INTERNAL AND EXTERNAL COMPARISON MEASUREMENTS FOR ENSURING ACCURACY AND TRACEABILITY IN FLOWMETER CALIBRATION

Walter Poeschel
Physikalisch-Technische Bundesanstalt (PTB)
Braunschweig, Germany

ABSTRACT

Comparison measurements become more and more important to perform traceability and to confirm the accuracy of flowmeter calibration facilities. But intercomparisons between different laboratories are very complicated, expensive and time consuming. Therefor a method is proposed to prove different components of the total measurement uncertainty by internal comparison measurements between different flow measurement devices within the laboratory.

The paper describes this method and its realization in the new primary standard for water flow measurement at PTB (being under construction now). This test rig represents a combined double type standard containing a gravimetric and a volumetric part, connected by a set of turbine meters in parallel, and offers the possibility for carrying out comparison measurements between two metrologically independent measurement devices, a gravimetric and a volumetric one. The set of turbine meters is needed to overcome differences in volume and flowrate ranges. The influence of the turbine meters on the uncertainty of comparison measurements (it is expected to be smaller than 0.01 %) was investigated using a special test installation. The results of various measurement series using this test installation are reported.

INTRODUCTION

Today, traceability in flowmeter calibration means traceability of volume, mass, density, temperature and time to the corresponding national standards in connection with an error analysis. But often the conventional error analysis, i. e. carrying out static tests of the main elements and assessing some additional error influences, does not lead to a realistic evaluation of the measurement uncertainty of flowmeter calibration devices, in particular when high accuracies are claimed. Even the strict application of the ISO-Guide [1] does not solve this problem. Indeed, the rules of the Guide offer a good basis for an unified treatment of known error influences, but whether all essential influences (especially those caused by the dynamic behavior of the measurement process) were taken into account and evaluated realistically remains as an unsolved question.

For this reason intercomparison measurements to confirm the stated uncertainties become more and more important and are actually carried out much more than in the past. But the intercomparisons carried out so far with the help of flowmeters as transfer standards show mainly two disadvantages:

- The long-term stability of the available transfer standards is not adequate to the needs for confirming test rig uncertainties of 0.1 % or better, which are claimed for accredited test laboratories.
- It is very complicated, expensive and time consuming to prove the total uncertainty of test facilities by intercomparisons extensively.

Therefor a method is proposed to subdivide the total uncertainty into different components which can be investigated partly by internal comparison measurements within the same laboratory (e. g. short-term, medium-term and long-term stability) [2], so that only constant systematic deviations between different laboratories and the influence of different flow conditions on the meter under test have to be checked by (external) intercomparisons. This method will be realized by the new primary standard for water flow measurement at PTB which is under construction now [3].

EVALUATION OF UNCERTAINTY

Application of the ISO-Guide

In the “Guide to the Expression Uncertainty in Measurement” [1] two methods for evaluating standard uncertainties are named: Type A, using the statistical analysis of a series of experimental data, and Type B, using other means for evaluating the uncertainty. Mostly, a combined method of both types is used on the basis of a mathematical model of the measurement process.

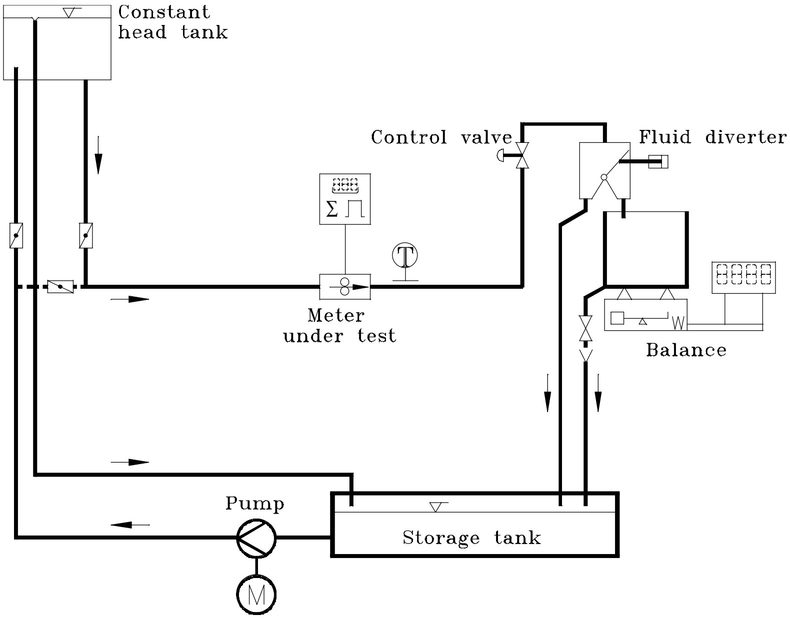

Fig. 1 Gravimetric flowmeter calibration facility

As an example a gravimetric calibration facility for liquid flowmeters is shown in Fig. 1, using a balance to measure the liquid quantity by static weighing and a fluid diverter for starting and stopping the measurement. The measurement process is started by diverting the flow into the weighing tank and starting the registration of the output of the meter under test and the time measurement. During the filling time of the weighing tank the liquid temperature and density are recorded continuously. By switching back the diverter, recording of meter output and time is stopped. Using all necessary corrections, the liquid volume V_M passed through the meter is calculated and the calibration factor of the meter under test.

The mathematical model of the process is given by Formula 1:

$$V_M = \frac{m_W + C_W}{\rho_{dest}\left(\vartheta_{LM}\right) + C_{\rho L}} C_B \left[1 + C_\vartheta + C_C + C_{Div}\right] + C_P + C_{ev} + C_G + \delta\, V_M \tag{1}$$

V_M	*Volume passed through the meter under test*
m_W	*balance indication*
C_W	*correction of balance indication*
ϑ_{LM}	*temperature at the meter under test*
$\rho_{dest}\left(\vartheta_{LM}\right)$	*density of distilled water at* ϑ_{LM}
$C_{\rho L}$	*density deviation from distilled water*
C_B	*Buoyancy correction*
C_ϑ	*correction for deviations of the thermometer indication from the real average water temperature*
C_C	*compressibility of water*
C_{Div}	*correction for diverter errors*
C_P	*correction for thermal expansion of the pipe between meter under test and diverter*
C_{ev}	*correction for volume loss by evaporation and splashing (neglected)*
C_G	*correction for gas within the water (neglected)*
$\delta\, V_M$	*repeatability of* V_M

This equation is the basis for estimating the uncertainty components $u_i(V_M)$ caused by the uncertainties $u(x_i)$ of the various influence quantities x_i

$$u_i\left(V_M\right) = \frac{\partial V_M}{\partial x_i} u(x_i) \tag{2}$$

and for calculating the combined uncertainty $u_c(V_M)$ of the measurement result (supposed the . $u_i(V_M)$ are uncorrelated):

$$u_c\left(V_M\right) = \sqrt{\sum_{i=1}^{n} u_i^2\left(V_M\right)} \tag{3}$$

Only the last term δV_M in Formula 1, the repeatability of V_M, can be treated by statistical methods (called Type A in the Guide), where a series of repeated measurements is analyzed and its mean and standard deviation calculated. Most of the uncertainty components must be evaluated by theoretical considerations and by non-statistical estimation of the limits of the effective influence quantities (called Type B method). That means, the result of such an error analysis (the combined uncertainty) is strongly affected by the personal understanding and experience of the scientist concerned.

For this reason the Guide gives an additional recommendation: “Because the mathematical model may be incomplete, all parameters should be varied to the fullest practicable extent so that the evaluation of uncertainty is based as much as possible on observed data. Whenever feasible, the use of empirical models of the measurement procedure founded on long-term quantitative data ... should be part of the effort to obtain reliable evaluations of uncertainty.” (see [1], par. 3.4.2). In other words: Not only the short-term standard deviation but also medium- and long-term standard uncertainties should be investigated and calculated from repeated and comparing measurements. To derive a value for the overall uncertainty from the results of such experimentally obtained data, a new concept is necessary for subdividing the total uncertainty into components mentioned above, based on a different (empirical) model of the measurement process. A proposal is given in the next paragraph.

A New Approach to a Realistic Evaluation of Uncertainty

In opposite to Formula 1, which represents the physical dependence of the measurement result from different measurement and influence quantities, the following empirical model of the measurement process can be drawn up:

$$V_M = V_M' + \delta_{sh}(V_M) + \delta_{m/sh}(V_M) + \delta_{lg/m}(V_M) + \delta_{ex/lg}(V_M) \qquad (4)$$

V_M	value of the measurand
V_M'	measurement result (including all known corrections)
$\delta_{sh}(V_M)$	short-term instability component (repeatability)
$\delta_{m/sh}(V_M)$	instability component, which is constant concerning short-term repeatability conditions but variable concerning medium-term conditions (e. g. day-to-day or week-to-week measurements during one recalibration period)
$\delta_{lg/m}(V_M)$	instability component, which is constant concerning medium-term conditions but variable concerning long-term conditions over several recalibration periods in the same laboratory
$\delta_{ex/lg}(V_M)$	component, which is constant concerning long-term conditions within the laboratory but variable in (external) comparison measurement programs between different test facilities and laboratories

The corresponding variances are forming the combined total measurement uncertainty $u_c(V_M)$ in the following way:

$$u_c^2(V_M) = u_{sh}^2(V_M) + u_{m/sh}^2(V_M) + u_{\lg/m}^2(V_M) + u_{ex/\lg}^2(V_M) \qquad (5)$$

In this formula the sign $u_{sh}(V_M)$ is standing for $u(\delta_{sh}(V_M))$, $u_{m/sh}(V_M)$ for $u(\delta_{m/sh}(V_M))$ and so on. The variance component for measurements carried out during medium-term conditions is then

$$u_m^2(V_M) = u_{sh}^2(V_M) + u_{m/sh}^2(V_M) \qquad (6)$$

and the long-term component:

$$u_{\lg}^2(V_M) = u_{sh}^2(V_M) + u_{m/sh}^2(V_M) + u_{\lg/m}^2(V_M)\,. \tag{7}$$

Except $u_{ex/\lg}^2(V_M)$, the variance components on the right side of equation (5) can be determined by internal comparison measurements, supposed a comparison measuring device is available which short-, medium- and long-term stability is as good as or better than those of the test facility to be investigated. Carrying out an appropriate measurement program over longer time and calculating the results with the help of the method of analysis of variance [4], the short-, medium- and long-term uncertainty components can be separated [2]. However, the variance component $u_{ex/\lg}^2(V_M)$ have to be investigated by external comparison measurements in an intercomparison program where different laboratories are involved.

For comparison purposes the different variance or standard uncertainty components can also be estimated by the method of analysis of error sources, based on Formula 1, and it is recommended to do so. In this case not only the limits of the different influence quantities have to be estimated, but also whether they are constant or variable with respect to the different test conditions mentioned above. Comparing the results of both methods is the best way for obtaining reliable evaluations of uncertainty. A practical example and the results of such a comparison for an experimental test installation are presented in [2].

PRACTICAL APPLICATIONS

Turbine Meters as Internal Comparison Meters

Specially chosen flowmeters (e. g. turbine meters, magnetic-inductive meters or other types), used as transfer standards for intercomparisons, can also serve as comparison meters for internal measurement programs to investigate uncertainty components. In particular turbine meters are suitable to determine the repeatability component of the uncertainty of flowmeter calibration facilities. With good exemplars of turbine meters relative standard deviations (short-term) can be achieved of smaller than 0.01 %. Using a tandem arrangement of two turbine meters in series, the component of the standard deviation, caused only by the test facility, can be separated by a special method [5]. In this way the short-term component of the relative standard uncertainty of a gravimetric test facility at PTB was determined to an average value of 0.004 %.

But the results of this measurements reported in [5] also showed, that the same method cannot be applied just like that for determining the medium-term uncertainty component of a high accuracy calibration facility. The reason is, that on different days correlation effects occur between the two meters in series, possibly an effect of variations in the ambient conditions, affecting the two meters in the same way. For long-term investigations of water standard facilities with higher accuracy turbine meters are also not a good choice. Used for measuring water, over longer time variations of the meter factor up to 0.2 % were observed.

The Concept of a Combined Double Type Standard

For carrying out high accuracy comparison measurements within a laboratory over longer time with the aim to determine uncertainty components in the magnitude of 0.01 %, therefore a combined double type standard was proposed and tested [2]. This concept will be realized in the new primary standard for water flow measurement at PTB Braunschweig being under construction presently [3].

The standard will comprise two metrologically independent standards of different types (a gravimetric system and a pipe prover as a volumetric one) both on similar accuracy level. The two standards are connected hydraulically and metrologically in such a way, that measurements of the same quantity (e. g. the liquid volume passed through a flowmeter in the test section) can be made by two different methods at the same time, so that two independent measurement results of the same quantity are available. The relative difference of the two results is a valuable objective criterion whether the measurement process is under statistical control or not. If yes, the uncertainty components defined above (related to the mean value) can be determined and if not, the reason for this have to be investigated. The use of two very different types of standards (which are affected by different error influences) makes it also possible to detect hardly recognizable errors, e. g. air bubbles in the liquid.

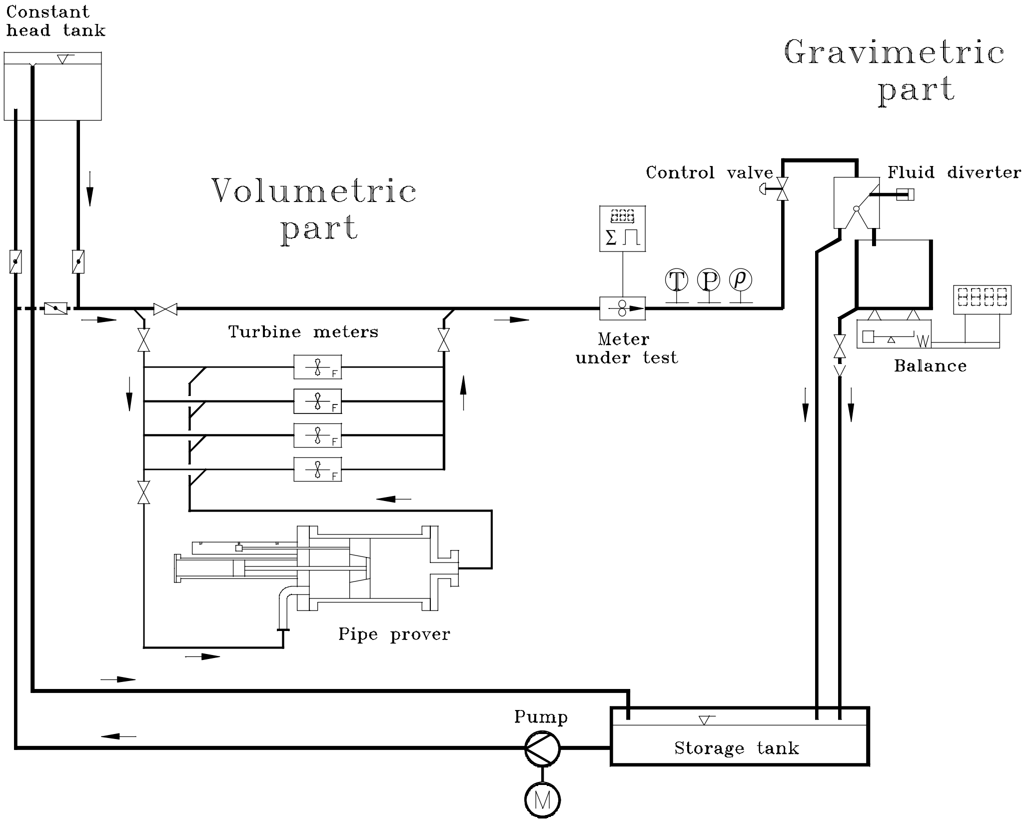

Fig. 2 High accuracy combined flow standard with volumetric and gravimetric part

Fig. 2 shows a simplified scheme, Fig. 3 a cut-away view and Table 1 the main specifications of PTB's new water test rig being under construction. The supply system with a constant head tank, the two standard devices, a volumetric and a gravimetric one, and between them the test section with the meter under test are seen. The volumetric part consists of a pipe prover (compact or small-volume prover) and a set of turbine meters in parallel, the gravimetric part has three balances of different sizes and three flow diverters.

Table 1. Claimed specification of the combined water flow standard

	Gravimetric system	**Pipe-prover**
Flowrate	0,3 m^3/h...2100 m^3/h	0,3 m^3/h...550 m^3/h
Measured quantity	75 kg...30.000 kg	250 litres
Pressure	up to 5 bars	up to 5 bars
Standard uncertainty	0.01 %	0.01 %
Standard deviation (repeatability)	0.004 %	0.004 %

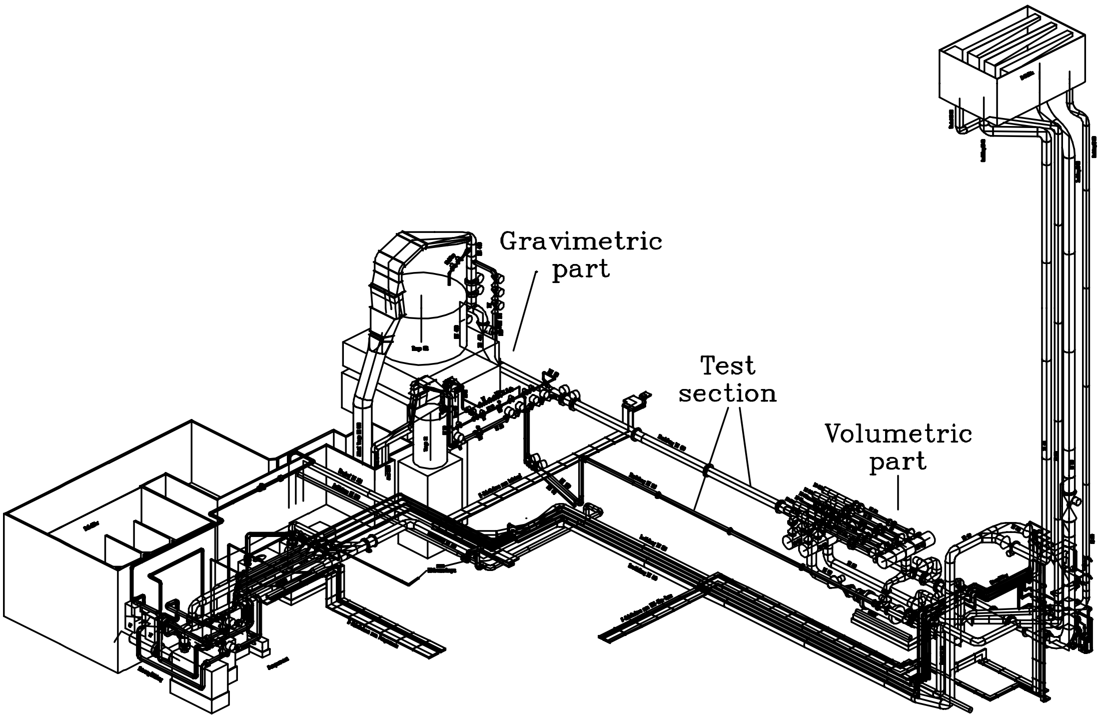

Fig. 3 Cut-away view of PTB's new calibration plant

The volume and flowrate ranges of the two standard devices are different (see Table 1). Therefor a set of four turbine meters in parallel is necessary to perform comparison measurements between the pipe prover and the gravimetric system throughout the whole volume and flowrate range of the test rig. Fig. 4 shows a scheme of this meter set connected in series to the prover and Fig. 5 the corresponding section of Fig. 3.

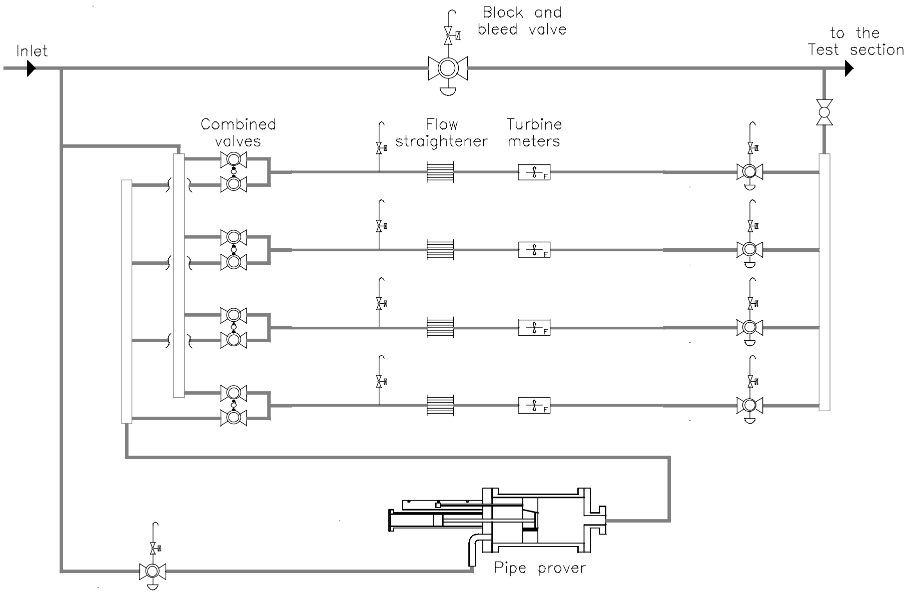

Fig. 4 Pipe prover in series with a set of 4 turbine meters

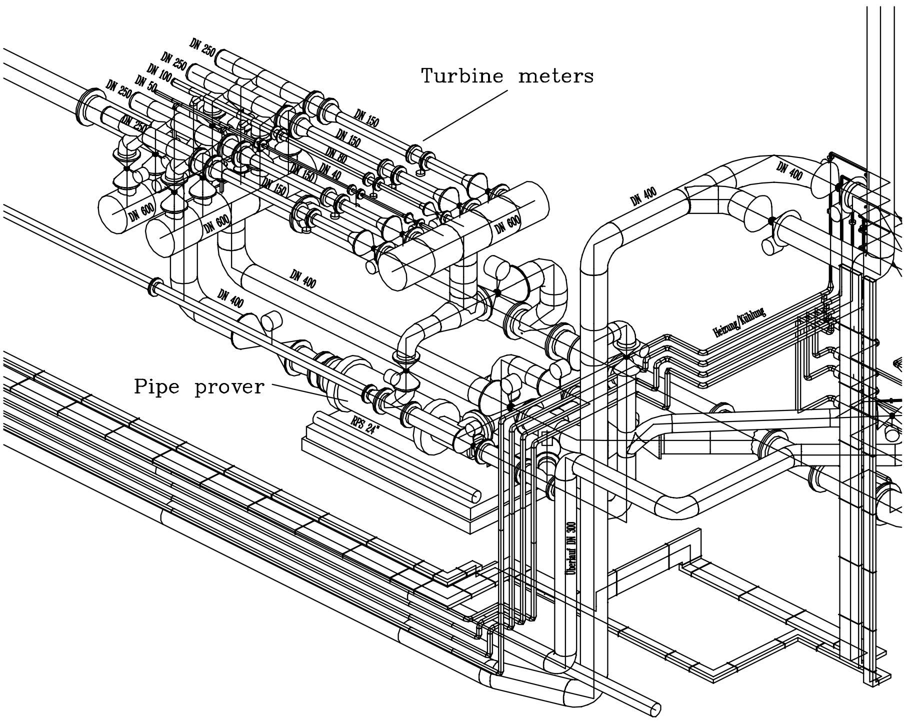

Fig. 5 Detail cut-away view: pipe prover and turbine meters in parallel

When a comparison measurement shall be carried out at higher flowrate, the liquid flow is directed through two, three or four turbine meters in parallel before the flow is passing on to the test section and the gravimetric system. Without changing or braking down the flowrate, each turbine meter can be connected with the pipe prover by switching over the corresponding combined valve (see Fig. 4). While keeping flowrate and the other parameters constant (repeatability conditions) the following comparison procedure is carried out several times:

At first the calibration factors of the turbine meters are determined on the basis of a comparison with the pipe prover by connecting successively each meter with the prover. Then (or partly at the same time, if possible), using this topical calibration factors, the sum of the volume output of the turbine meters is compared with the volume value measured by the gravimetric system. The relative difference is a measure for the error difference of the volumetric and the gravimetric standard, provided that random and systematic error influences caused by the turbine meters can be neglected with respect to this comparison procedure. This assumption is a critical matter for the application of the proposed method. Therefor a test installation with three turbine meters in parallel was set up and random and systematic errors investigated experimentally. This is described in the next paragraph.

Tests on turbine meters for internal comparison measurements

Test installation. A scheme of the test installation is shown in Fig 6. The set of three turbine meters in parallel was connected to the test section of the existing gravimetric water test facility of PTB. The pressure loss of the planned pipe prover was simulated by a control valve.

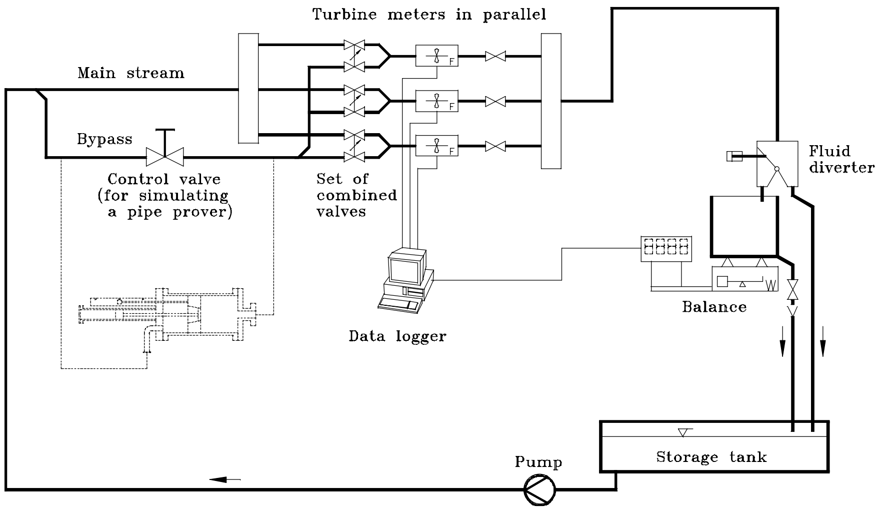

Fig. 6 Test installation with 3 turbine meters in parallel

The design of the meter set for testing, shown in Fig. 7 was similar to that which is planned for the new primary standard. The nominal diameter of the turbine meters were 50 mm. The combined valves were ball valves with a joint drive. They fed the upstream pipework of the turbine meter either from the main stream or from the by-pass exactly in the same direction (see Fig. 7) to ensure uniformity of flow conditions upstream the meter in both switching positions.

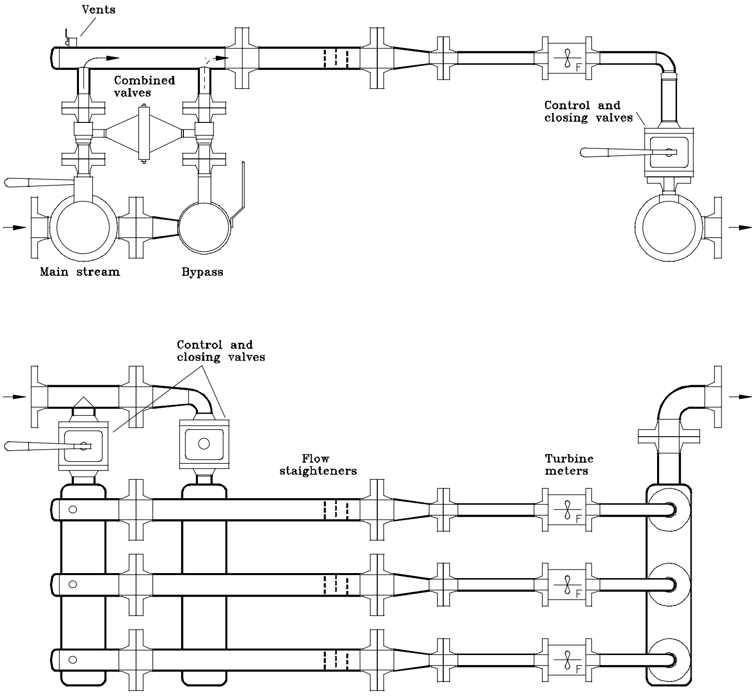

Fig. 7 Setup of turbine meter arrangement

Accuracy requirements. Internal comparison measurements between the volumetric and the gravimetric standard are intended to determine standard uncertainty components for medium- and long-term conditions which are expected to be smaller than the claimed overall standard uncertainty of 0.01 % (Table 1). Therefor the relative standard uncertainty of the comparison method u_{comp} (including a repeatability component u_r and a systematic component u_s caused by the turbine meters) should be also clearly smaller than that value.

A relative standard deviation for single comparison measurements of 0.01 % assumed, the standard deviation of the mean of 10 measurements (i. e. u_r) is about 0.003 %. More serious is a possibly occurring systematic error of the turbine meters as a result of varations in the flow conditions caused by switching the combined valves from main stream to by-pass and back. Supposed the relative systematic error is within the limits ±0.01 %, the corresponding standard uncertainty component u_s can be estimated (assuming equal probability - see [1], clause 4.3.7) as:

$$u_s = \frac{0.01}{\sqrt{3}}\% = 0.0058\%.$$

Then the relative standard uncertainty of the comparison method u_{comp} is

$$u_{comp} = \sqrt{u_r^2 + u_s^2} \approx 0.0065\%.$$

This may be acceptable, but the question is, can we really meet these requirements? To answer this question, a large number of measurements were carried out using the test installation described above to investigate repeatability and systematic effects in this special test arrangement.

Repeatability measurements. With an exemplar of turbine meter (nom. diameter 50 mm) 10 measurement series were carried out on different days. Table 2 shows the results. The relative standard deviations are between 0.003 % an 0.008 %, the average value is 0.0055 %. This is a very good result and shows, that a short-term repeatability better than the expected limit of 0.01 % can be achieved constantly, when during the series of measurements the measurement conditions are constant and the results are distributed randomly. In other cases, e. g. when a drift occur from temperature changes or other reasons or when solid particles in the meter or in the flow straightener cause a slight variation of the flow conditions, we have to look for the reason, to remove it and to repeat the measurement series.

Table 2. Repeatability measurements with a turbine meter DN 50 at 30 m^3/h

Series No	Number of measurement	Mean value of meter error %	relative standard deviation %
1	30	0.053	0.0080
2	50	0,061	0.0043
3	40	0.052	0.0056
4	41	-0.149	0.0033
5	51	0.009	0.0072
6	50	0.015	0.0055
7	50	0.044	0.0051
8	40	0.071	0.0032
9	50	0.069	0.0029
10	50	0.065	0.0075

The measurement results in Table 2 gives also an information about the medium-term uncertainty by calculating the standard deviation of the mean values of measurement series carried out on different days over two weeks. Even when we remove the outlyer (No 4), the relative standard deviation of the mean values is 0.023 % and the medium-term component $u_m = 0.024$ %.

Testing systematic effects. An important supposition for an effective use of the planned arrangement of turbine meters in parallel is, that the flow conditions upstream the meters do not change significantly when the flow is diverted by one of the combined valves from the main stream to the by-pass (see Fig. 6), so that a systematic effect on the calibration factor of the turbine meters is negligible (or within the limits of ±0.01 %).

At first tests were made using only one turbine meter. Measurement series were carried out alternating at both positions of the corresponding combined valve ("main stream" and "by-pass"). The relative differences measured at both positions (mean values of 15 measurements) are given in Table 3. It can be seen, that all obtained differences are within the claimed limits of ±0.01 %, even those carried out without flow straightener. The average results with and without flow straightener differ slightly from each other (0.011 %).

Table 3. Effect of changing valve position from ”main stream” to ”by-pass” (see Fig. 6) on a turbine meter

Flowrate	Number of measurements	Relative differences ”main stream” – ”by pass”			
		with flow straightener		without flow straightener	
		mean value	stand.dev. of mean	mean value	stand. dev. of mean
m^3/h		%	%	%	%
15	15	0.008	0.0062	-0.008	0.0018
	15	0.007	0.0015	-0.008	0.0018
	15	0.010	0.0013	-0.004	0.0033
30	15	0.003	0.0031	-0.000	0.0038
	15	0.004	0.0007	-0.001	0.0020
	15	0.005	0.0014	-0.005	0.0018
60	15	0.007	0.0014	-0.005	0.0018
	15	0.005	0.0017	-0.008	0.0021
	15	0.006	0.0011	-0.002	0.0031
90	15	0.003	0.0018	-0.007	0.0038
	15	0.003	0.0022	-0.007	0.0025
	15	0.005	0.0013	-0.012	0.0028
	average	0.0053		-0.0055	

In another test program the three turbine meters in parallel were working together. The summarized indications of the meters were tested against the gravimetric system. Alternating the following measurements were carried out:

- all meters connected to “main stream” (reference value)
- first meter connected to “by-pass”, the other to “main stream”
- second meter connected to “by-pass”, the other to “main stream”
- third meter connected to “by-pass”, the other to “main stream”.

Table 4. Tests using 3 turbine meters in parallel (reference value: all meters connected to ”main stream”)

Flowrate	Number of measurements	Relative differences to reference value (mean values)		
		1. meter at ”by-pass”	2. meter at ”by-pass”	3. meter at ”by-pass”
m^3/h		%	%	%
43	10	-0.001	-0.003	0.002
	10	0	-0.003	-0.002
	10	0.001	0.004	0.001
	10	0.005	0.004	0.001
	10	0	0.002	0.002
66	10	-0.008	-0.008	-0.009
	10	0	-0.003	0
	10	0.001	0.001	0.001
	10	-0.001	0.006	0.002
	10	-0.001	0	-0.008
90	10	0.004	0.005	0.006
	(10	0.010$^{*)}$	0.016$^{*)}$	0.005)
	10	0.006	0.005	0.001
	10	-0.001	-0.007	-0.004
	10	0	0.001	-0.007
	average	0.0003	-0.0003	-0.001

$^{*)}$ bad series (standard deviations 0.05 %!)

Table 4 shows the relative differences of measurements carried out with one meter at ”bypass” (the other at ”main stream”) from the reference value (all meters at ”main stream”). Except one figure (result of a bad measurement series with standard deviations of 0.05 %!), all differences are within the limits of ±0.01 %, the total average is smaller than 0.001 %.

That means, an average bias caused by the different positions of the combined valves on the turbine meters can be considered as negligible relating to the used test installation. But for high accuracy comparing measurements always mean values of 10 or more measurement results should be used to meet the accuracy requirements mentioned above.

CONCLUTION

The proposed method to prove different components of total uncertainty by internal (within the laboratory) and external comparison measurements is suitable to confirm the stated uncertainty of flowmeter calibration facilities in an objective way. Realizing the concept of a combined double type standard, comprising two metrologically independent parts (a gravimetric standard and a pipe prover as a volumetric one), it will be possible to determine by internal comparison measurements the short-, medium- and long-term uncertainty component of PTB’s new primary standard for water flow measurement. A further uncertainty component (comparability between various laboratories) however, has to be investigated by external comparisons.

To perform global comparability in flow measurement an international flow standard should be established as a group of national standards liked by intercomparisons. On the basis of such an international standard a new kind of traceability could be realized: traceability in terms of the unit flowrate, back to a national flow standard which itself is a part of the international flow standard.

ACKNOWLEDGMENTS

The design of test installations by P. Mahr and G.Binder as well as carrying out a lot of measurements by G. Gusewski, U. Rehfeld and M. Becker is gratefully acknowledged.

REFERENCES

1. ISO International Organisation for Standardisation (1993/1995). Guide to the Expression of Uncertainty in Measurement. Geneva, Switzerland.
2. Poeschel, W. (1989). Investigation and Verification of the measurement Uncertainty of High Accuracy Flowmeter Calibration Devices by Internal Comparing Measurements. Report of the 5th International IMEKO-Conference on Flow Measurement FLOMEKO, Duesseldorf, Germany. *VDI-Berichte No 768*, p. 247-256, VDI-Verlag, Duesseldorf, Germany.
3. Poeschel,W. and Engel, R. (1998). The Concept of a New Primary Standard for Liquid Flow Measurement at PTB Braunschweig. Proceedings. The 9th International Conference on Flow Measurement FLOMEKO’98, p. 7-12, Lund, Sweden.
4. Sachs, L. (1974). Angewandte Statistik (Applied Statistics). Springer-Verlag Berlin, Germany.
5. Poeschel,W. (1995). Testing the Repeatability of Flowmeter Calibration Devices. Proceedings of the 2nd Brazilian Symposium on flow Measurement, p. 171-181, Sao Paulo S.A., Brazil.