluating a laboratories measurement capability. And national flow laboratories should evaluated themselves from time to time by organizing/taking part in a suitable round robin exercise where a meter is calibrated at different flow facilities.

The most convincing way for a national flow laboratory to reach acceptance lies in the possibility to get a calibration result related to a key-comparison, i.e. a degree of equivalence to a key comparison reference value (KCRV) . However, for hot water and low flow rates no key-comparison has been arranged yet and a reference value does not exist. Thus the scope of this round robin was to achieve a limited reference value representative for the participating laboratories.

Sammanfattning

Föreliggande rapport beskriver en jämförelsemätning av två varmvattenmätare i serie som kalibrerats vid 5 europeiska laboratorier. Mätarna, en coriolis massflödesmätare och en magnetisk induktiv mätare, är avsedda för mycket låga flöden. Mätningarna utfördes vid 50 °C och tre flöden (6, 12 och 25 L/h). Varje deltagande laboratorium fick två dagar för sin kalibrering, vilket ansågs motsvara normala rutinförhållanden och projektledaren tog ansvar för hela transporten. Projektet hölls öppet för fler deltagare men avslutades efter drygt två år. Sammanlagt utvärderades 3 signaler vid 3 flöden. Den sammanlagda överensstämmelsen mellan de fem deltagarna med hänsyn till referensvärden från två mätare och alla tre flöden låg inom en marginal på $\pm 0,2$ %. Tre av sammanlagt 42 mätpunkter låg utanför $\pm 0,2$ % marginalen och betraktas som outliers.

1 Scope

Flow meters for hot water are mostly used for measuring an accumulated volume and they are predominantly calibrated to a static mass or volume. Flow, however, is a dynamic quantity. Only via an inter-comparisons flow calibration laboratories can show that their calibrations give correct judgements of the flow meters capability and thus assure a calibration quality. The main purpose of this project was to demonstrate that the participating laboratories can provide traceability to the secondary laboratories and inspection bodies in their countries. There is no absolute calibration curve serving as a reference for the meters. The three reference values found at three flow rates is made up of the results from the participants.

1.1 Participating laboratories

This inter-comparison was started in 2006 and open to all interested laboratories within Europe under Euramet project number 877. The application form is shown in appendix 1.

The participating laboratories from 5 countries are listed below in the order of performing the calibration measurements. Originally further laboratories were interested to take part. For different reasons, however, those measurements were postponed and finally when planning the visit the interest to participate was withdrawn. This was a major cause for the delay in the project that was finished in early 2009.

Table 1 Participating laboratories

Laboratory		Country	Calibration method	Volume [L]	Date of Calibration
BEV	Bundesamt für Eich- und Vermessungswesen	Austria	Gravimetric	5 & 10	June 2006
SP*	Technical Research Institute Sweden	Sweden	Volumetric	5	October 2006
Metas	Federal Office of Metrology	Switzerland	Gravimetric	5 & 10	November 2006
SMU	Slovak Institute of Metrology	Slovak Republic	Gravimetric	3 & 5	November 2006
LEI	Lithuania Energy Institute	Lithuania	Gravimetric	5 & 10	November 2008

* Pilot Laboratory

2 Experimental design

An important aspect for a inter-comparison experiment is the way it is planned and performed. In this exercise the aim was to make all participants to run the calibration of the meters as a routine task like any other calibration. The idea was to simulate a usual calibration situation giving the operators only two days to complete the job. It also implied that no prior information on the meters were given except their type and size, the temperature and the flow rates for the calibration. No time was intentionally given to get to know the meters and to possibly improve measurement techniques or to make additional test runs to confirm prior results.

2.1 Transport

Problems with failing meters due to transportation and handling are not unusual and have limited the possible outcome in earlier comparisons. For this project it was decided to deliver the

meters personally by car. The project leader also supervised the mounting, the electrical connection just to avoid harm to the meters. The ambition was purely to make sure no simple handling problems should lower the gained information. The coriolis meter with its 1,5 mm pipe diameter is particularly sensitive to thrust. No direct help of any kind was given. After the two days the meters were collected the same way. For detecting eventual drift intermediate calibrations were performed at SP.

2.2 Measurement protocol

The main purpose was to perform the measurements in a stipulated way concerning flow rates and to present the results using the ordinary calibration certificate. No translation to English was asked for. Further no details needed to be delivered on how the measurements were performed in detail.

This is quite different from how inter-comparisons often are conducted, especially key comparisons, where a technical protocol first has to be decided on by all participants and the BIPM.

2.3 The meters

It has become tradition to use at least a pair of meters in series as a transfer package when performing flow meter calibration comparisons. Moreover the two meters preferably are based on different measurement principles. For this experiment a coriolis mass flow meter and an magnetic inductive volume flow meter also referred to as Mag-meter were used. Both have been used at SP for several years and showed good stability over time. They also have served as reference meters for audit measurement with laboratories inside and outside of Sweden. The meters are shown in the figure 1 and 2 below.



Fig. 1 Danfoss Mass 2100 - Coriolis mass flow meter; Pipe size Ø 1,5 mm



Fig. 2 Endress and Hauser - Magnetic inductive volume flow meter; Pipe size Ø 2 mm

2.4 Installation and measurement conditions

The meters were to install in series and their respective errors were to determine simultaneously. Whereas the Endress and Hauser meter provided a volume flow signal, the Danfoss meter, being a coriolis meter could provide two signals, one for mass and one for volume flow rate. For the volume signal the internal density measurement signal can be used. This renders the possibility to compare three simultaneous signals.

Due to the low flow rates the water temperature drops very easily, which is a major error source. To prevent this a special arrangement was suggested. The coriolis meter was connected to the test line via insulated hoses and heated via the support in its bottom. Hot water from a separate bath, where the temperature could be adjusted to the temperature in the test line, was circulated

through this support (see figure 3a). To produce a constant ambient temperature independent of the flow rate, the meter was placed in a insulating box of polystyrene foam (see figure 3b).

The Mag-meter was mounted the normal way in the test line, but with additional insulation (see figure 3c).

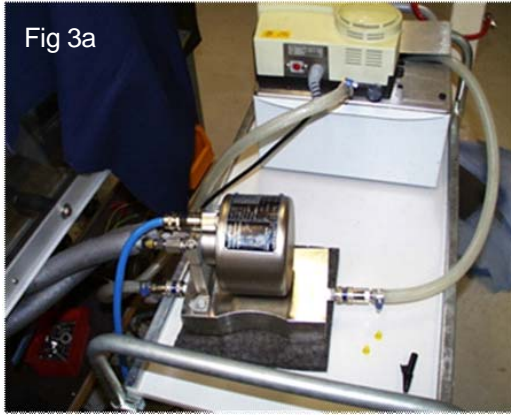


Fig. 3a Coriolismeter with heated support at the bottom and insulated connection hoses.

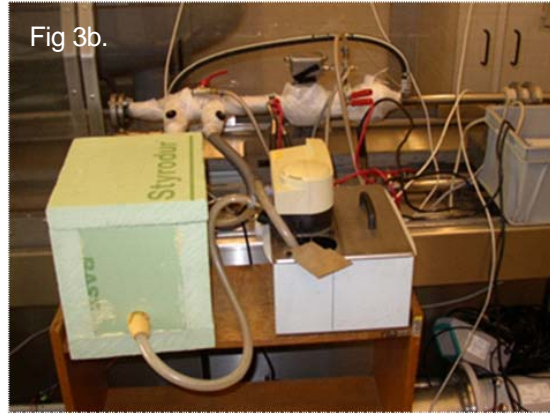


Fig. 3b Coriolis meter in an insulation box with support connected to a controlled water bath.



Fig. 3c Mag-meter with extra insulation in the test line left and coriolis meter on a trolley connected in series via hoses.

2.5 Calibration information

2.5.1 Flow signal

The required information was the measurement error E at three different flow rates based on at least 5 repeated runs. This means the meters were calibrated with respect to the passed liquid volume under actual conditions giving an actual k-factor in pulses per litre or pulses per kg. The measured k-factors $K_m(q)$ then were compared to nominal ones set by SP to 1000 p/L and 1000 p/kg respectively (for the coriolis meter both apply). The meter errors $E_m(q)$ were then calculated according to the following equation making the three signals from the two meters directly comparable for different flow rates.

$$E_m(q) = \frac{K_m(q) - K_{nom}}{K_{nom}} \cdot 100 \quad [\%] \quad (1)$$

Here q denotes the actual flow rate in L/h and the index m denotes one of three meter signals (mass or volume from the coriolis- and volume from the Mag-meter).

Depending on the measurement method $K_m(q)$ can be based on a collected volume or a weighed mass of water. For the majority the later was the case. Thus a transformation from mass to volume was necessary based on the water density, which was calculated from the measured temperature and standard water tables.

2.5.2 Flow rates

The three flow rates for which the meter error was determined were at 6, 12 and 25 L/h. These are usual points for heat meter testing at q_{\min} . Especially at low flow rates it is difficult to keep a

constant hot water temperature, which makes this situation critical. The amount of liquid passing the meter and the reference is not totally equivalent if temperature changes along the pipe line. A correction of the change in contained volume between the meter and the volume standard is often tricky, but necessary. To correct the passed volume for the temperature loss between the meter and the reference is not enough.

2.5.3 Temperature

The water temperature of 50 °C chosen for this comparison is a typical temperature for test purposes. It is also interesting as former meter comparisons at higher flow rates were performed at this temperature, thus giving a chance to judge the outcome of this exercise with others, especially concerning the degree of equivalence between laboratories that can be achieved.

2.5.4 Time schedule

For this comparison measurements no time schedule was erected. The reason was twofold. All the participants were not decided at the beginning and some planned calibration events were postponed. Thus they were planned in agreement with each participant at convenient occasions and in connection with other commissions for the project leader to keep transportation costs low.

2.5.5 Meter zeroing

Meters drift with time and temperature. The two meters used in this inter-comparison were never zero-adjusted, something that is recommended by the suppliers. The reason is to be able to follow up the drift. Likewise the meters were never zeroed during the calibration at any of the participating laboratories.

3 Results

The numeric outcome of the comparison is collected in the tables 2 to 4. The data is collected from the different calibration certificates. The three tables concentrate each on one of three signals and present the calibrated meter factor error for three flow rates and the corresponding uncertainty stated by the participating laboratories. These are on a 95 % confidence level ($k=2$) and indicated by italic style.

The lower part of the tables contains alternative suggestions for a comparison reference value (CRV) for the meter factor errors and the uncertainty involved in this reference value. The preferred choice for the CRV is always the weighted mean (see chapter 3.3), which is recommended 5 times. The arithmetic mean is calculated purely for comparison. In four of nine individual comparison situations the average median is decided to represent the best reference value. The chosen reference value is indicated in bold style and by marking the corresponding cells in the table light yellow. For details see chapter 3.2.2 and 3.2.3

3.1 Comparison results in numeric form

3.1.1 Danfoss mass signal

Table 2 presents the meter factor error of the mass signal of the coriolis meter for three different flow rates at 50 °C. The error was calculated by each laboratory according to equation (1) and given as percentage of the nominal value. As can be seen the error is largest at low flow rates, which is typical for a mass flow meter and depends largely on the zeroing of the meter, which was not performed at any place.

Table 2 Error in Danfoss mass signal at 50 °C

Laboratory	E(6L/h)	U(6L/h)	E(12L/h)	U(12L/h)	E(25L/h)	U(25L/h)
	[%]	[%]	[%]	[%]	[%]	[%]
BEV	0,46	0,17	0,31	0,11	0,05	0,11
SP	0,42	0,19	0,29	0,17	0,08	0,15
SMU	0,50	0,22	0,46	0,14	0,29	0,14
METAS	0,52	0,14	0,54	0,17	0,41	0,20
LEI	0,61	0,17	0,17	0,15	0,10	0,15
Weighted mean	0,507	0,077	0,345	0,064	0,150	0,064
Arithmetic mean	0,502	0,073	0,354	0,149	0,186	0,160
Av. Median (MCS)	-	-	0,338	0,100	0,138	0,106
Range in CRV [%]	1,0		4,6		30,4	

Only at the lowest flow rate all reported errors led to a consistent result (see chapter 3.2.2) in the sense that the estimated uncertainties cover the differences to the reference value (compare figure 5). In this case the weighted mean is assumed to produce the best reference value with a meter factor error of 0,507 %. For the other two flow rates no consistency could be reached among the results and the reference value was determined by a Monte Carlo Simulation (MCS) on the median (see chapter 3.2.2) and averaged over 5 times 5000 simulations. This is the recommended procedure in such a situation [1]. Thus a best meter error for reference is given in the light yellow cells in table 2 together with its calculated uncertainty. For comparison also an arithmetic mean is calculated.

For the lowest flow rate no MC-simulation was performed. It is interesting to note that the way of determining the CRV leads to clearly different values for the comparison. The range, i.e. the difference between the highest and lowest CRV is largest for the high flow rate and amounts to 30,4 % of the average for the three calculated CRV's. Such big differences are unusual and depend in this case on the relative small number of laboratories.

3.1.2 Danfoss volume signal

Besides the mass flow the coriolis meter can also measure the density of the passing liquid. This allows to produce a volume flow signal, which integrated over the time of a run gives the volume of passed liquid. The accumulated number of pulses divided by the accumulated volume gives the meter factor, which then is compared to the nominal one. The relative deviation is shown in table 3.

Table 3 Error in Danfoss volume signal at 50 °C

Laboratory	E(6L/h)	U(6L/h)	E(12L/h)	U(12L/h)	E(25L/h)	U(25L/h)
	[%]	[%]	[%]	[%]	[%]	[%]
BEV	0,44	0,18	0,28	0,11	-0,07	0,11
SP	0,38	0,19	0,24	0,17	-0,03	0,15
SMU	0,26	0,19	0,18	0,15	-0,05	0,13
METAS *	-	-	-	-	-	-
LEI	0,57	0,19	0,10	0,17	-0,08	0,17
Weighted mean	0,413	0,094	0,219	0,072	-0,058	0,067
Arithmetic mean	0,413	0,155	0,200	0,094	-0,058	0,027
Av. Median (MCS)	0,412	0,113	0,182	0,107	-0,059	0,083
Range CRV [%]	0,2		18,5		1,7	

* Metas could not report on this signal as their data acquisition system did not have enough channels.

Table 3 shows better agreement between the participants. The choice of method to determine a reference value is of much less concern. The weighted mean is chosen as the best estimator for the reference value (indicated error in measured volume) at all three flow rates. This is a surprising result that was not expected. When measuring water flow it is usually assumed that the density is best determined from water tables and a temperature measurement. This seems overruled in this case as the meter definitely shows more stable values of the volume signal, whereas the mass signal is the one predominantly used.

3.1.3 Endress and Hauser volume signal

The reported error values for the Mag- meter are arranged in table 4 in the same way as in the previous one with three columns accompanied by their respective estimated uncertainty. The errors are almost the same for all three flow rates.

Table 4 Error in Endress and Hauser volume signal at 50 °C

Laboratory	E(6L/h)	U(6L/h)	E(12L/h)	U(12L/h)	E(25L/h)	U(25L/h)
	[%]	[%]	[%]	[%]	[%]	[%]
BEV	0,68	0,20	0,67	0,11	0,38	0,13
SP	0,76	0,19	0,87	0,17	0,85	0,15
SMU	0,52	0,22	0,79	0,17	0,68	0,13
METAS	0,93	0,23	0,94	0,18	1,01	0,12
LEI	0,94	0,20	0,68	0,17	0,89	0,17
Weighted mean	0,766	0,092	0,761	0,068	0,755	0,062
Arithmetic mean	0,762	0,180	0,790	0,120	0,766	0,249
Av. Median (MCS)	0,823	0,126	0,783	0,113	0,825	0,120
Range CRV [%]	7,8		3,7		9,0	

Best agreement can be stated at the middle flow rate which allowed for the weighted mean to represent the reference. For the other flow rates the Chi-squared test failed demanding for an average median calculated by a Monte Carlo Simulation explained below.

3.2 Determining the comparison reference value, CRV

From a statistical point of view the five results at each flow rate can be considered as a small sample of all possible ones defining a distribution around the most probable result for each meter and flow rate. The central value having the highest probability is considered the best reference value for this comparison. As later or similar comparisons can relate to such a reference value it is of considerable importance to determine it well.

Generally, if the results from all laboratories were given with the same uncertainty, i.e. had the same merit, the arithmetic average would be the given estimator.

As this is obviously not the case, good results, i.e. those with low uncertainty should have a larger impact on the average than results with large uncertainties. Thus a weighted mean, using the measurement uncertainty as weighing factor, seems the most suitable estimator.

This decision, however, requires that the uncertainty estimates from the laboratories are good ones and that there are no significant systematic errors in the results, with other words that the results are consistent with each other. Statistically all results are expected to be a part of a normal distribution with the calculated mean as central value and the calculated uncertainty as the characteristic width of the distribution. This consistency can be tested using a chi-squared test.

If no consistency can be stated a third, more robust estimator is required. In this case a so called Monte Carlo simulation procedure is used based on the median. The simulation can be run with different statistic programs. It can also be carried out with Excel using an add-in program like Pop-tools. From the reported results and belonging uncertainties principally many thousand random sample results are produced. For each sample the median is build and the reference value is given by the average over all these medians. Programs that can produce such a MC-simulation calculate directly a 95 % probability numerically not assuming any specified distribution.

If the results are consistent (see 3.2.2) the weighted mean is the suggested method to calculate the reference value KCRV in key-inter-comparisons [1]. If they are not consistent the average median should be used for the determination of the KCRV, which involves a Monte Carlo Simulation.

3.2.1 Arithmetic mean

Considering the average of the reported errors as reference value and the belonging uncertainty then equation (2) to (4) apply, where all uncertainties in $U(E_i)$ are considered equal.

$$E_{CRV} = \frac{E_1 + \dots + E_5}{5} \quad (2) \quad U_{\text{comp}}(k=2) = 2 \cdot \sqrt{\frac{\sum_{i=1}^5 (E_i - E_{CRV})^2}{5-1}} \cdot \frac{t_{1\sigma}^4}{\sqrt{5}} \quad (3)$$

$$U_{CRV}(k=2) = 2 \cdot \sqrt{\left(\frac{U_{\text{comp}}}{2}\right)^2 + \left(\frac{U(E_i)}{2}\right)^2} \quad (4)$$

This is rarely the case and thus these definitions are normally not used.

3.2.2 Consistency test and weighted mean

The consistency test before using the weighted mean as estimator is performed by comparing the observed χ^2 -value, χ^2_{obs} with a theoretical one $\chi^2(\nu)$ with respect to the actual degree of freedom $\nu=n-1$. The theoretical value can be found with the excel-function $\text{CHI2INV}(0,05;4)$, where 0,05 gives the significance level of 5 % and 4 the actual degree of freedom.

$$\text{The observed value is calculated as } \chi^2_{\text{obs}} = \frac{(E_1 - E_{CRV})^2}{u^2(E_1)} + \dots + \frac{(E_5 - E_{CRV})^2}{u^2(E_{CRV})} \quad (5)$$

The criterion for the test is $\Pr\{\chi^2(\nu) > \chi^2_{\text{obs}}\} < 0,05$, which tells that if the theoretical value is larger than the observed one then the test is accepted with high probability and all five results belong to the same normal distribution and thus are consistent. One example for this calculation is shown in table 5. It concerns the Danfoss meter and the mass flow rate of 6 L/h. The interesting information of the test is found in the yellow cells of the sixth column.

The calculation starts with the reported errors E_i and belonging uncertainties U_i in the second and third column. The following columns 4 and 5 contain the expressions indicated in the head row. It should be observed that u_i equals $U_i/2$. These columns are summed up. The sum in column 5 divided by the sum of column 4 constitutes the reference value E_{CRV} as given by equation (6) and indicated in the bottom of table 5. The fourth column also holds the reversed sum and its square root making up standard uncertainty of reference value $u(E_{CRV})$ according to equation (7), also indicated at the right bottom side. (Observe again u for $k=1$; U for $k=2$).

The sixth column contains the individual terms and the their sum marked yellow giving the observed Chi squared value as defined in equation (5). The theoretical value of the chi-squared

function for 4 degrees of freedom and a significance level of 5 % is found just below is also marked yellow. The observed value $\chi^2_{\text{obs}}=2,65$ is clearly lower than the calculated one 9,49 suggesting that the data in this case pass the consistency test.

Table 5. Calculation procedure for consistency test and weighted mean.

1	2	3	4	5	6	7	8	9
					$\frac{(E_i - E_{\text{CRV}})^2}{\left(\frac{U_i}{2}\right)^2}$			$E_n = \frac{E_i - E_{\text{CRV}}}{\sqrt{U_i^2 + U_{\text{CRV}}^2}}$
Laboratory	E_i [%]	U_i [%]	$\frac{1}{u_i^2}$	$\frac{E_i}{u_i^2}$		$D_i = E_i - E_{\text{CRV}}$	$U(D_i) = 2\sqrt{u_i^2 - u_{\text{CRV}}^2}$	
BEV	0,46	0,17	138,4	63,67	0,309	-0,0473	0,1516	0,25
SP	0,42	0,19	110,8	46,54	0,844	-0,0873	0,1737	0,43
SMU	0,50	0,22	82,6	41,32	0,004	-0,0073	0,2061	0,03
METAS	0,52	0,14	204,1	106,12	0,033	0,0127	0,1169	0,08
LEI	0,61	0,17	138,4	84,43	1,461	0,1027	0,1516	0,55
Average	0,502	SUM	674,3	342,08	2,651	χ^2_{obs}		
		1/SUM	0,0015		9,488	$\chi^2(0,05;4)$		
E_{CRV}	0,507	$u(E_{\text{CRV}})$	0,0385		$E_{\text{CRV}} = \frac{E_1/u^2(E_1) + \dots + E_5/u^2(E_5)}{1/u^2(E_1) + \dots + 1/u^2(E_5)}$		$u_{\text{CRV}} = \sqrt{\frac{1}{\sum_{i=1}^5 \frac{1}{u_i^2}}}$	
$U(E_{\text{CRV}}) k=2$	0,077							

The seventh column holds the degree of equivalence for each laboratory and the eighth column the corresponding uncertainty. The last column finally holds the E_n -value defined as indicated (see also equation (11)). It is almost identical with the degree of equivalence divided by its uncertainty. The slight difference lies in the – and +sign respectively. The E_n value only states a normalized deviation from any reference value assuming no correlation between the individual results and the reference value. In the case of the weighted mean a clear correlation exists between each result and the reference value. This leads to a –sign for the uncertainty combination related to the degree of equivalence and its uncertainty gets a little larger than the E_n -value.

After acceptance of the χ^2 -test the weighted mean and the belonging uncertainty is calculated with the help of equations (6) and (7) below. This as well is shown in table 5.

$$E_{\text{CRV}} = \frac{E_1/u^2(E_1) + \dots + E_5/u^2(E_5)}{1/u^2(E_1) + \dots + 1/u^2(E_5)} \quad (6) \quad U_{\text{CRV}}(k=2) = 2 \cdot \sqrt{\frac{1}{\sum_{i=1}^5 \frac{1}{u^2(E_i)}}} \quad (7)$$

3.2.3 Average median and Monte Carlo Simulation

In this simulation PopTools produced 5 columns (one for each laboratory) with 5000 random values in each. The random numbers were produced based on a normal distribution around the reported errors with a standard deviation taken as half the reported uncertainty ($k=1$). These 25000 random numbers represent 5000 samples with 5 results in a row. From each sample the median was taken and the result sorted in ascending order. PopTools then depicted the average as the best representative reference and also gave a number for the upper and lower percentile making up 95 % of all values around the average. A picture of the distribution from 5000 simulated medians is shown in figure 4 together with the calculated characteristics for the distribution. The uncertainty in this reference value was calculated using equation (8)

$$U(k=2) = \frac{97,5 \text{ percentile} - 0,25 \text{ percentile}}{2} \quad (8)$$

This procedure was repeated at least five times for each non-consistent result (one flow rate, one signal). The median values defining the CRV-values in tables 2 and 4 are the mean of these averages and the same is valid for the stated uncertainties for these reference values .

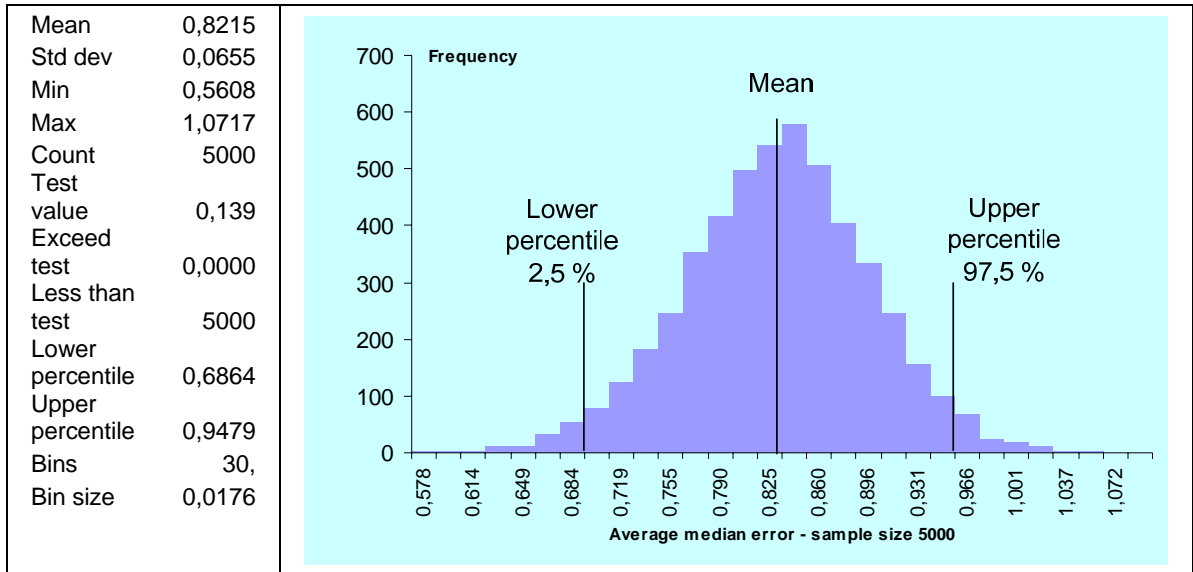


Fig. 4 Result of one Monte Carlo Simulation – average median for a sample size of 5000 to find the comparison reference value and its uncertainty – example E&H-meter, 6 L/h

3.3 Comparison results in graphical form

3.3.1 Danfoss mass signal

The content of table 2 is graphically displayed in figure 5 with the lowest flow rate 6 L/h left and the highest 25 L/h to the right. For the low flow rate the specified uncertainty limits all cover the CRV, which is the weighted mean indicated by a triangular symbol and a blue horizontal line. This is a graphical argument for the consistency between the various results.

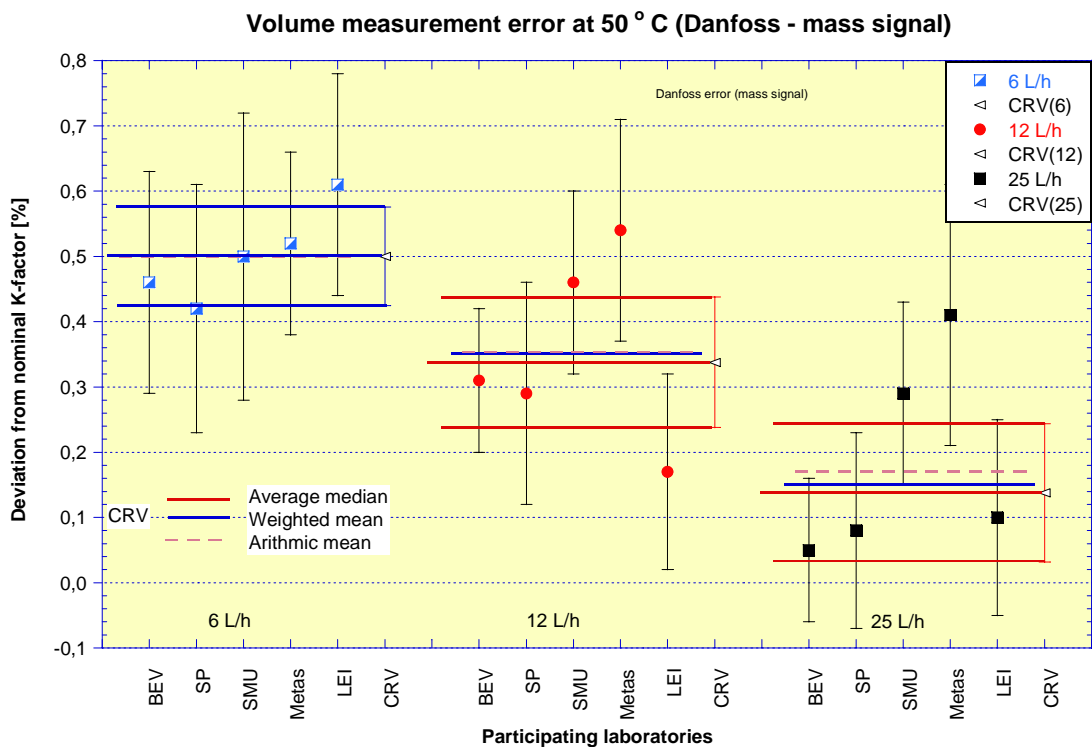


Fig. 5 The meter factor errors of the mass signal from the Danfoss meter at 50 °C as specified by the participating laboratories at three flow rates – data from table 2.

The CRV for the medium and high flow rate is the average median marked by a triangular symbol and a red line. This CRV is calculated using a Monte Carlo Simulation. For comparison the alternative reference values are presented as well along with the chosen one. The uncertainties with each reference value is marked by the red horizontal lines.

Figure 5 does not show a clear trend for one laboratory to have a deviating result. The only pattern that can be seen is that Metas has two high values for the two larger flow rates. The corresponding En-values (see table 6) are 1,02 and 1,2. This is just outside the border of metrological acceptance and above. The En-value provides a somewhat kinder criterion than the degree of equivalence and its uncertainty (compare to 3.5.1 and 3.5.2). The third questionable result is the value of LEI for 12 L/h, which corresponds to an En-value of 0,93 and considered acceptable from a metrological point of view.

3.3.2 Danfoss volume signal

The volume signal of the coriolis meter was recorded simultaneously with the mass signal. One laboratory (METAS) did not have the capability for this recording. Thus one result is missing in the reported errors for the three flow rates (figure 6). Here the four remaining results all overlap the reference value, extremely well at the high flow rate, thus allowing for the weighted mean as best estimator in all cases, represented by the blue line and marked with a triangular symbol. The largest deviation (LEI) at 6 L/h still corresponds to an En-value of 0,74 i.e. still a fully acceptable comparison result.

The general pattern with a decreasing error when flow rate is increasing is the same as in figure 5 for the mass signal.

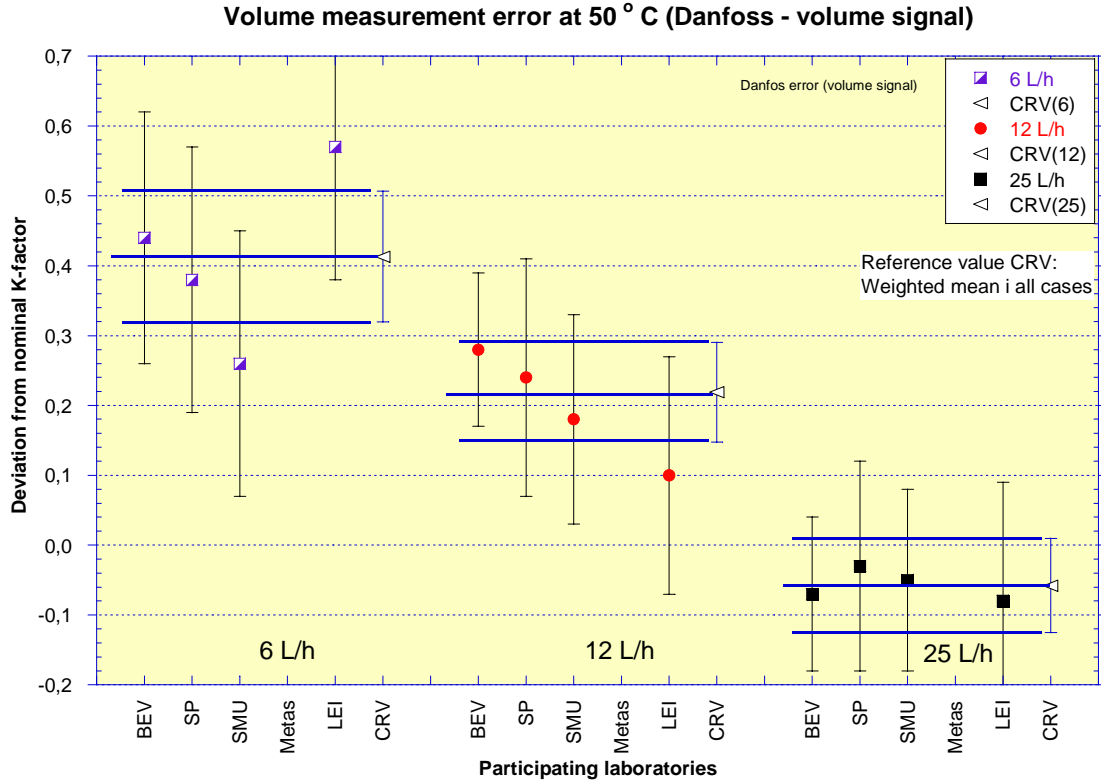


Fig. 6 The meter factor error for the volume signal of the coriolis mass flow meter at three measured flow rates – data from table 3. The corresponding uncertainties are indicated by vertical bars. The reference value is always the weighted mean given by the blue horizontal line. Its uncertainty is indicated by the symmetric blue horizontal lines.

3.3.3 Endress and Hauser volume signal

Based on the comparison reference value the magnetic inductive meter shows a quite constant error over the flow range. This is however not true when looking to the laboratories results separated. The reported errors vary much more between the flow rates and in comparison with other laboratories. Even so the reported uncertainty is typically higher this behaviour causes a situation where consistency, i.e. all uncertainty bars overlap the reference value, only exists for the medium flow for which the weighted mean was the best choice for the CRV. It is marked by a triangular symbol and the blue horizontal line. For the other two flow rates the average median was chosen and calculated via a MC-simulation, indicated with red lines.

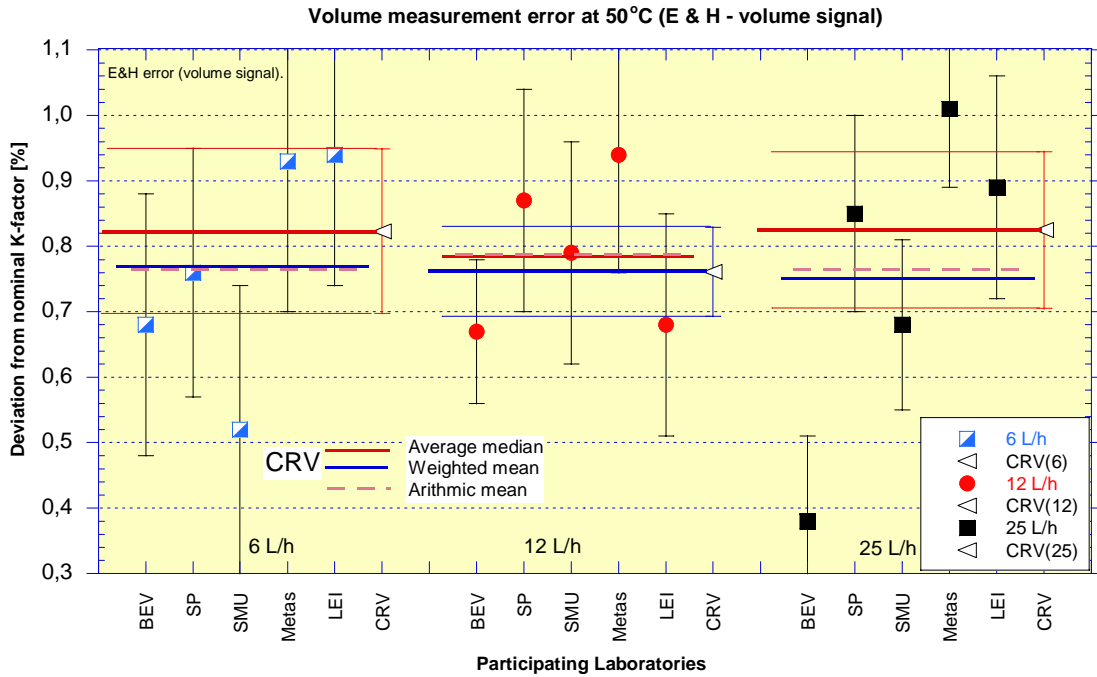


Fig. 7 Reported error in meter factor for the volume signal of the Mag-meter with reported uncertainty bars. Data from table 4. The reference values are marked with a triangle and red or blue lines. The belonging uncertainty limits are given with a pair of red or blue lines depending on the chosen reference value.

3.4 Stability of the meters

The two meters were used irregularly but several times per year and thus calibrated in-between. The red symbols in figures 8, 9 and 10 show the variation in the determined meter factor error over time at approximately 50 °C. They numbering a), b) and c) refer to the three flow rates used for comparison. The first measurements registered for time 0 were performed roughly 100 days before the first calibration belonging to the inter-comparison, the last one after about 1100 days. The white symbols represent the actual data from the participants belonging to the comparison as a function of time .

3.4.1 Danfoss mass- and volume signal

The maximum change over time amounting to 0,1 % for the mass signal and 0,13 % for the volume signal of the Danfoss meter are roughly half the size of SP's measurement uncertainty. A small drift can be seen in both meter signals, but figures 8 and 9 also show that it is smaller than the spread between the laboratories. Thus one can state that the coriolis meter was suitably stable for the comparison.

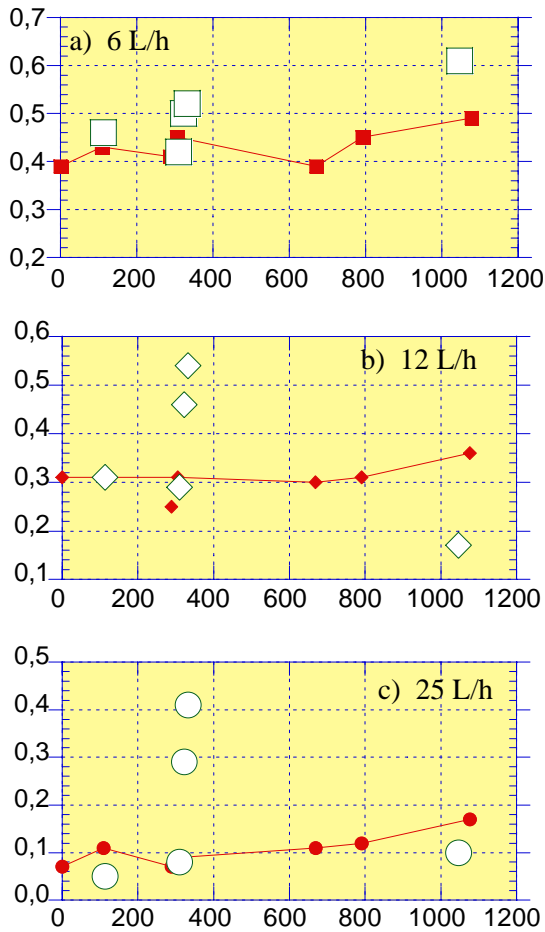


Fig. 8 Meter error of the coriolis mass signal in [%] over time [days].

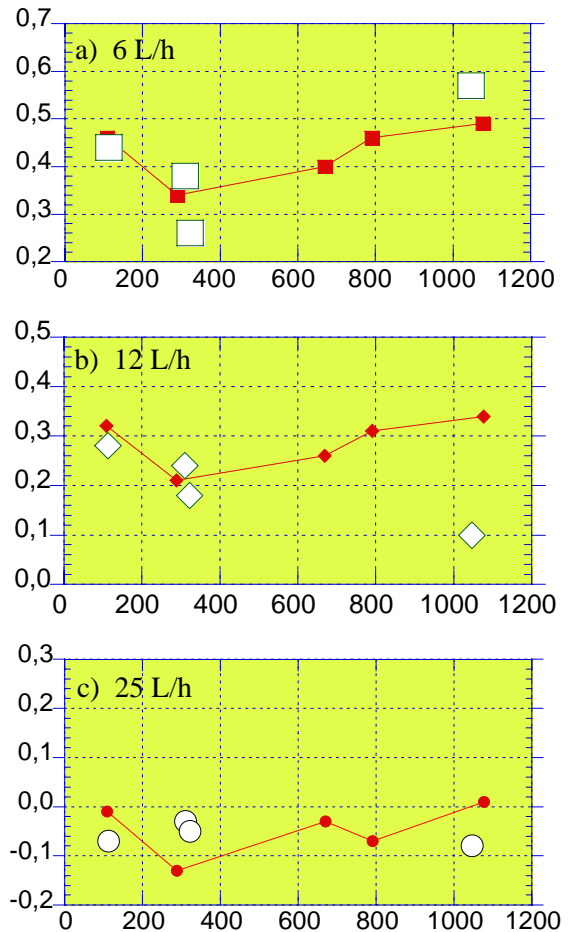


Fig. 9 Meter error of the coriolis volume signal in [%] over time [days].

3.4.2 Endress and Hauser volume signal

The magnetic inductive meter is not as stable as the coriolis meter. This statement applies both to the repeatability and a drift over time. Figure 10 a) to c) shows the inter-comparison data (open symbols) on the background of intermediate calibrations (filled symbols). For all three flow rates a significant tendency for an increasing meter factor error can be seen, when compared to the specified measurement uncertainty.

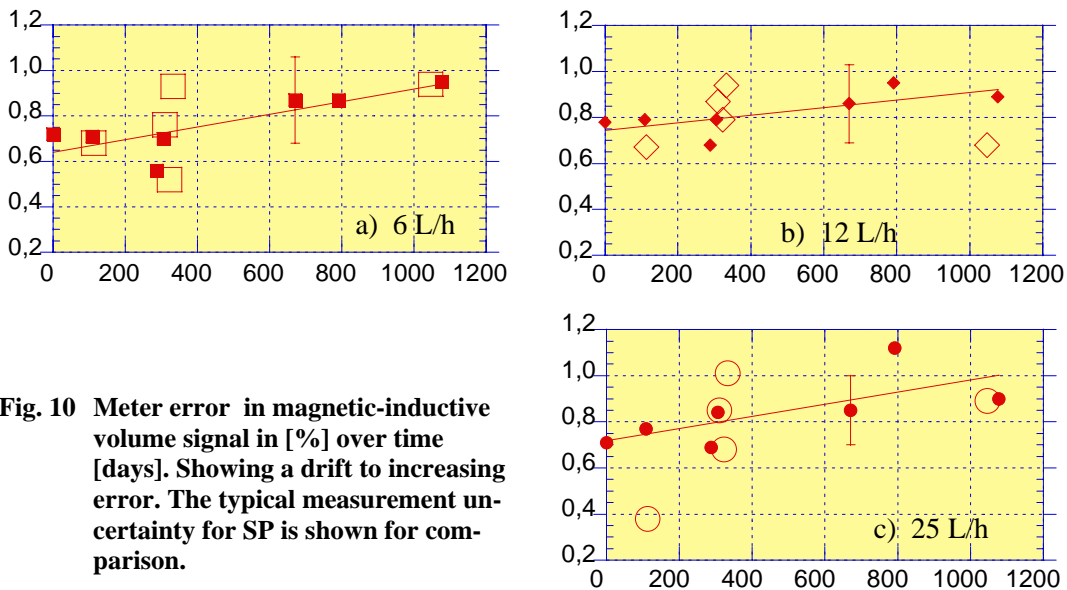


Fig. 10 Meter error in magnetic-inductive volume signal in [%] over time [days]. Showing a drift to increasing error. The typical measurement uncertainty for SP is shown for comparison.

This drift is most pronounced at the largest flow rate with 0,32 % and least for the middle flow rate with 0,16 % over the time interval of the comparison. Although the intermediate calibrations were performed at temperatures between 43 and 50 °C the spread is not caused by the varying temperature and neither is the change caused by zero point drift. Figure 10 also shows that, except for the middle flow rate, the inter-comparison data match that trend. The result that would be mostly affected of the meter drift is the last one from LEI. However, it does not seem justifiable to apply a trend correction. Only for the lowest flow rate (figure 10 a) a drift correction with -0,2 % would bring the results closer to each other and improve the over all equivalence.

An immediate conclusion is that the less stable Endress and Hauser Mag-meter constitutes a less good transfer standard when compared to the Danfoss coriolis meter. A degree of equivalence should therefore be based on the coriolis meter in the first place.

3.5 Comparison of results

The most important outcome of the comparison is the closeness of each result to the reference value, which is considered the best representative for the meter and flow rate in question. This value can be compared with the stated uncertainty from each laborator. Two measures can be used to characterize the result, the degree of equivalence and the En-value.

3.5.1 Degree of equivalence

The degree of equivalence DoE or simply d_i is the difference of one comparison result to the reference value. It is defined by equation (9). The index i counts the laboratory, m the flow meter in question and q stands for the actual flow rate. Thus this comparison delivers a number of DoE-values (se table 6). If possible the Mutual recognition Arrangement (MRA), assigned treatment between many countries to accept each others measurement results, also asks for a DoE with respect to a key comparison reference value (KCRV). For water at 50 °C no such value yet exists. The corresponding uncertainty $U(\text{DoE})$ or $U(d_i)$ in this difference from a reference value is defined by equation (10), where E_i and E_{CRV} are the uncertainties estimated by each laboratory and the calculated reference value for the comparison respectively.

$$d_i = E_i(m, q) - E_{\text{CRV}}(m, q) \quad (9)$$

$$U(d_i) = 2 \cdot \sqrt{u^2(E_i(m, q)) + u^2(E_{\text{CRV}}(m, q))} \quad (10)$$

It should be observed that U refers to a 95 % confidence level, whereas u relates to a standard deviation level, which in case of a normal distribution means a 68 % confidence level. The factor 2 (strictly statistically 1,96) is the transformation between both levels in case we assume a normal distribution.

3.5.2 En-values

Alternatively of using two separate measures d_i and $U(d_i)$ a comparison can be based on the En-value, which combines both. The En-value is the difference between a single result E_j and its reference value E_{ref} normalized over the uncertainty (at 95 % confidence) defined by equation (11).

$$E_n = \left| \frac{E_j(q_k) - E_{\text{ref}}(q_k)}{\sqrt{U^2(E_j) + U^2(E_{\text{ref}})}} \right| \leq 1 \quad (11)$$

The interpretation is the following. As long as the uncertainty in the difference, which is the combination of the two uncertainties of E_j and E_{ref} , is larger than the difference itself one cannot really state a difference and thus E_j and E_{ref} are equivalent in a metrological sense.

3.5.3 Tabulated results for DoE and En-value

Table 6 below contains the degree of equivalence for each laboratory at the three flow rates together with the corresponding uncertainties (blue figures). Beneath the comparison reference values are displayed in bold and their corresponding uncertainties.

Table 6. Comparison of different measures to characterize the outcome of the round robin.

Danfoss massa sign.	DoE			U(DoE)			En-value		
	6 L/h	12 L/h	25 L/h	6 L/h	12 L/h	25 L/h	6 L/h	12 L/h	25 L/h
BEV	-0,047	-0,028	-0,088	0,15	0,05	0,03	0,25	0,19	0,58
SP	-0,087	-0,048	-0,058	0,17	0,14	0,11	0,42	0,24	0,32
SMU	-0,007	0,122	0,152	0,21	0,10	0,09	0,03	0,71	0,87
Metas	0,013	0,202	0,272	0,12	0,14	0,17	0,08	1,02	1,20
LEI	0,103	-0,168	-0,038	0,15	0,11	0,11	0,55	0,93	0,21
REF VALUE	0,507	0,338	0,138	0,077	0,1	0,106			
Danfoss volume sign.	DoE			U(DoE)			En-value		
	6 L/h	12 L/h	25 L/h	6 L/h	12 L/h	25 L/h	6 L/h	12 L/h	25 L/h
BEV	0,027	0,061	-0,012	0,15	0,08	0,09	0,13	0,46	0,09
SP	-0,033	0,021	0,028	0,17	0,15	0,13	0,16	0,11	0,17
SMU	-0,153	-0,039	0,008	0,17	0,13	0,11	0,72	0,23	0,05
Metas									
LEI	0,157	-0,119	-0,022	0,17	0,15	0,16	0,74	0,64	0,12
REF VALUE	0,413	0,219	-0,058	0,094	0,072	0,067			
E & H volume sign.	DoE			U(DoE)			En-value		
	6 L/h	12 L/h	25 L/h	6 L/h	12 L/h	25 L/h	6 L/h	12 L/h	25 L/h
BEV	-0,143	-0,091	-0,445	0,16	0,09	0,03	0,61	0,70	2,46
SP	-0,063	0,109	0,025	0,15	0,16	0,08	0,28	0,60	0,13
SMU	-0,303	0,029	-0,145	0,18	0,16	0,03	1,21	0,16	0,80
Metas	0,107	0,179	0,185	0,20	0,17		0,41	0,93	1,06
LEI	0,117	-0,081	0,065	0,16	0,16	0,11	0,50	0,44	0,31
REF VALUE	0,825	0,761	0,823	0,12	0,068	0,126			

The right part of table 6 shows the calculated En-values. It should be noted that two different perspectives are applied concerning the uncertainty estimation. For the degree of equivalence, being based on the weighted mean in the first place, the uncertainty of the reference value is subtracted (due to correlation between CRV and all results). This leads to very small uncertainties and in one case to a disappearing one. For the En-value the uncertainties of the CRV and the single results are assumed independent (+ sign). This has considerable influence on the En-values. Five results (indicated in red bold) exceed the stipulated limit of 1.

3.6 Youden Plots

An empirical analysis tool to look at tandem measurements was suggested by Youden in 1959 for studying systematic differences between results [2]. In figures 11 and 12 the DoE's from table 6 are plotted for one meter along the x-axis and a second meter (or signal) along the y-axis. The black diamond in the right upper corner of figure 11 for example represents the values 0,272 from the Danfoss mass signal on the x-axis and the value 0,185 from the E&H- signal on the y-axis belonging to the flow rate 25 L/h. These two values are marked in table 6 with a ring. In that way the two reference values form the centre for each axis. Normally this type of cross-correlation plots are constructed symmetrically to the arithmetic means and the origin of coordinates is also the "centre of gravity" for all points. In this case with a reference value formed as a weighted mean or an average

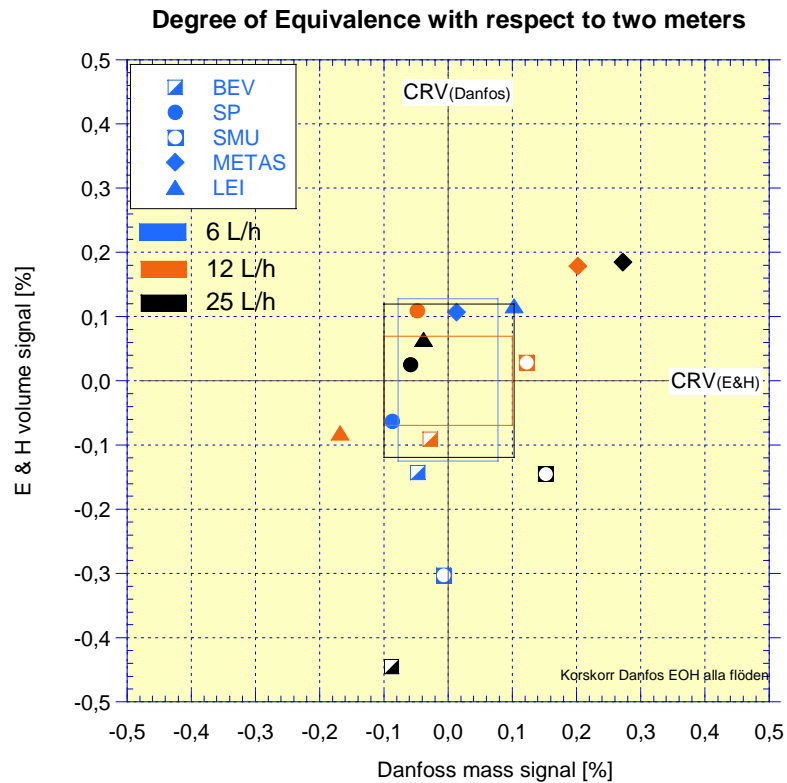


Fig. 11 Cross-correlation plot showing the simultaneous degree of equivalence for the Coriolis- and Mag- meter. The rectangles represent the uncertainty in the reference values.

median this need not be the case. Here the results from all three flow rates are presented simultaneously, but separated by colour (light blue 6 L/h, red 12 L/h, black 25 L/h). The different laboratories are characterized by different symbols. The uncertainty of the respective reference value is indicated by the coloured rectangles.

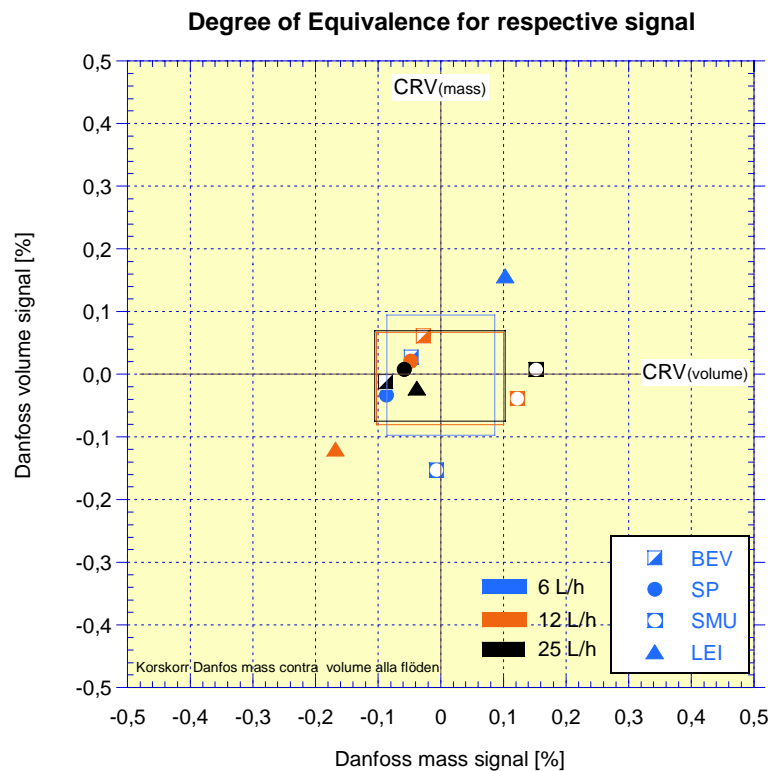


Fig. 12 Cross-correlation plot showing the simultaneous degree of equivalence with respect to two signals from the Coriolis meter. The rectangles represent the uncertainty in the reference values.

Of 15 results in figure 11 two are found outside a $\pm 0,2$ % range both from the $CRV_{(Danfoss)}$ and $CRV_{(E\&H)}$. Two results lie outside a 0,3% and a 0,4 % range with respect to one but within a 0,1 % range with respect to the other meter. One result is on the 0,2 % border for one meter and close for the other. Two third of the results are within a $\pm 0,15$ % range. The plot shows that the spread in results is larger for the volume signal of the Mag-meter than for the mass signal of the coriolis meter, but there is no indication for a systematic error behaviour.

In figure 12 results from Metas are missing as it was not possible to measure three signals simultaneously. The 12 remaining DoE's lie clearly within $\pm 0,2$ % and there does not seem to be a systematically shifted result. The results from this comparison are very similar to an other hot water calibration inter-comparison performed amongst 10 Nordic laboratories [3], that gave a very good agreement.

4 Interpretation of results and conclusions

The inter-comparison carried on almost two and a half years between the first and last measurement. This was caused by the hope to involve more participants. But in the end only five laboratories took part. During this time several intermediate calibrations were performed to follow an eventual meter drift. Within the measurement uncertainty no real drift for the coriolis meter can be stated. For the Mag-meter an increase of the error could be seen at all flow rates ($\sim 0,2$ %). However, the results from the last laboratory did not indicate a systematic drift. Thus no drift correction was applied. A zeroing of the meters, as often recommended by the suppliers, would perhaps have helped, especially concerning the temperature sensibility at low flow rates. A zeroing procedure with hot water is critical because both temperature and flow need to be absolutely stable, which in practice is almost impossible to achieve. The actual meters have never been zeroed since the first time usage in order to keep an unbroken record. This strategy is based on earlier experience.

The design of the comparison simulating real calibration conditions did not cause any problem for the staff at the different laboratories. At some places it was difficult to keep a stable temperature of 50 °C, which probably increased the experimental spread in the results.

The uncertainty claims are of comparable size between the laboratories. They tend to be larger for a flow rate of 6 L/h and smaller at 25 L/h. But as figure 7 and table 6 show the largest DoE's (differences to the reference value) are found at the highest flow rate, which seems to indicate a general underestimation of the uncertainty. It is reasonable to assume that too little concern is given to evaporation and temperature corrections.

Other comparisons have shown that it is not likely to achieve total consistency or agreement in results with only 5 participants. Of the nine partial results (three signals and three flow rates) agreement with respect to uncertainty was achieved in five cases and thus the weighted mean could form the best reference value. For the rest a median was used calculated by 25000 Monte Carlo Simulation trial. Looking to the data, altogether 70 partial results, three clear outliers were found giving En-values exceeding 1 and five values are just on the border.

Surprisingly the volume signal of the coriolis meter, which often is not used at all, seems to give the lowest spread and best overall agreement between the laboratories, although one would expect extra uncertainty contributions from the usage of the density signal.

A direct comparison with the laboratories CMC's (calibration measurement capability) is not possible. Only Metas, BEV and SP cover with their values the actual temperature and flow range specifying 0,2 %, 0,05 % and 0,2 % respectively. LEI has no value for elevated temperatures and the CMC-value of 0,12 % applies for flow rates above 400 L/h. SMU specifies a value of 0,12 % but for flow rates above 20 L/h. From this comparison one can conclude that the participating laboratories are capable to calibrate heat meters with good reproducibility at 50 °C and at these low flow rates within a 0,15 to 0,25 % uncertainty margin.

5 References

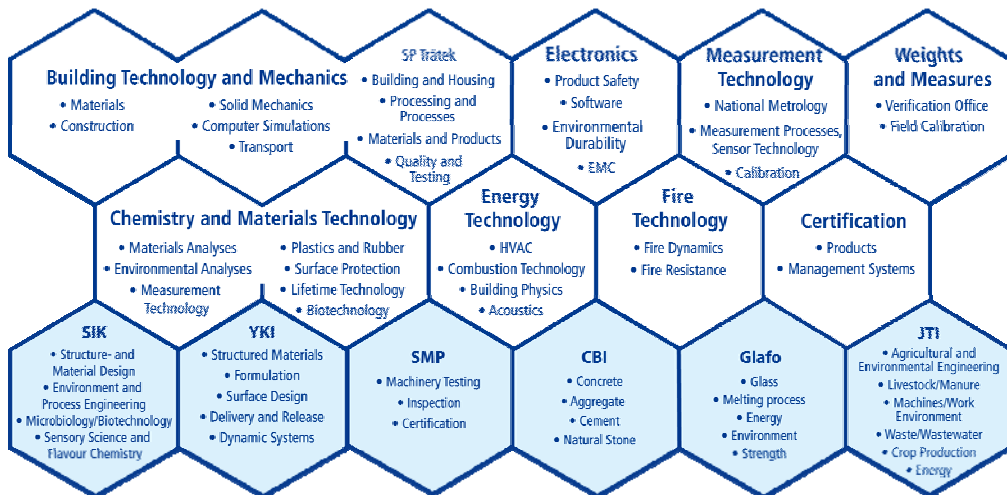
- [1] Cox M., "The evaluation of key comparison data", *Metrologia*, 2002, 39, 589-595
- [2] Youden, W.J., *Industrial Quality Control*, Vo.XV, No.11, May 1959 Graphical Diagnosis of Interlaboratory Test Result.
- [3] P. Lau, K. Stolt, Inter-comparison on water-/heat meters at 50 °C and 1-20 m³/h, EUROMET Project No. 863, SP Report: 2007:01

Appendix 1 - Application form

PROPOSED EUROMET PROJECT		
1. Ref. No.:	2. Subject Field: Flow	
3. Type of collaboration:	Inter-comparison of measurement	
4A. Partners:	BEV, PTB, SMU, SP	4B. CEC funded?
5. Participating countries:	AT, DE, SK, SE	
6. Title:	Inter-comparison on Water/Heat meters at 50 °C, 6-25 l/h	
7. Description:	<p>The meter package consists of two meters; one Mag-meter and one Coriolis-meter. This proposal is initiated by the common wish of the four laboratories to compare their calibration results in reference to a well established pair of meters. The meters have been frequently used for international assessment visits at several test facilities for hot water and heat meters in Europe. As there is no financial support for this project is not planned as a usual round robin. Instead the calibrations should be performed at occasions when the participating laboratories are visited by the pilot for other reasons. Due to the sensitivity of the meters and the value of their long calibration history, the pilot wishes to transport and be present at least when mounting and dismounting the meters and also to receive preliminary data on site. The allowed time for a calibration is maximum two days.</p> <p>The meter package consists of a Mag-meter followed of a Coriolis-meter working in the flow range 6-25 l/h.</p> <p>Proposed test points are 6, 12 and 25 l/h.</p> <p>The number of pilot measurements at SP will depend on the final travelling scheme, but at least 2 measurements are planned.</p> <p>No instructions are given in advance. The pilot plans to provide necessary information on site. A traditional evaluation will be performed and a draft distributed before finalizing the report.</p> <p>A time schedule cannot be given, but the pilot expects the exercise to be finished latest in December 2006.</p>	
8. Additional remarks:	<p>The project was proposed by the subgroup for water at the Euromet meeting in March 2005.</p> <p>The meters will be transported by car between the participating laboratories.</p>	
9. Proposer´s name:	Krister Stolt	
	<p>Address: SP – Swedish National Testing and Research Institute Measurement Technology Box 857 SE-501 15 BORÅS Sweden</p> <p>Telephone: ++46 33165482 Fax: ++46 33106973 E-mail: krister.stolt@sp.se</p>	
10. Proposer´s signature:	11. Date: 2005 10 28	12. Proposed starting date: 2006

SP Technical Research Institute of Sweden

Our work is concentrated on innovation and the development of value-adding technology. Using Sweden's most extensive and advanced resources for technical evaluation, measurement technology, research and development, we make an important contribution to the competitiveness and sustainable development of industry. Research is carried out in close conjunction with universities and institutes of technology, to the benefit of a customer base of about 9000 organisations, ranging from start-up companies developing new technologies or new ideas to international groups.



SP consists of eight technology units and six subsidiary companies. Three of the companies, CBI, Glafo and JTI are each 60 % owned by SP and 40 % by their respective industries.



SP Technical Research Institute of Sweden

Box 857, SE-501 15 BORÅS, SWEDEN

Telephone: +46 10 516 50 00, Telefax: +46 33 13 55 02

E-mail: info@sp.se, Internet: www.sp.se

www.sp.se

Measurement Technology

SP Report 2009:49

ISBN 435-237 8869

ISSN 0284-5172