The Implementation and Result Processing of Highway Engineering Metrology Comparison

Chunxia Zhang*, Biaozhi Zhang and Zhidan Lin
Research Institute of Highway Ministry of Transport, Beijing 100088, China

*Corresponding author e-mail: zcx314@163.com

Abstract. With the need of highway industry measurement standard value transmission and traceability, this paper surveys and analyses the measurement comparison activities carried out inside and outside the industry and study the implementation and evaluation of the measurement comparison. To improve the analysis precision of the method and the credibility of the results of the analysis, we establish the evaluation model, which is verified by an example. The results provide support for the establishment and improvement of highway engineering measurement system.

1. Introduction
As a basic measure to evaluate the accuracy, scientificity and effectiveness of various testing instruments and equipment, metrological verification forms a closed and effective system together with standards, inspection, testing and certification, which supports the development and safety of quality, and is the fundamental guarantee of quality control and evaluation of traffic engineering. Measurement comparison, as an effective tool to ensure the unification of industrial quantity values, is a process of comparison, analysis and evaluation among the quantity values reproduced or maintained by the measurement standards of the same quantity under prescribed conditions, which is an important way to verify laboratory capacity. In 2008, the General Administration of Quality Supervision, Inspection and Quarantine of the People’s Republic of China (formerly) issued the order of No. 107, Measurement Comparison Management Method. [1] The implementation of the method has played a positive role in ensuring uniform, accurate and reliable measurement standards and strengthening supervision and management of measurement comparison.

2. Survey of measurement comparison
Since the 1980s, in order to achieve the consistency of measurement standards in different countries, mutual recognition of calibration certificates issued by national metrology institutes, and mutual recognition agreement (MRA) of International Committee of Weights and Measures (CIPM), metrology departments in various countries have carried out comparisons in different measurement fields. In the metrology disciplines of mechanics, length, thermal engineering, electricity and optics, the key quantity comparison (KC) in bilateral or multilateral manner or organized by CIPM has been carried out. In this way, it can promote international trade, eliminate the trade barriers caused by inconsistent measurement, and reduce the cost of international trade.

In March 2004, the General Administration of Quality Supervision, Inspection and Quarantine of the People’s Republic of China (formerly) issued Specification for Comparison of Measuring
Instrument (JJF 1117-2004). In April 2007, the Ministry of Railways (former) issued Rule for the Operation of Measurement Comparisons of Railway Laboratory (JJF (railway) 603-2007). In August 2008, Measurement Comparison Management Method is issued. [1] Since then, China's metrological comparison work has officially opened the curtain of orderly management. The Measurement Comparison Management Method has two remarkable features. First, it clearly describes the responsibilities and obligations of the leading laboratory, reference laboratory and the organizer of metrological comparison, and emphasizes the confidentiality of the results of metrological comparison. Second, it has very strict, detailed and clear provisions for the procedure, step and participant of measurement comparison, including the determination of the route of measurement comparison and the transportation conditions of standard equipment, so as to ensure the consistency of measurement comparison and the authenticity and fairness of results. In June 2010, the General Administration of Quality Supervision, Inspection and Quarantine of the People’s Republic of China (formerly) issued Measurement Comparison (JJF 1117-2010), implemented from December 10, which not only enriches and enriches the 2004 edition in terms of content, but also puts forward and updates several technical concepts. [2]

The data processing of comparison mainly refers to Statistical Methods for Use in Proficiency Testing by Interlaboratory Comparisons (GB/T 28043-2011), [3] which is equivalent to the international standard ISO 13528: 2005. The standard elaborates the determination of reference value and its uncertainty, the calculation of standard deviation and statistics and the evaluation criteria. In 2015, ISO released ISO 13528: 2015, a revision to the 2005 edition. The new edition basically kept the original format, but made many adjustments in content, such as the consideration of statistical design, the more complicated Q method in the aspect of robust statistical technique, expanding the criterion of statistical test of uniformity and stability, the percentage of standardized tolerance, and the treatment of outliers.

3. Implementation of highway engineering measurement comparison

3.1. Traceability characteristics of measuring instruments for highway engineering

1) System integration device with integrated output value (derived quantity)

There is a large number of non-detachable large equipment with complex control system for special measuring instruments for highway engineering. Even if the equipment can be disassembled and measured separately for each part, the reliability of the integrated system cannot be guaranteed. For the measurement of this kind of equipment, it is unavoidable to test the accuracy of its overall performance in order to ensure the reliability of test results. That is to say, in addition to checking the accuracy of its functional components, the key is to check the overall performance of the components in the state of test work. However, it is often difficult to verify the accuracy of the comprehensive output. First, the test working conditions are difficult to simulate. Especially for vehicular automation equipment, it is generally necessary to use large-scale road sites to test the dynamic performance of the components in the state of test work. Second, it is difficult to find a stable and reliable measurement, such as vehicle-mounted pavement laser flatness meter and rut meter. The special measuring instruments for highway engineering take road, bridge, culvert, safety facilities and other entities as the measurement objects, whose technical indicators are often difficult to be reproduced and difficult to ensure its stability even if achieved.

Regarding the method of quantity value traceability for such instruments and equipment, the following principles should be mastered.

(1) We should inspect the overall performance of the equipment and verify the accuracy of its comprehensive output.

(2) For the comprehensive output value or the main influencing factors, the trace of the quantity value to the source should be on the basis of the national verification system.
(3) It is preferred to use relatively stable real (or simulated) measured objects in the field to simulate the actual working conditions of such equipment. If it is really difficult to realize, indoor simulation by vibration and other means can be considered.

2) Instruments and equipment measured separately by multiple social metrology categories
For some special measuring instruments for highway engineering, their measuring parameters are often multi-quantity values, which involve various professional fields such as length, optics, mechanics and temperature. Only by tracing all the measuring parameters to the source, can the reliability of the measuring values be guaranteed.

Such instruments are often some special of small and medium-sized equipment, and even some tools without value output ability, such as automatic penetration instrument, compaction instrument and rubber sand mixer. The key about the quantity traceability of this kind of equipment is to analyze the influencing factors comprehensively, to highlight the key influencing factors, and to consider the operability of the measurement parameter verification method in combination with the hardware structure.

3) Nondestructive testing when the fixed relationship between the direct measured value and the derived quantity is difficult to determine

With the improvement of understanding about engineering quality and safety, for some concealment and a large number of masonry structures in highway construction projects, the quality evaluation parameters are often derived through complex functional relations from single or multiple measurement values gotten by indirect and non-destructive technical means. Generally, the factors that affect the accuracy and stability of the quantity value of such equipment are complex and changeable, which is difficult to be fully understood. The main principle in the metrological verification of such equipment is to check the accuracy of its output.

4) Equipment whose output value is a unique value or limited physical quantity
The output value of some equipment is the special value of highway engineering. For some equipment, although the output value is a relatively clear physical quantity, it is generated under certain conditions. Therefore, the output value of such equipment cannot be directly traced to the social common basis (standard), which brings certain difficulties to its metrological verification. For example, British pendulum number (BPN) of pendulum friction coefficient tester is the unique value, and the single-wheel deflection friction coefficient tester output friction coefficient is the mechanical value under certain conditions, etc.

The traceability method of this kind of equipment should start from two aspects. The first is the accurate relationship between output value and conventional physical quantity through strict theoretical analysis, and the second is to manufacture standard materials with fixed quantity value, so as to create stable test environment conditions.

5) Quasi-instrumental equipment without output capacity of quantity value
For some test items in highway engineering, in the process of sample preparation, some auxiliary tools are needed, or some agreed attribute tools are used in the process of test. This kind of tool often does not have the output capacity of quantity value, but it has great influence on the test result. One kind of equipment is small and with fixed quantity value of weighing device property, such as ring knife, sand laying apparatus and sand filling cylinder. The other kind is a processing tool used for test sample preparation, such as accelerating polishing machine and plastic sand mixer.

For the metrological verification method of these kinds of equipment, we need to grasp the following points. First, for the small instruments with the property of weighing instrument, we should decide the fixed value of the physical quantities it represents, such as volume, area, length, and so on. Second, the tools should be prepared according to the methods from multiple social metrology categories. Third, if it is not easy to carry out quantity value traceability, although it has an impact on the test and detection, it is not necessary to establish the measurement standard reluctantly.
3.2. Reference value of highway engineering measurement comparison

The two main sources of reference value are authoritative data and the quantity value obtained from multiple reference laboratories. The mode of comparison depends on the conditions of comparison and the reference laboratory.

1) The quantity value of the authoritative data is taken as the reference value, and the uncertainty of the reference value is the uncertainty of the measurement result of the authoritative laboratory. The quantity value from the laboratory with the measurement benchmark or the upper level measurement standard or the uncertainty advantage is selected as the reference value.

   (1) The quantity value of the metrological benchmark or the upper metrological standard is taken as the reference value. When the quantity value of the reference laboratory is transmitted by the same quantity value of a laboratory (directly or indirectly), the quantity value of that laboratory shall be used as the reference value. The quantity value is usually the national measurement standard or the upper level measurement standard.

   (2) The quantity value of the laboratory with an uncertainty advantage is taken as the reference value. If the quantity value of a laboratory has a significant measurement uncertainty advantage and is supported by national comparison results, or its uncertainty evaluation is effectively confirmed, the quantity value shall be used as the reference value.

   (3) The quantity value of a higher level laboratory is taken as the reference value. If the quantity value of a recognized higher level laboratory is available, we should seek a way to use the quantity value as the reference value.

   When using this method, the uncertainty of the authoritative laboratory should be significantly better than that of the participating laboratory, the reference value should not be affected by the results of the participating laboratory, and outliers need not be identified and excluded. However, sometimes using this method may cause higher economic costs. Therefore, it is advisable to use this method when conditions are available.

2) The reference value from the quantity value of multiple reference laboratories. When determining the reference value, the independence of the quantity value of each laboratory shall be fully considered, and the reference value shall be determined by the independent quantity value. The basic principle is to avoid the participation of laboratories whose quantity value is not independent in the calculation of the reference value.

   (1) When there are multiple replicating quantity value laboratories in the reference laboratory, these laboratory quantities are used to determine the reference value.

   (2) The correlation between each laboratory quantity value should be fully considered to avoid the case that the weight of the laboratory quantity value related to a group of quantities is too heavy.

   (3) When it is difficult to clearly evaluate the quantity value reproduction and traceability of the reference laboratory, the reference value can be determined by using the quantity value of the reference laboratory in part or in whole.

3.3. Evaluation of uncertainty of highway engineering

During tests by using special measuring instruments, the detection conditions are complicated (for example, the difference of temperature is big and the laboratory conditions are poor), the control parameters are too many (for example, the force, displacement, time, temperature, and quality parameter combinations need to be controlled), and the effect of the operator's skill level on the results is very large. Therefore, in addition to the permissible error of unit value for measuring instruments, the factors affecting the accuracy, uniformity and reliability of test results include the measured objects, such as uniformity, stability, consistency of test conditions, and the operator's skill level. As a result, the experiment of measuring instrument used for highway project measurement uncertainty evaluation cannot apply the traditional assessment mode of general measuring instruments, which only assesses the uncertainty of measurement introduced by measuring instruments device. We must assess the uncertainty of measurement introduced both by measuring instruments device and by the
differences of technical personnel operating level, and more importantly assess the measurement uncertainty introduced by control inconsistencies. This is the characteristic of highway engineering.

Since a lot of measuring instruments and standard device are vehicular and dynamic measurement, the measured object is difficult to achieve the stability of the public areas of measurement, and many measurement uncertainty components produced in the process are difficult to control and integrate for measuring parameters. So there are technical difficulties in the measurement uncertainty evaluation. Taking vehicle-mounted pavement laser flatness meter verification device as example, the measurement uncertainty includes the components introducing from experimental group piece of repeatability, a group of environmental stability, precision level and total station.

Based on the above situation, the evaluation model of the uncertainty of the highway engineering quantity should first ensure the repeatability test conditions. Then repeated tests were performed to obtain a set of data. Finally, the measurement uncertainty of the whole test method is obtained by calculation.

1) The repeat calculation of the uncertainty components introduced by standard uncertainty and by measurement repeatability because of resolution. The reasonable way is to take the one with the larger value between the two.

2) When evaluating the standard uncertainty caused by measurement repeatability, the arithmetic mean is used as the estimated value, and the experimental standard deviation shall be divided by $\sqrt{n}$, where $n$ is the number of measurement where the arithmetic mean value is taken. If a single measurement is used as the estimated value, the experimental standard deviation need not be divided by $\sqrt{n}$.

3) When evaluating the standard uncertainty introduced by the standard devices, if the nominal value or indication value is chosen in the use of the standard divices, the maximum allowable error of the standard is used in the evaluation, and if the actual value in the use, the standard uncertainty is used in the evaluation.

4. Evaluation and analysis of measurement comparison results

Comparison is the test of the laboratory's measurement standards, environmental conditions, personnel levels, management capabilities, testing methods, data processing, material supply and other levels and capabilities, and examine the consistency of measured values. Appropriate statistical methods should be selected for the analysis of the results of inter-laboratory comparison. We need to consider the distribution characteristics of comparison data, which is related to the number of participating laboratories to some extent. In the case of a large number of participating laboratories, the robust statistical method can be used to give a strict evaluation conclusion. When the number of participating laboratories is small, it is suggested to use the Grubbs test to eliminate outliers first, and consider the maximum allowable error specified by relevant method standards or product standards, and then evaluate and analyze the remaining data.

In general, the consistency of the measurement results of a reference laboratory and its uncertainty is evaluated by normalized deviation $E_n$. Based on the analysis of the traditional evaluation model, this paper makes a new exploration on the evaluation model to study and improve the existing technologies and methods, and analyzes its applicability.

4.1. Calculation model of $E_n$ values

Measurement Comparison (JJF 1117-2010) has strict rules and instructions for the comparison evaluation, whose results use the normalized deviation $E_n$ values or Z score value judgement. For highway industry, $E_n$ value is used commonly. $E_n$ is defined as the ratio of equivalence degree to uncertainty, and the calculation formula is

$$E_n = \frac{y_{\bar{\mu}} - y_{rl}}{\sqrt{u^2(y_{\bar{\mu}}) + u^2(y_{rl})}}$$

(1)
where $Y_{ji}$ represents the measurement result of the $j$th laboratory at the measuring point $i$, $Y_{ri}$ represents the reference value of the measuring point $i$ in the comparison laboratory, $U(Y_{ji})$ represents the extended uncertainty of the reference laboratory at the measuring point $i$, $U(Y_{ri})$ represents the extended uncertainty of the reference value at the $i$th measurement point.

The consistency evaluation principle of the comparison results is that when $|E_n| \leq 1$, the ratio between the difference between the measurement results and the reference value and the uncertainty is within reasonable expectations, and the comparison results are acceptable, and when $|E_n| > 1$, the ratio does not reach the reasonable expectation and the reasons should be analyzed.

Since in the calculation of $E_n$ the denominator is taken as the uncertainty with $k=2$, the confidence probability is about 95%.

### 4.2. Applicability of $E_n$ evaluation

From mathematical meaning, the $E_n$ value indicates whether the measurement uncertainty is within a certain range of uncertainty, and not indicates that measurement is closest to the reference value, because the laboratory with small uncertainty has the similar $E_n$ value with the laboratory with big uncertainty. When the measurement value is the closer to the reference value, the $E_n$ value may not be smaller. Conversely, the smaller $E_n$ value may not mean that the measurement result is closer to the reference value. Therefore, the evaluation method needs to unify the evaluation methods of uncertainty and avoid ignoring or excessively considering the uncertainty component.

In practice of using $E_n$ value, since the evaluation principle is too simple, the judgement maybe cannot truly reflect the real level and ability of some laboratories. We need to notice the following point.

1) The prerequisites for comparison are to ensure that metrology standard equipment has a consistent level for the entire reference laboratory (or the targeted adjustment is implemented for this situation in the comparison report). In many cases, the metrological verification agencies with the same accuracy level measurement standard in different areas have uneven equipment because financial resources. This increases the risk that $E_n$ value is greater than 1 for some comparison result.

2) In the process of some comparisons, the leading laboratory only evaluates the comparison results by value, without considering the measurement capability of the reference laboratory itself, which is not mentioned in the comparison report. Although some reference laboratories claimed that a certain level of measurement capability is reached and the result value was less than 1, the reported measurement uncertainty revealed a lack of measurement capability.

3) In a series of similar measurements, $E_n$ values are expected to be normally distributed. Therefore, when considering the $E_n$ results significantly greater than 1, all results issued by this laboratory should be evaluated to see if there is a systematic deviation. For example, we should check whether the $E_n$ values are always positive or negative.

4) Because the formula contains measurements and measurement uncertainty, from a survey perspective, the $E_n$ value, known as the normalized deviation, has the uncertainty. Thus when we accept the result based on $|E_n| \leq 1$, the instructions for improvement for actual measurement capability is lack for some reference laboratory and the real role of comparison is not fully reflected.

### 4.3. Improvement of $E_n$ value evaluation method

When the uncertainty of $E_n$ value needs to be considered, its reliability should be determined from the inclusion probability of the uncertainty of measurement result, so as to get a clear value of $E_n$ value and to make the determination of measurement ratio reliable and effective. For the combination uncertainty as the denominator in Eq. (1), $U(Y_{ji})$ and $U(Y_{ri})$, independent of each other, should have the same inclusion factor. In the calculation of $E_n$ value in JJF 1117-2010, the selection of $k$ value is generally 2. From JJF1059.1-2011, it can be known that when the probability distribution of output is approximately normal, the inclusion probability of $k=2$ is approximately 95%.

In order to solve the impact of uncertainty comparison on the result determination, Ref. [5] sets the standard uncertainty of the reference value as the lower limit of the uncertainty claimed by the
reference laboratory, that is, \( u_{\text{ref}} \) is taken as the minimum standard uncertainty \( u_{\text{min}} \). However, sometimes \( u_{\text{ref}} \) of the reference laboratory contains an inhomogeneity and an uncertainty component, so the claimed uncertainty may be less than \( u_{\text{ref}} \).

According to the \( \chi^2 \) test, 1.5\( u_{\text{ref}} \) can be taken as the upper limit of the uncertainty of the reference laboratory, the maximum standard uncertainty \( u_{\text{max}} \). The comparison organizer may adjust the upper and lower limits as appropriate according to the objectives and circumstances of the comparison, and provide the reasons for the adjustment. If the uncertainty reported by the reference laboratory exceeds this range, the evaluation procedure for measuring the uncertainty should be examined.

Even if the uncertainty provided by the reference laboratory is between \( u_{\text{min}} \) and \( u_{\text{max}} \), the claimed uncertainty cannot be considered valid if the uncertainty assessment scheme is wrong.

4.4. Calculation of En value model

Due to standardized processes and methods, the processing and discrimination of data comparison can be easily realized by computer programming. Currently popular computer languages, such as Python and R, have very professional and powerful data statistics and analysis library. This article uses the Python language to implement the calculation program of the En value. Python is an interpreted, object-oriented, dynamic data type of high-level programming language. With its unique simplicity, readability and extensibility, Python has become one of the most popular programming languages, widely used in data analysis, software development, big data analysis, machine learning, website construction, games and other fields. Python language ensures the reliability and compatibility of the program, which is easy to maintain and upgrade.

The En value calculation program is mainly compiled based on the theoretical model in this section. The reference value, extended uncertainty of reference value, laboratory measurement value and extended uncertainty of laboratory measurement value are taken as input, and the consistency evaluation result of comparison result is given as output. This program can process large-scale data quickly and stably, and optimize the input and output design of the program.

Considering the convenience and uniformity of data entry, this program supports data in standard CSV format. CSV format can be easily exported from databases and Excel files. The program input file gives the detailed explanation to the data format, in order to help operators to master the program quickly and accurately. The program adds validation to the input data and verifies the common errors in the actual input data, to ensure the robustness of the operation. The data output is as straightforward as possible and easy to read and understand.

The program in the compilation fully considered the extensibility, providing the interface in the input and output, which is very easy to integrate with other programs. For the time being, we assume that the input data meets the requirements for authenticity and uncertainty. In future extended studies, the validity of data can be verified before En value calculation to eliminate unreasonable results. After the calculation of En value, combined with the calculation results and other analysis methods, the causes of the results are analyzed and detailed reports and comparison suggestions are given. These pre- and post-calculations can be easily docked with the program.

5. Case analysis

5.1. Participating laboratory

This comparison was conducted by laboratory No. 1 as the leading laboratory, and a total of 6 units participated in the comparison work:

- Leading laboratory: Laboratory 1
- Participating laboratories: Laboratory 2, Laboratory 3, Laboratory 4, Laboratory 5, Laboratory 6.
5.2. Comparison plan

In order to improve the quality control of the measurement laboratory and ensure the calibration ability of the laboratory, the leading laboratory proposes and drafts the comparison implementation plan, discusses the plan with the reference laboratory, and finally determines the plan of this project.

1) Transfer standards

The selection of transmission standard is mainly considered from two aspects. First, its performance is stable to ensure the safety and reliability of comparison. Second, the selected transmission device should be widely representative. The sample preparation should be conducted by the leading laboratory.

2) Comparison location and schedule

Due to the simple structure, easy handling and excellent stability of the standard devices used in this comparison, this comparison is conducted in a centralized manner. The reference laboratory shall bring the standard devices to the comparison site respectively for comparison experiments.

3) Comparison items

The rising rate of the bearing ratio detector is determined according to the implementation plan.

5.3. Reference value and data processing method

1) The comparison reference value and standard uncertainty

The reference value is the mean value of all participating comparative laboratory measurements, as

\[ Y_{r1} = \frac{1}{n} \sum_{j=1}^{n} Y_{ji} \]  

(2)

where \( j \) is the \( j \)th laboratory that contributes to the reference value, \( i \) is the \( i \)th measurement point of the comparison experiment, \( n \) is the number of laboratories contributing to the reference value, and \( Y_{ji} \) is the measurement results at the \( i \)th measuring point reported by the \( j \)th laboratory.

Since the reference value is the average value of the measured value in the laboratory participating in the comparison, according to the law of uncertainty propagation, the calculation formula of the standard uncertainty of the comparison reference value is

\[ u_{r1} = \frac{1}{n} \sqrt{\sum_{j=1}^{n} u^2_{ji}} \]  

(3)

where \( u_{ji} \) is the standard uncertainty of the measurement result declared by the \( j \)th laboratory at the measuring point \( i \) and \( u_{r1} \) is the standard uncertainty of the reference value at the \( i \)th measurement point.

2) Uncertainty evaluation of comparison parameters

(1) Mathematical model

The rising rate of lifting platform is

\[ v = \frac{h}{t} \]

where \( v \) is the lifting speed, mm/min; \( h \) is the lifting height, mm; and \( t \) is the time, 1min.

(2) Propagation law and sensitivity coefficient

\[ u_e^2 = c_1^2 \times u^2(h) \]

where \( c_1 = 1/t \).

(3) Analysis and calculation of standard uncertainty components

a. Standard uncertainty due to measurement repeatability

Under the condition of repeatability, the bearing ratio detector was measured for 1min lift, and the measurement was repeated for 10 times. The measurement results were

| Number of measurement | 1  | 2  | 3  | 4  | 5  | 6  | 7  | 8  | 9  | 10 | Average | Standard deviation |
|-----------------------|----|----|----|----|----|----|----|----|----|----|---------|-------------------|
| Displayed value (mm)  | 1.038 | 1.032 | 1.040 | 1.057 | 1.033 | 1.040 | 1.040 | 1.048 | 1.038 | 1.050 | 1.0416 | 0.0078 |

According to the regulation, this indicator is only tested for 3 times, and the corresponding uncertainty is
\[
u_1 = 0.0078 / \sqrt{3} = 0.0107 \text{ mm/min}
\]
b. Standard uncertainty introduced by a dial indicator
The maximum indication error of the dial at any 1mm is ±0.01mm, subject to the uniform distribution
\[
u_2 = 0.01 / \sqrt{3} = 0.0058 \text{ mm}
\]
(4) Uncertainty of synthesis standard
\[
u_c = c_1 \sqrt{\nu_1^2 + \nu_2^2} = 0.012 \text{ mm/min}
\]
(5) Extended uncertainty
We take the confidence factor \(k=2\), and
\[
U = 2 \times 0.012 = 0.024 \text{ mm/min}
\]
3) Evaluation of comparison results
The consistency of the results of comparison between the measured values in the laboratory refers to the consistency between the measured values and the reference values, as shown in Eq. 1.

5.4. Comparison results
1) Reference laboratory measurement results

| Reference laboratory | Measurement data | Average |
|----------------------|-----------------|---------|
| Laboratory 1         | 1.010 1.000 1.000 1.010 0.990 1.010 0.990 1.010 1.000 1.010 | 1.003 |
| Laboratory 2         | 1.002 1.000 0.998 0.998 0.996 0.998 1.000 1.000 0.998 0.998 | 0.999 |
| Laboratory 3         | 1.070 1.050 0.975 1.025 1.010 0.980 0.982 1.048 0.975 1.070 | 1.019 |
| Laboratory 4         | 1.010 1.020 1.020 1.010 1.010 1.000 1.000 1.010 1.000 1.010 | 1.009 |
| Laboratory 5         | 1.012 1.021 0.989 0.975 1.004 1.012 1.033 1.011 0.992 0.998 | 1.005 |
| Laboratory 6         | 1.009 1.008 0.998 0.999 0.997 0.999 0.998 0.999 1.008 1.007 | 1.002 |

2) Reference laboratory measurement errors and comparison results

| Reference laboratory | Mean value of measurement (mm/min) | Comparison reference value (mm/min) | Difference between the measured value and the reference value | Extended uncertainty of measurement (k=2) | Extended uncertainty of comparison reference value (k=2) | | En | Judgement result |
|----------------------|----------------------------------|-----------------------------------|-------------------------------------------------------------|------------------------------------------|----------------------------------------------------------|---|-----------------|
| Laboratory 1         | 1.003                            |                                   | 0.003                                                       | 0.03                                     |                                                          |   | 0.04 Satisfaction |
| Laboratory 2         | 0.999                            |                                   | 0.003                                                       | 0.012                                    |                                                          |   | 0.11 Satisfaction |
| Laboratory 3         | 1.019                            |                                   | 0.013                                                       | 0.026                                    |                                                          |   | 0.20 Satisfaction |
| Laboratory 4         | 1.009                            |                                   | 0.003                                                       | 0.036                                    |                                                          |   | 0.04 Satisfaction |
| Laboratory 5         | 1.005                            |                                   | 0.001                                                       | 0.016                                    |                                                          |   | 0.02 Satisfaction |
| Laboratory 6         | 1.002                            | 1.006                             | 0.004                                                       | 0.03                                     |                                                          |   | 0.06 Satisfaction |
6. Conclusion
In order to give full play to the efficiency of measurement comparison and promote the scientific and effective development of measurement comparison activities in China's highway industry, it is necessary to further improve the management system and mechanism of the measurement comparison, establish an authoritative information platform, and strengthen the research and development of measurement comparison technology.

We should further improve the measurement comparison system and establish a unified measurement comparison system of the industry.

As an effective tool to ensure the unification of industrial quantity values, the measurement comparison is an important way to verify laboratory capacity and should be carried out in a normal state. It is necessary to guide the senior management of the laboratory to understand the significance of measurement comparison and to effectively play the role of external evaluation of measurement comparison.

In view of the low utilization rate and acceptance degree of the results of measurement comparison in the laboratory of the highway industry, we suggest establishing a unified measurement comparison platform, so that the project information, process and results of measurement comparison in the industry can be shared, thereby effectively improving the standardization and effectiveness of measurement comparison activities.

We should strengthen the communication activities about the organization, management and technical levels of measurement comparison, and strive to improve the quality of measurement comparison work, in order that each measurement comparison project can play its due role in improving the laboratory capacity and management level.

Acknowledgments
This research was supported by the Special funds for science and technology innovation in Research Institute of Highway Ministry of Transport (Approval No.: 2018-E0042).

References
[1] General Administration of Quality Supervision, Inspection and Quarantine of the People’s Republic of China. Measurement comparison management method [S]. Beijing: Order of the General Administration of Quality Supervision, Inspection and Quarantine No. 107, 2008.

[2] National Technical Committee on Legal Metrology Management. JJF 1117-2010 Measurement comparison [S]. Beijing: General Administration of Quality Supervision, Inspection and
Quarantine of the People’s Republic of China, 2010.

[3] China National Institute of Standardization. GB/T 28043-2011 Statistical methods for use in proficiency testing by interlaboratory comparisons [S]. Beijing: China Standards Press, 2011. (ISO 13528: 2005 IDT).

[4] China National Accreditation Service for Conformity Assessment. CNAS-GL002:2018 Guidance on statistic treatment of proficiency testing results and performance evaluation [M]. Beijing: China National Accreditation Service for Conformity Assessment, 2018.

[5] Ma Jie, Guan Shujun et al. Ability verification and result processing and evaluation [M]. Beijing: China Quality Inspection Press/China Standards Press, 2016:65-66.

[6] Failleau G, Beaumont O, Razouk R. A metrological comparison of Raman-distributed temperature sensors. Measurement, 2017, 116: 18-24.

[7] Gonzalez-Jorge H, Rodriguez-Gonzalvez P. Metrological comparison between Kinect I and Kinect II sensors, Measurement, 2015, 70: 21-26.

[8] Ferrero, Prioli, Salicone. A Metrological Comparison among marketers the Methods for Harmonic Pollution Metering. IEEE the TRANSACTIONS ON INSTRUMENTATION AND MEASUREMENT, 2012, 61: 2972-2981.