The comparison of the gas flow secondary standard facilities at high pressure

Mengna Li[1], Bodo Mickan[2*], Chunhui Li[1*], Jia Ren[3], Yan Wu[4], Ming Xu[5]

[1]National Institute of Metrology (NIM), Beijing, China
[2]Physikalisch-Technische Bundesanstalt (PTB), Braunschweig, Germany
[3] Chengdu Natural Gas Sub-Station (Chengdu), China,
[4] Nanjing Natural Gas Sub-Station (Nanjing), China
[5] Wuhan Natural Gas Sub-Station (Wuhan), China)
E-mail (corresponding author): lich@nim.ac.cn; Bodo.Mickan@ptb.de

Abstract

The first formal comparison was organized in China for the purpose of determination of the degree of equivalence of the gas flow secondary standard facilities during 2016~2020. There were 4 participants from China, and PTB was invited as the link lab to connect this comparison with the serial key comparisons of CCM.FF.K5. Based on the flow range of the existing secondary standard facilities of participating laboratories, 3 turbine flow meters were used as transfer standards. Totally 298 sets of measured data were obtained with Reynolds number range of (6.3×10^4~1.4×10^7). Results of all participants were considered for the determination of the reference value and the uncertainty of the reference value. The operation conditions were represented by the Reynold number, while the measured value was represented by the relative error of the meter *e* in (%). The fitted curve based on the relationship between *e* and the Reynolds number for each single meter was obtained. The degree of equivalence of E_n was finally evaluated. Among all 298 sets of measured results, there were 282 sets of results with $E_n\leq1$, while there were 9 sets of results with $1< E_n\leq1.2$.

1. Introduction

Natural gas plays two important roles as the world transitions to a low-caron energy system: increasing the speed at which fast-growing emerging economies reduce their dependency on coal, and providing a source of low-carbon energy when combined with carbon capture, use and storage (CCUS) [1]. Natural gas is deeply involved with international trade, commerce, and regulatory affairs. To realize the fair trade of natural gas, it is necessary to ensure the accuracy of the flow meter in use by the regular verification or calibration with gas flow working standard facility, following the quantity value traceability.

CIPM Mutual Recognition Arrangement (CIPM MRA) is the framework through which National Metrology Institutes (NMIs) demonstrate the international equivalence of their measurement standards and the calibration and measurement certificates they issue [2]. The outcomes of the arrangement are the internationally recognized Calibration and Measurement Capabilities (CMCs) of the participating institutes. The technical basis of the CIPM MRA is the set of results obtained over the course of time through key comparisons. The serial key comparisons CCM.FF.K5, for the working standard facility of gas flow national standard of high pressure, were conducted during 2004~2012 [3~6] to fulfill the requirements of the CIPM MRA, which were all piloted by PTB. The turbine meters with nominal diameter of 150 mm to 300 mm were chosen as the transfer meters, and the degree of equivalence E_n was evaluated based on the meter deviation.

The first formal comparison was organized for the purpose of determination of the degree of equivalence of the primary and secondary standard facility for high-pressure gas flow measurement in China from 2016 to 2020. Following the successfully conducted primary standard facility comparison, the secondary standard facility comparison can further promote the development of natural gas and ensure the fair trade in natural gas field.

There were 4 participants from China, and PTB was invited as the link lab to connect this comparison with the serial key comparisons of CCM.FF.K5. The information of participants is shown in Table 1. Each laboratory completed the measurements and sent the transfer standards to the next laboratory.

Table 1: Participants information

Country	**Lab**	**Pressure range [kPa]**	**Working fluid**	**Date of calibration**
China	NIM	690~2500	Air	Sep. to Oct., 2019
Germany	PTB	2000~5000	Natural gas	Nov., 2016 to Sep., 2018

China	Chengdu	2000~4500	Natural gas	July, 2021
China	Nanjing	6500	Natural gas	Dec., 2020
China	Wuhan	5500~8500	Natural gas	May, 2019

2. The comparison schemes

2.1 Secondary standard facilities

The technical specifications of secondary standard facilities are presented in this section.

2.1.1 NIM (National Institute of Metrology, China)

There are 2 sets secondary standard facilities in NIM, including sonic nozzle facility and closed loop facility. The technical specification of the facilities is shown in Table 2.

Figure 1: Sonic nozzle facility and closed loop facility in NIM

Table 2: Technical specification of Sonic nozzle facility and closed loop facility in NIM

Facility	Sonic nozzle facility	Closed loop facility
Pressure	(200~2500) kPa	(200~2500) kPa
Temperature	(20±5) ℃	(20±5) ℃
Flow rate	(20~400) m^3/h	(20~1400) m^3/h
Uncertainty for MUT	≥0.15% (k=2)	≥0.18% (k=2)

2.1.2 PTB (Physikalisch -Technische Bundesanstalt)

The tests were conducted in Turbine flow meter facility with natural gas. The technical specification of the facility is shown in Table 3.

Figure 2: Turbine flow meter facility in PTB

Table 3: Technical specification of Turbine flow meter facility in PTB

Pressure	(1500~5500) kPa
Temperature	(20±5) ℃
Flow rate	(16~5600) m^3/h
Uncertainty for MUT	(0.13~0.16) % (k=2)

2.1.3 Chengdu (Chengdu Natural Gas Sub-Station)

The tests were conducted in the sonic nozzle facility with natural gas. The technical specification of the facility is shown in Table 4.

Figure 3: Sonic nozzle facility in Chengdu

Table 4: Technical specification of Sonic nozzle facility in Chengdu

Pressure	(400~5500) kPa
Temperature	(20±5) ℃
Flow rate	(5~5155) m^3/h
Uncertainty for MUT	≥0.16% (k=2)

2.1.4 Nanjing (Nanjing Natural Gas Sub-Station)

The tests were conducted in the sonic nozzle facility with natural gas. The technical specification of the facility is shown in Table 5.

Figure 4: Sonic nozzle facility in Nanjing

Table 5: Technical specification of Sonic nozzle facility in Nanjing

Pressure	(2000~6500) kPa
Temperature	(20±5) ℃

Flow rate	(20~2500) m^3/h
Uncertainty for MUT	≥0.22% (k=2)

2.1.5 Wuhan (Wuhan Natural Gas Sub-Station)
The tests were conducted in the turbine flow meter facility with natural gas. The technical specification of the facility is shown in Table 6.

Figure 5: Turbine flow meter facility in Wuhan

Table 6: Technical specification Turbine flow meter facility in Wuhan

Pressure	(2500~10000) kPa
Temperature	(20±5) ℃
Flow rate	(20~9600) m^3/h
Uncertainty for MUT	≥0.16% (k=2)

2.2 Transfer standard
Considering the flow range of the secondary standard facilities for all participants, three turbine flow meters with the range (20 to 1600) m^3/h are used as transfer standards. The information of transfer standards is shown in table 7.

Table 7: Technical specification of the transfer standards

Transfer standards	**DN [mm]**	**Flow range [m^3/h]**
Elster turbine meter	100	(20~400)
RMG turbine meter	100	(20~400)
Elster turbine meter	200	(80~1600)

3. Measurement results

There were totally 298 sets of measured results. The expanded CMC uncertainty of all the measured results was (0.13~0.23) % (k=2). The long-term stability and installation of the transfer meters are considered with the uncertainty u_{TM}=0.05%. The reported uncertainty for the comparison of a participant is $u_{test}^2 = u_{CMC}^2 + u_{TM}^2$.

3.1 Measurement result of Elster DN100
There were 112 measured points.

- NIM: There were 39 measured points, and the expanded uncertainty was (0.15~0.18) % (k=2).
- PTB: There were 19 measured points, and the expanded uncertainty was 0.13% (k=2).
- Chengdu: There were 16 measured points, and the expanded uncertainty was 0.19% (k=2).
- Nanjing: There were 8 measured points, and the expanded uncertainty was 0.23% (k=2).
- Wuhan: There were 30 measured points, and the expanded uncertainty was 0.16% (k=2).

3.2 Measured result of RMG DN100
There were 114 measured points.

- NIM: There were 39 measured points, and the expanded uncertainty was (0.15~0.18) % (k=2).
- PTB: There were 19 measured points, and the expanded uncertainty was 0.13% (k=2).
- Chengdu: There were 16 measured points, and the expanded uncertainty was (0.19~0.22) % (k=2).
- Nanjing: There were 8 measured points, and the expanded uncertainty was 0.23% (k=2).
- Wuhan: There were 32 measured points, and the expanded uncertainty was 0.16% (k=2).

3.3 Measured result of Elster DN200
There are 72 measured points.

- NIM: There were 22 measured points, and the expanded uncertainty was 0.18 % (k=2).
- PTB: There were 18 measured points, and the expanded uncertainty was 0.13% (k=2).
- Chengdu: There were 16 measured points, and the expanded uncertainty was 0.19% (k=2).
- Nanjing: There were 8 measured points, and the expanded uncertainty was 0.23% (k=2).
- Wuhan: There were 8 measured points, and the expanded uncertainty was 0.16% (k=2).

4. Comparison Evaluation

4.1 Evaluation procedure
The reference value was determined for each flow meter separately. Results of all participants were considered for the determination of the reference value and the uncertainty of the reference value [7,8].

The challenge for the evaluation of this comparison is caused by the complex situation where the data base and the data processing are necessary to satisfy the needs of statistical concepts. The conventional situation for comparisons in the flow community is:

➢ Only measurement values generated at the same operation point (regarding flow rate, pressure and gas) are compared, the so-called point-to-point evaluation. The reference value is therefore a single value for each operation point separately.

- Only independent (non-correlated) values are taken into account for the reference value. This requires each participant provides only one value for each operation point and the comparison is evaluated separately for each operation point.

The situation in this secondary comparison is different because the participants had made their measurements at the operation points according to their possibilities without a pre-specification of the number of measurements. Consequently, there is no fixed operation point to be used in the evaluation, and the reference value has to be determined by fit function of the operation conditions.

The operation conditions are represented in this comparison by the Reynolds number and the measurement value is represented by the relative error of the meter *e*. There is an available approach which can provide an overall fit of the relationship of relative error of the meter and Reynolds number: the model of PTB (see chapter 4.1.1 for a detailed description of the model).

Another fact which also has to be considered carefully is that there is more than one value of each participant to be used in the calculation. In this case, the condition of non-correlated data is not fulfilled anymore. The data of one lab have all common sources of uncertainty which are static over time. The task to evaluate the comparison data ends as a parameter optimization based on the generalized linear least square fitting of correlated data [9,10].

4.1.1 Evaluation model

The calibration data was fitted using the formula as

$$e_{Re} = \sum_{j=0}^{n} a_j \left[\log \frac{Re}{10^6}\right]^j \qquad (1)$$

where e_{Re} is the meter deviation that is dependent of the Reynolds number. The coefficients a_j (*j*=0... *n*) are obtained by curve fitting. The value of *n* is chosen from 1 to 4, depending on the best fit result. The factor 10^6 is introduced to obtain arguments of the logarithm that are around 1, which supports the numerical stability of the least-square approximation [11].

The basic principles for evaluation of comparison data sets with curves is explained in detail by Cox [12] using the Generalised Least Square (GLS). We discuss here only two points of our application which are different to the situation described in [12].

In [12], there have been fit curves calculated for the overall set of data as well as for each subset of the participants based on the chosen fitting model. In such case, it is possible to express the difference to the comparison reference value for each participant simply as a difference of curves. In our case, we refrain from this because our data set does not fulfil all conditions to apply this approach. Hence, we only calculate the comparison reference curve as an overall GLS to the complete data set for one meter. The difference d_i of each single measurement point to the reference value (curve) and its related uncertainty $u(d_i)$ rsp. Degree of Equivalence is then consequently the residual with its related uncertainty, and we focus on these residuals in our evaluation.

The GLS provides the so-called projection matrix ***P*** (often in literature also call hat-matrix)

$$\boldsymbol{P} = \boldsymbol{A} \cdot \left(\mathbf{A}^{\mathrm{T}} \cdot \boldsymbol{V_e}^{-1} \cdot \boldsymbol{A}\right)^{-1} \cdot \boldsymbol{A}^{\mathrm{T}} \cdot \boldsymbol{V_e}^{-1} \qquad (2)$$

using the design matrix *A*

$$\boldsymbol{A} = \left(\frac{\partial \boldsymbol{e}_{Re}}{\partial \boldsymbol{a}}\right) = \left(1 \;\; \lg\frac{Re}{10^6} \;\; \left[\lg\frac{Re}{10^6}\right]^2 \;\; \left[\lg\frac{Re}{10^6}\right]^3 \;\; \left[\lg\frac{Re}{10^6}\right]^4\right) \qquad (3)$$

and the variance-covariance matrix $\mathbf{V}_e$ of the all the measured meter deviations e_i. With this, we get the vector fitted meter deviations $\mathbf{e}_{fit}$ (i.e. the points at our reference curve) by:

$$\boldsymbol{e}_{fit} = \boldsymbol{P} \cdot \boldsymbol{e} \qquad (4)$$

And the vector of residuals ***d*** is given by

$$\boldsymbol{d} = \boldsymbol{e} - \mathbf{e}_{fit} = \boldsymbol{e} - \boldsymbol{P} \cdot \boldsymbol{e} = (\boldsymbol{I} - \boldsymbol{P}) \cdot \boldsymbol{e} \qquad (5)$$

providing the variance-covariance matrix of d_i

$$\boldsymbol{V}_d = (\boldsymbol{I} - \boldsymbol{P}) \cdot \boldsymbol{V}_{\mathrm{e}} \cdot (\boldsymbol{I} - \boldsymbol{P})^{\mathrm{T}} \qquad (6)$$

The diagonal elements of $\boldsymbol{V}_d$ containing our uncertainties of residuals rsp. the differences d_i for each measured meter deviation e_i to the comparison reference curve,

$$u(d_i) = \sqrt{v_{d,i,i}} \qquad (7)$$

hence in our convention the absolute value of normalized difference [13,14] is

$$E_{n,i} = \frac{|d_i|}{2u(d_i)} \qquad (8)$$

Up to here, the approach is straight forward but the central question is the non-diagonal values of the variance-covariance matrix $\boldsymbol{V}_e$ of the measured meter deviations e_i (the diagonal elements are just the squared standard uncertainties of these values)

$$\boldsymbol{V}_e = (v_{i,j}) = \begin{pmatrix} v_{1,1} & \cdots & cov_{i,j} \\ \vdots & \ddots & \vdots \\ cov_{i,j} & \cdots & v_{n,n} \end{pmatrix} \quad (9)$$

In Appendix A of Cox [18], it is proposed to derive the covariant elements based of the uncertainty budget of each participant and the assumption that the total uncertainty of a measured value is a superposition of stochastic (*s*) and constant (*c*) effects

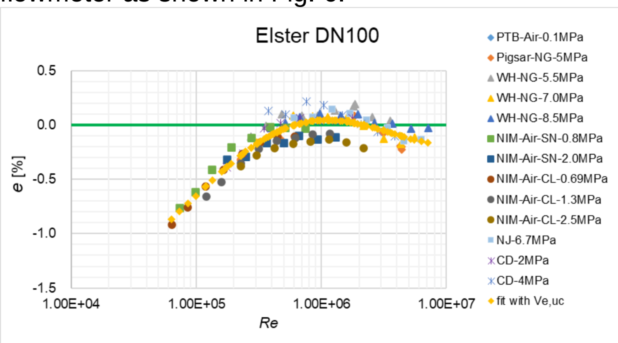
$$u^2(y_i) = s^2 + c^2(y_i) \quad (10)$$

In the comparison here, it is not an appropriate solution to go through the uncertainty budgets because it is not for all effects quite clear if an effect is in the specific situation more random or more constant or have their very specific dependency on time [and with this a specific dependency on *Re* (due to the sequence of measurements) e.g. temperature effects or leakages]. Therefore, we apply the basic idea of the so-called Feasible GLS (FGLS) with a heuristic approach to estimate the covariant element of $\boldsymbol{V}_{\mathrm{e}}$ out of the data itself.

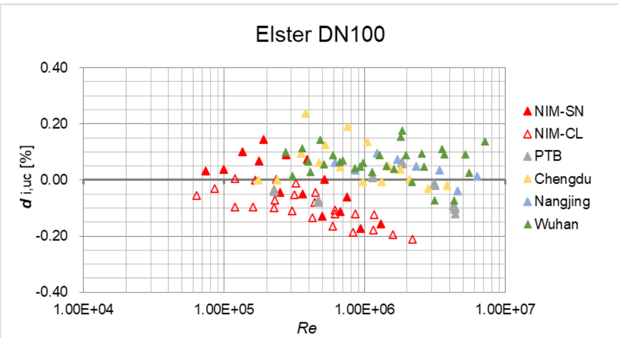
The first step of FGLS is the calculation of fit curve and residuals using the GLS assuming uncorrelated data.

$$\boldsymbol{V}_{e,uc} = (v_{i,j}) = \begin{pmatrix} v_{1,1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & v_{n,n} \end{pmatrix} \quad (11)$$

With this, we get fit (reference curve) and the residuals $d_{\mathrm{i,uc}}$ for all points e.g. for Elster DN100 flowmeter as shown in Fig. 6.



(a) Measured results and curve fitting



(b) Residual

Figure 6: Fitting values and residuals with uncorrelated data variance-covariance matrix for Elster DN100

If we look at the $d_{\mathrm{i,uc}}$ for one participant as shown in Fig 7 e.g. for NIM together with the error bars of measurements, it is visible that the d_{i} are not completely randomly distributed around Zero but following also a trend versus the Reynolds number and the scatter of the trend is smaller than the original reported uncertainty of the measurements.

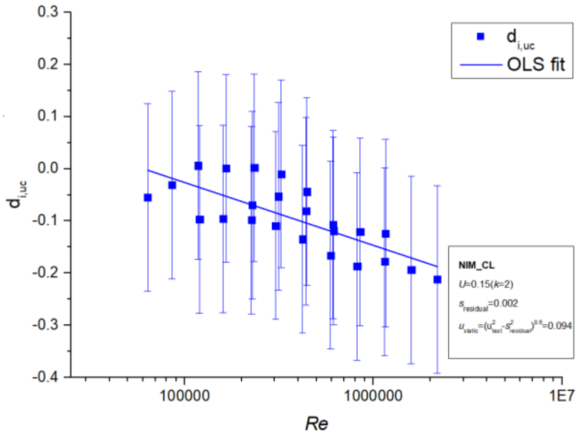


Figure 7: Residuals and fitting line of residuals for NIM Closed Loop facility

The $d_{\mathrm{i,uc}}$ for the data sub-set of on participant can now be approximated again by a function versus Reynolds based on OLS to get an expression of the systematic effect $\hat{d} = f(Re)$. It is of course always a difficult discussion what functionality should be used for the OLS of d_{i}; in case of Fig 6 the application of a line versus log(*Re*) seems to be sufficient. Generally, this question should be answered testing various models in connection with statistical criteria such as the Bayesian Information Criteria (BIC).

The application of OLS on the set of d_{i} for one participant provides a standard deviation *s* and a variance-covariance matrix for the fitted point, i.e. here fitted $\hat{d}$

$$\boldsymbol{V}_{\hat{d}} = \begin{pmatrix} v_{\hat{d}_{i,i}} & v_{\hat{d}_{i,j}} \\ v_{\hat{d}_{j,i}} & v_{\hat{d}_{j,j}} \end{pmatrix} \quad (12)$$

Applying the formal relationship $r_{i,j} = \frac{v_{\hat{d}i,j}}{\sqrt{v_{\hat{d}i,i} v_{\hat{d}j,j}}}$ we get the matrix of correlation coefficients

$$\boldsymbol{r}_{\hat{d}} = \begin{pmatrix} 1 & r_{i,j} \\ r_{j,i} & 1 \end{pmatrix} \quad (13)$$

Coming back to variance-covariance matrix of measured meter deviations $\boldsymbol{V}_{\mathrm{e}}$, we can split formally this matrix into following parts:

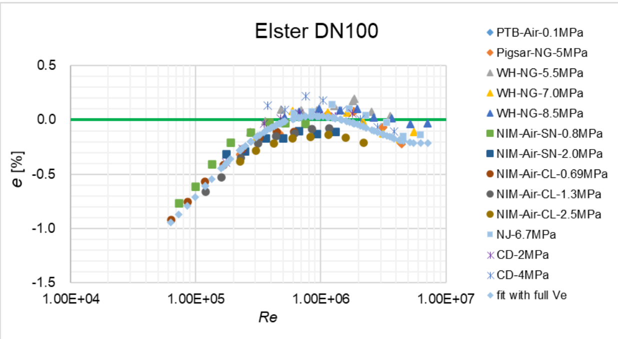
$$\boldsymbol{V}_e = \begin{pmatrix} u_i^2 & cov \\ cov & u_j^2 \end{pmatrix}$$

$$= \begin{pmatrix} s^2 & 0 \\ 0 & s^2 \end{pmatrix} + \begin{pmatrix} c_i & 0 \\ 0 & c_j \end{pmatrix} \begin{pmatrix} 1 & r_{ij} \\ r_{ij} & 1 \end{pmatrix} \begin{pmatrix} c_i & 0 \\ 0 & c_j \end{pmatrix} \quad (14)$$

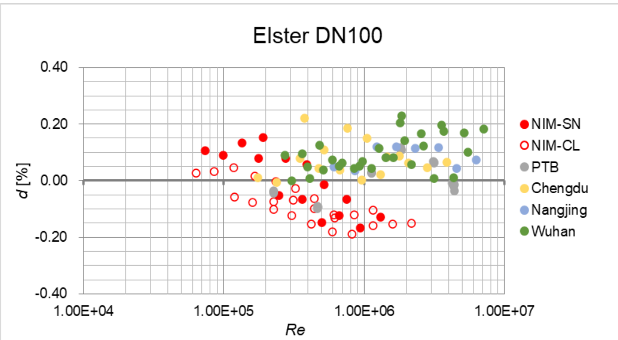
With $u^2(e_i) = s^2 + c_i^2$ rsp. $c_i^2 = u^2(e_i) - s^2$ [1]

The next step is to take the correlation coefficients r_{ij} from the $\boldsymbol{r}_{\hat{a}}$ and the estimate for s out of the OLS of $d_{i,uc}$ above.

With that, the new estimate of the full matrix $\boldsymbol{V}_e$ is achieved[2] and the calculation of the GLS can be applied again. The outcome is shown in Fig 8.



(a) Measured results and curve fitting

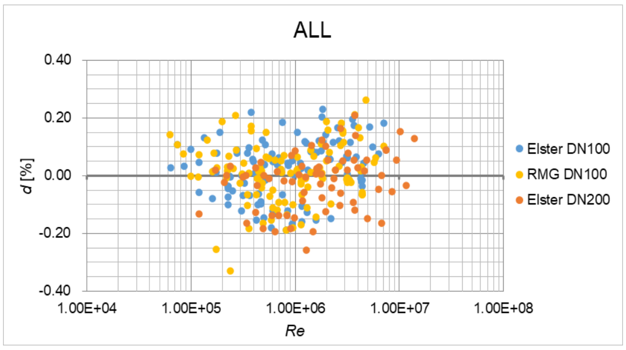


(b) Residual

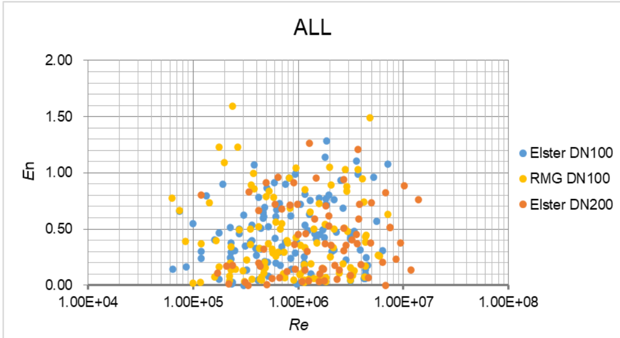
Figure 8: Fitting values and residuals with full variance-covariance matrix for Elster DN100

4.2 Comparison result evaluation

With the above evaluation scheme, the difference between the measured values and the fitted values d, and the normalized deviation E_n were calculated. The absolute values of the difference were all smaller than 0.33%. Among all 298 sets of measured results, there were 282 sets of results with $E_n \leq 1$, while there were 9 sets of results with $1 < E_n \leq 1.2$.



(a) Residual



(b) Normalized deviation

Figure 9: Comparison results for all participants

4.2.1 Elster DN100 result evaluation

There were 112 measured points, the absolute values of the difference were all smaller than 0.23%.

- There were 107 points with $E_n \leq 1$.
- There were 4 points with $1 < E_n \leq 1.2$,
- There were 1 point with $E_n > 1.2$.

4.2.2 RMG DN100 result evaluation

There were 114 measured points, the absolute values of the difference were all smaller than 0.33%.

- There were 105 points with $E_n \leq 1$.
- There were 5 points with $1 < E_n \leq 1.2$,
- There were 4 points with $E_n > 1.2$.

4.2.3 Elster DN200 result evaluation

There were 72 measured points, the absolute values of the difference were all smaller than 0.26%.

- There were 70 points with $E_n \leq 1$.
- There were 2 points with $E_n > 1.2$.

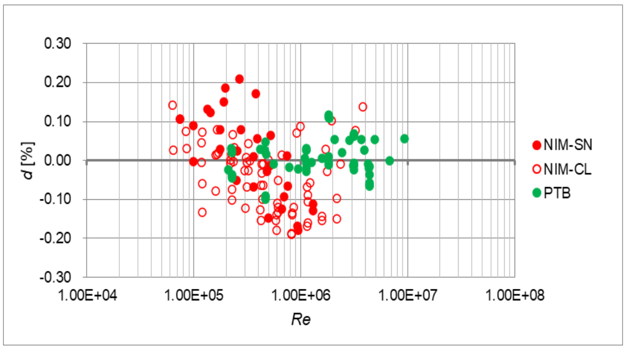
4.3 NIM and PTB result evaluation

Since NIM is the pilot lab and PTB is the link lab in this comparison, the comparison results are further analysed between NIM and PTB to explore the consistency of the secondary standard facility.

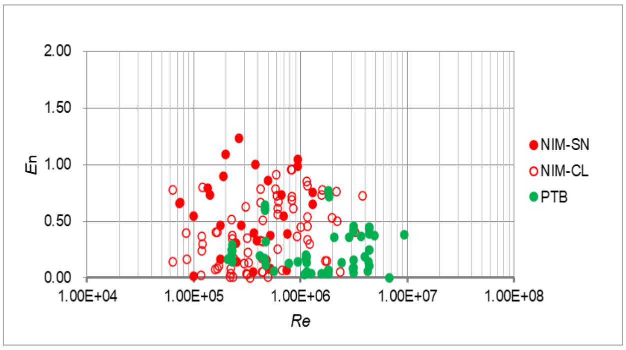
[1] Please note that it must be checked that c_i^2 is always positive. In case of stronger heteroscedastic data sets it is not always fulfilled and this approach has to be extended to deal with heteroscedasticity.

[2] There is of course still no co-variance present between data of two independent labs.

The difference of measured points and E_n values of all the measured points of transfer meters were summarized for NIM and PTB. The residuals of NIM were in the range of (-0.190~0.21) %, and except one measured point with E_n=1.23, all E_n values were smaller than 1.2. The residuals of PTB was (-0.100~0.117) %, and the E_n values wereall smaller than 0.77.



(a) Residual



(b) Normalized deviation

Figure 10: The residual of measured values and normalized deviation for NIM and PTB

5. Conclusion

The first formal comparison of gas flow secondary standard facilities was conducted in China during 2016~2020. There were 4 participants from China, and PTB was invited as the link lab to connect this comparison with the serial key comparisons of CCM.FF.K5. Based on the flow range of the existing secondary standard facilities of participating laboratories, 3 turbine flow meters were used as transfer standards. Totally 298 sets of measured data are obtained with Reynolds number range of (6.3×10^4~1.4×10^7). Among all 298 sets of measured results, there were 282 sets of results with $E_n \le 1$, while there were 9 sets of results with $1<E_n\le1.2$. With consideration of the 95% probability for the uncertainty for each participating laboratory, above 97% of results within the consistent area, so the comparison result is sufficient to support the uncertainty of each participating laboratory.

It is a great achievement as the first formal comparison of gas flow secondary standard facilities at high pressure in China. For further determination the degree of equivalence of high-pressure gas measured in China, optimized comparison procedure should be put forwarded in the next step based on the current results.

References

[1] bp Energy Outlook 2022 edition. https://www.bp.com/content/dam/bp/business-sites/en/global/corporate/pdfs/energy-economics/energy-outlook/bp-energy-outlook-2022.pdf

[2] BIPM. Mutual recognition of national measurement standards and of calibration and measurement certification issued by national metrology institutes. [R]1999.

[3] CCM.FF-K5.a, Metrologia, 2006, 43, Tech. Suppl., 07001

[4] CCM.FF-K5.a.1, Metrologia, 2007, 44, Tech. Suppl., 07006

[5] CCM.FF-K5.b, Metrologia, 2006, 43, Tech. Suppl., 07008

[6] CCM.FF-K5.a.2 , Metrologia, 2013, 50, Tech. Suppl., 07004

[7] B. Mickan, Jean-Pierre Vallet, C.H. Li, J. Wright, Extended data analysis of bilateral comparisons with air and natural gas up to 5 MPa, Flomeko, Sydney, 2016

[8] M Benkova, S Makovnik, B Mickan, R Arias, K Chahine, T Funaki, C H Li, H M Choi, D Seredyuk, C M Su, D Windenberg and J Wright, Comparison of the primary (national) standards of low-pressure gas flow, Metrologia, 2014, 51, Tech. Suppl., 07004, CCM.FF-K6.2011 Final Report, 2014, 36 pages

[9] C. H. Li, Aaron Johnson, The bilateral comparison between NIM's and NIST's gas flow capabilities, Journal of Metrology Society of India, 2011, 26(3):211-224

[10] C. H. Li, Bodo Mickan and Ilya Isaev, The comparison of low pressure gas flowrate among NIM, PTB and VNIIR, Flomeko, Sydney, 2016

[11] Grinten et al., An intercomparison between Primary High-Pressure Gas Standards with Sub-Permille Uncertainty, 37th International North Sea Flow Measurement Workshop, 2019

[12] M G Cox and P M Harris, The evaluation of key comparison data using key comparison reference curves 2012 Metrologia, 49, 437

[13] Cox M.G., Evaluation of key comparison data. Metrologia, 2002, 39, 589-595.

[14] Cox M.G., The evaluation of key comparison data: determining the largest consistent subset, Metrologia, 2007, 44, 187-200.

Appendix 1:

- **The calculation result of fit curve and residuals using the GLS with or without considering data correlation.**

The data correlation effect was considered in the analyses of the comparison as shown in section 4.1.1. The impact of data correlation on of E_n was evaluated, and ΔE_n was calculated as Eq. (1-1)

$$\Delta E_\mathrm{n} = E_\mathrm{n,uncorr} - E_\mathrm{n,corr} \qquad (1\text{-}1)$$

Where $E_\mathrm{n,uncorr}$ is the E_n value using the GLS assuming uncorrelated data; $E_\mathrm{n,corr}$ is the E_n value using the GLS assuming correlated data. The difference of E_n was shown in Figure 1-1

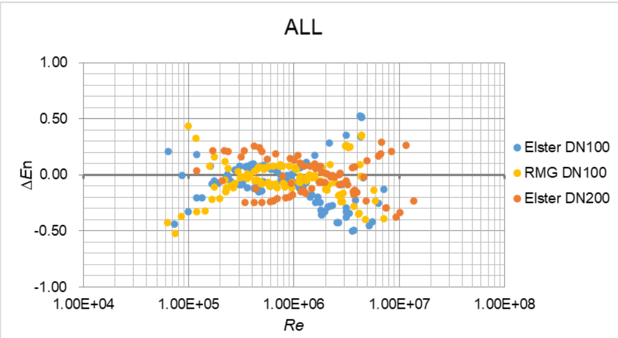
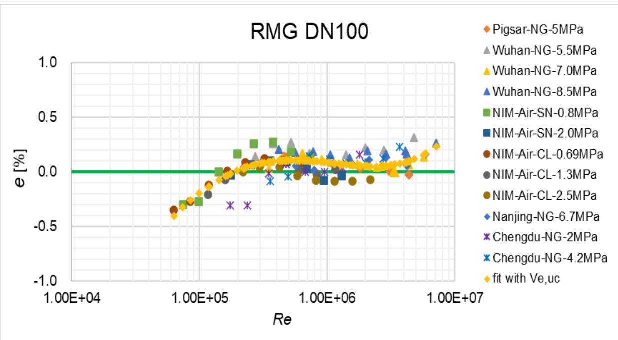
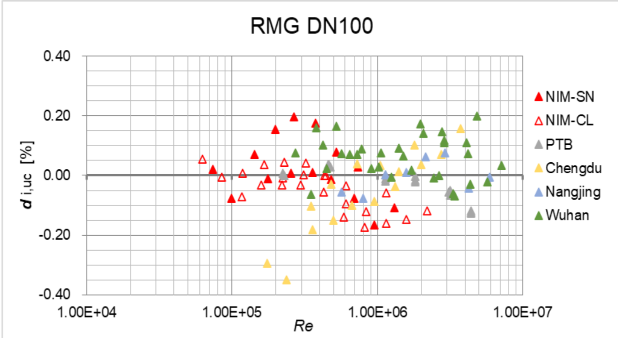


Figure 1-1: Difference of E_n values with or without considering data correlation.

- **The result of fit curve and residuals using the GLS assuming uncorrelated data.**



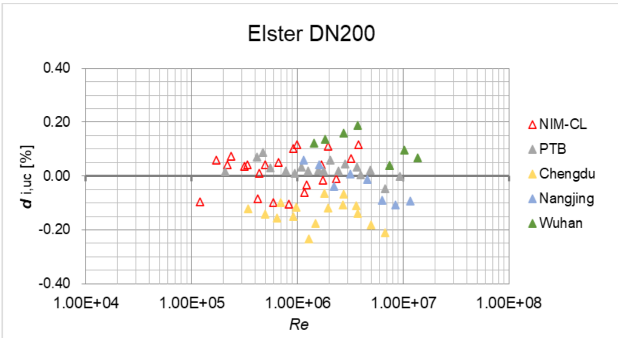
(a) Measured results and curve fitting



(b) Residual

Figure 1-2: Fitting values and residuals with uncorrelated data variance-covariance matrix for RMG DN100

(a) Measured results and curve fitting



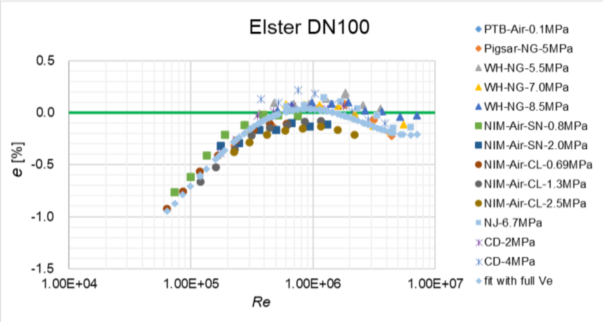
(b) Residual

Figure 1-3: Fitting values and residuals with uncorrelated data variance-covariance matrix for Elster DN200

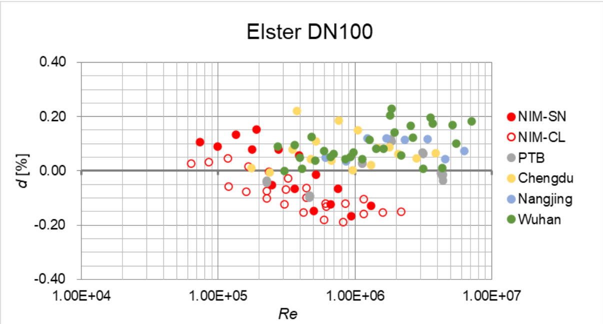
Appendix 2：Comparison results obtained using the GLS with consideration of data correlation for all transfer meters

- **Elster DN100**

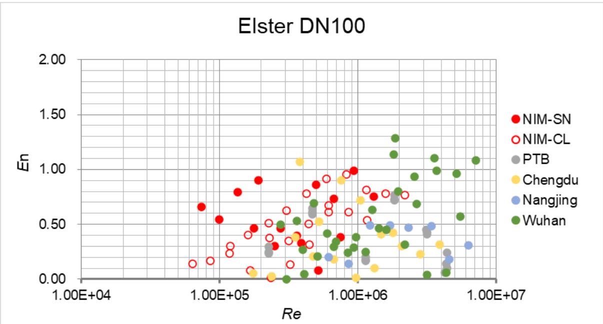
There were 112 measured points, the absolute values of the residuals were all smaller than 0.23%. Among all measured points, 107 points with $E_n≤1$, 4 points with$1<E_n≤1.2$, and1 point with $E_n>1.2$.



(a) Measured results and curve fitting



(b) Residual

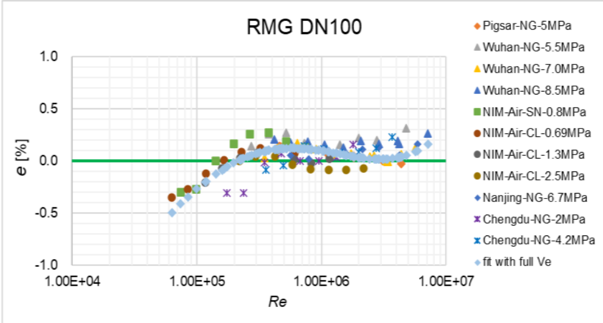


(c) Normalized deviation

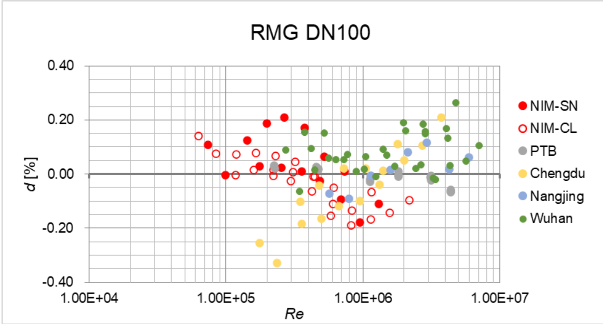
Figure 2-1: Comparison result of Elster DN100

- **RMG DN100**

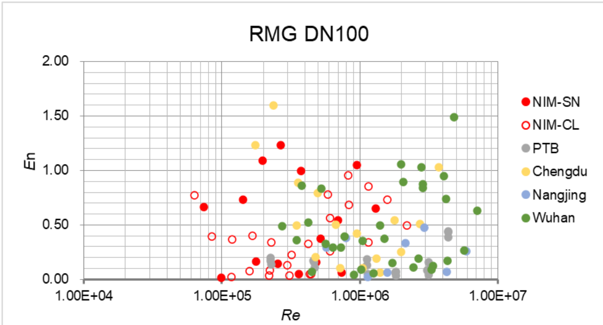
There were 114 measured points, the absolute values of the residuals were all smaller than 0.33%. Among all measured points, 105 points with $E_n≤1$, 5 points with$1<E_n≤1.2$, and 4 points with $E_n>1.2$.



(a) Measured results and curve fitting



(b) Residual

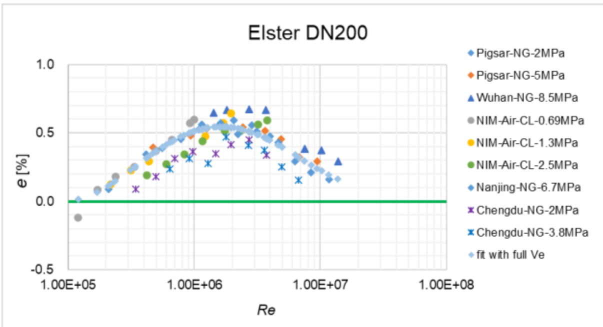


(c) Normalized deviation

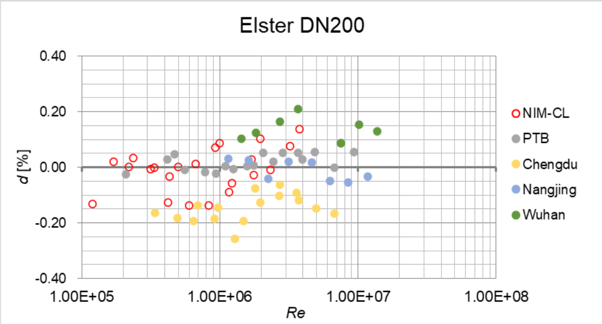
Figure 2-2: Comparison result of RMG DN100

- **Elster DN200**

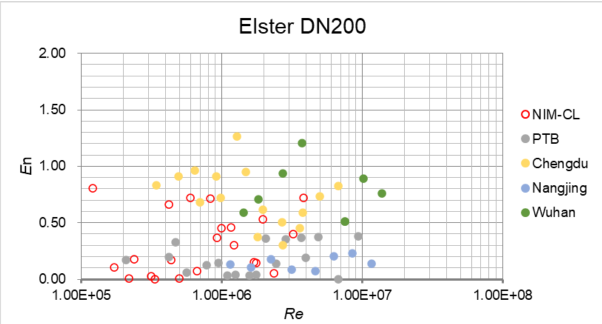
There were 72 measured points, the absolute values of the difference were all smaller than 0.26%. Among all measured points, 70 points with $E_n≤1$, and 2 points with $E_n>1.2$.



(a) Measured results and curve fitting



(b) Residual



(c) Normalized deviation

Figure 2-3: Comparison result of Elster DN200