Study on a Method for Evaluating Accuracy and Reproducibility Based on Uncertainty Theory

Yang Liu[1,3,a], Chan Huang[1,3,b], Wei Lin[2,c*]

[1]CCCC Second Harbor Engineering Company Ltd., Wuhan, Hubei, 430040, China
[2]CCCC Ltd., Southern Branch, Shenzhen, Guangdong, 510220, China
[3]Key Laboratory of Large-span Bridge Construction Technology, Wuhan, Hubei 430040, China
[a]1946509102@qq.com, [b]357041065@qq.com, [c*]yunqishi77@163.com

Abstract. For engineering experiments such as marine trials where there is no explicit error function or it is difficult to establish a mathematical model for reproducibility, a computational method is proposed to assess their accuracy and reproducibility. First, calculate the combined uncertainty of various physical quantities in the experiment based on the experimental conditions. Utilize theoretical analysis to calculate the uncertainty of the theoretical values. Then, through repeated measurements during the experiment, obtain the mean and standard deviation of the measured values, thereby determining the uncertainty of the measured values. Then, based on probability methods such as the central limit theorem, determine the accuracy and reproducibility of the experiment. Finally, the aforementioned method was applied to a suspended tunnel experiment, achieving the assessment of accuracy and reproducibility for the trial.

Keywords: uncertainty; accuracy; reproducibility.

1 Introduction

Study on accuracy control in engineering experiments is a primary focus of high-accuracy experimental studies. In the field of mechanical engineering, Li et al.[1] conducted an analysis on the propagation of geometric and positional errors of assembly units in the assembly process of CNC machine tools. This analysis facilitates the calculation and storage of various error propagation models in the error propagation network diagram. Furthermore, they constructed a link matrix for the entire product assembly error propagation. In the field of instrumentation and equipment engineering, N. Orkut M et al.[2] developed two instrumental variable (IV) estimators and a multi-factor error structure for dynamic panel data models with exogenous covariates. The main idea is to project common factors from the exogenous covariates of the model and construct instruments based on decomposed covariates. In the field of civil engineering, Zhang et al. [3] improved the comparability of the results of free swelling tests on different soil samples by varying variables such as soil particle size composition and environmental temperature, thereby reducing experimental errors.In the field of aerospace engineering, Hasan M N et al. [4] addressed the vibration of flexible appendages while considering issues such as inertial uncertainty, environmental disturbances, and actuator-related

F. Ding et al. (eds.), *Proceedings of the 2024 International Conference on Civil Engineering Structures and Concrete Materials (CESCM 2024)*, Advances in Engineering Research 247,
https://doi.org/10.2991/978-94-6463-564-5_10

ances, and actuator-related faults. They proposed an adaptive fault-tolerant attitude control scheme based on an improved fast nonsingular terminal sliding mode control (MFNTSM) to achieve high-precision attitude tracking and finite-time convergence of error trajectories. In summary, study on accuracy control in engineering experiments across various engineering fields mainly unfolds within two categories: theoretical [5,6] and practical [7,8].

Reproducibility is an important indicator for assessing the feasibility of engineering experiments [9,10]. Zhang [11] proposed a method for evaluating the repeatability of sensors by fitting data to obtain the temperature and strain sensitivity of the sensors. Through sensitivity analysis, they assessed the repeatability of the sensors, leading to the development of both linear and nonlinear sensor repeatability evaluation methods. Meng[12] established a quadratic programming mathematical model for redundant robotic arm repetitive motion, considering dual problems and convex optimization mathematical models compared to previous ones. They developed a dual neural network and derived dynamic differential equations based on linear variational inequalities to establish the original dual neural network dynamics. Based on these dynamic differential equations, they derived a neural network solver and obtained a computational method for its reproducibility.

Researchers have conducted extensive analysis and study on the accuracy and reproducibility of engineering experiments. However, most of the research has been focused on mechanical and aerospace domains where there are clear functional relationships between error propagation or experimental steps. There has been relatively less study on engineering experiments such as marine trials where there are no explicit error functions or establishing reproducibility mathematical models is challenging. This paper introduces the concept of theoretical value uncertainty and proposes a method to assess the accuracy and reproducibility of such experiments.

2 Uncertainty and Error Propagation Analysis

2.1 Uncertainty Analysis

The meaning of uncertainty refers to the degree of inability to ascertain the measured value due to the presence of measurement errors. Conversely, it also indicates the level of confidence in the result. Generally, the measurement result of a quantity x can be represented as:

$$\text{(The measured value of x)} = x_{best} \pm \delta x \tag{1}$$

In the equation, x_{best} is the best estimate value of x, and δx is the measurement uncertainty of x.

The method of evaluating standard uncertainty by performing statistical analysis on the observed data columns is referred to as Type A uncertainty evaluation. The corresponding standard uncertainty obtained from this method is called Type A uncertainty component, denoted by the symbol uA. The uncertainty of the average measurement result A is:

$$u_A = S / sqrt(n) = \sqrt{\frac{\sum_{i=1}^{n}(s_i - \bar{s})^2}{(n-1)n}} \tag{2}$$

Evaluating standard uncertainty using methods different from statistical analysis on observed data columns is referred to as Type B uncertainty evaluation. The corresponding standard uncertainty obtained from this method is called Type B uncertainty component, denoted by the symbol uB.

When the measurement result is obtained from the values of several other quantities, the standard uncertainty calculated based on the variances and covariances of these other quantities is called the combined standard uncertainty. It is an estimate of the standard deviation of the measurement result and is denoted by the symbol uc. Generally, the uncertainty uc can be expressed as:

$$u_C = \sqrt{\left(u_A^2 + u_B^2\right)} \tag{3}$$

2.2 Error Propagation Analysis

Error propagation commonly involves sum and difference, product and quotient, and power and exponential error propagation. Assuming measurements are taken for x,......,w, with uncertainties $\delta x, ..., \delta w$ respectively, the calculated value for sum and difference can be expressed as:

$$q = x + ... + z - (u + ... + w) \tag{4}$$

In any case, δq will never be greater than the ordinary sum of the original uncertainties:

$$\delta q \le \delta x + ... + \delta z + \delta u + ... + \delta w \tag{5}$$

The calculation for the product and quotient of measured values is as follows:

$$q = \frac{x \times ... \times z}{u \times ... \times w} \tag{6}$$

In any case, δq will never be greater than the ordinary sum of the original uncertainties.

$$\frac{\delta q}{|q|} \approx \frac{\delta x}{|x|} + ... + \frac{\delta z}{|z|} + \frac{\delta u}{|u|} + ... + \frac{\delta w}{|w|} \tag{7}$$

For uncertainty related to powers, if a measurement is taken for x with uncertainty δx, and the power is calculated using the measured value $q = x^n$, then the ratio uncertainty q is $|n|$ times the uncertainty of x.

$$\frac{\delta q}{|q|} \approx |n| \frac{\delta x}{|x|} \tag{8}$$

3 Assessment Methods for Uncertainty of Theoretical Values and Reproducibility

To determine the accuracy and reproducibility of experimental results, this paper proposes the concept of uncertainty of theoretical values. The specific calculation method is as follows:

Step One: The quantities relevant to the experimental results, denoted as D, E, F, G, H, J, K, L, etc., are derived through theoretical analysis. From these, the physical quantities of interest in the experiment, denoted as A, B, and C, are obtained. Suppose A=D+E-mF, B=mGH/J, $C=K^v+cL^y$, where a, b, c, m, v, and y are constants;

Step Two: Confirm whether the physical quantities D, E, F, G, H, J, K, L can be directly measured. If they cannot be directly measured, they are listed separately. Assuming the physical quantity D cannot be directly measured, derive the relationship between D and other quantities relevant to the experimental results. Suppose $D=EF^t$;

Step Three: For the directly measurable physical quantities, identify suitable testing instruments. Determine their Type B uncertainty through manufacturer specifications or calibration certificates. Assuming the instruments required for measurements are O, P, Q, R, S, then their corresponding Type B uncertainties are δo, δp, δq, δr, δs;

Step Four: Match the directly measurable physical quantities E, F, G, H, J, K, L with the corresponding Type B uncertainties of their respective measuring instruments. Let's assume these uncertainties are δo, δp, δq, δr, δs, δs and δs;

Step Five: Perform repeated measurements on the directly measurable physical quantities. After excluding obviously erroneous data, calculate their mean and standard deviation, denoted as (e, δe), (f, δf), (g, δg), (h, δh), (j, δj), (k, δk), (l, δl) respectively;

Step Six: Calculate the combined uncertainty, $u_C = \sqrt{\left(u_A^2 + u_B^2\right)}$, which is $u_e = \sqrt{\left(\delta_e^2 + \delta_o^2\right)}$. Similarly, the mean and combined uncertainty of the directly measurable physical quantities are represented as (e,u_e), (f,u_f), (g,u_g), (h,u_h), (j,u_j), (k,u_k), (l,u_l);

Step Seven: For the physical quantity D that cannot be directly measured, according to its relationship with other physical quantities D=EFt, we can deduce that $\frac{u_d}{d} = \frac{u_e}{e} + t\frac{u_f}{f}$. Therefore, $u_d = d(\frac{u_e}{e} + t\frac{u_f}{f})$;

Step Eight: Similarly $u_a = u_d + u_e + mu_f$, $u_b = mb(\frac{u_g}{g} + \frac{u_h}{h} + \frac{u_j}{j})$, $u_c = c(v\frac{u_k}{k} + y\frac{u_l}{l})$, the physical quantities of interest in the experiment can be represented as (a, ua), (b, u_b), (c, u_c). At this point, we obtain the theoretical value range of the physical quantities of interest under the experimental conditions.

Step Nine: For the physical quantities of interest in the experiment, identify the corresponding testing instruments W, X, Z. Repeat steps three to six to obtain the measured value range obtained from the experiment, which can be represented as (ac, u_{ac}), (bc, u_{bc}), (cc, u_{cc});

Step Ten: According to the Central Limit Theorem, it is known that the results obtained from repeated testing follow a normal distribution. Thus, the distribution of the theoretical and measured values of the physical quantities of interest in the experiment can be represented as N~(a, u_a), N~(b, u_b), N~(c, u_c) and N~(ac, u_{ac}), N~(bc, u_{bc}), N~(cc, u_{cc});

Step Eleven: Taking the physical quantity A as an example, its accuracy can be assessed as $pe=|a-ac|/a$. The calculation of reproducibility can be seen as finding the probability of the intersection of N~(ac, u_{ac}) within the intervals ($a-2u_a$, $a+2u_a$) and ($ac-2u_{ac}$, $ac+2u_{ac}$). This probability is denoted as qe. The assessment of accuracy and reproducibility for physical quantities B and C follows the same procedure.

4 Experimental Application

4.1 Introduction to Suspended Tunnel Testing

The layout of a suspended tunnel test site is shown in Fig. 1. The standard dimensions of the materials are as follows: steel rod length of 12 meters, diameter of 5 centimeters, and mass of 187.3 kilograms; foam outer diameter of 156 millimeters. The boundary conditions are either fixed or simply supported at both ends.

Fig. 1. Experimental Site Layout

4.2 Evaluation of Accuracy and Reproducibility of Frequency Measurements

Step One: When the structure is in an elastic working state, the relevant physical quantities in the experiment include foam diameter D, foam density ρ, steel rod diameter d, steel rod length L, steel rod mass m, elastic modulus E, moment of inertia I, etc. Determine the required physical quantity for the experimental result as frequency ω, with the fundamental frequency theoretical value being $\omega = (\frac{\pi}{l})^2 \sqrt{\frac{EI}{m_0}}$ (taking simply supported as an example, fixed support is similar);

Step Two: Confirm that the physical quantities foam diameter D, steel rod diameter d, steel rod length L, and steel rod mass m can be directly measured (foam density is provided by the manufacturer). Elastic modulus E and moment of inertia I cannot be directly measured. For $I = \frac{\pi d^4}{64}$ and $E = \frac{FL^3}{3\delta I}$,the measurement is based on the formula for deflection of a cantilever beam under the action of a concentrated force at the beam end, derived from the principles of materials mechanics.

Step Three: For the directly measurable physical quantities, find appropriate testing instruments. Determine their Type B uncertainty through manufacturer specifications or calibration certificates. The table for this experiment is as table 1:

Table 1. Type B Uncertainty Test Results

Instruments	accuracy	Type B Uncertainty
Vision Measurement	0.02mm	0.01
Accelerometer	0.0001grms	0.00005
Tape Measure	1mm	0.5
Vernier Caliper	0.02mm	0.01
Electronic Scale	0.01KG	0.005

Steps Four to Six: After repeated measurements, the mean and uncertainty of directly measurable physical quantities are shown in the table 2:

Table 2. Uncertainty Test Results

Material	Physical quantity	Mean of measurement	Type A Uncertainty	Type B Uncertainty	combined uncertainty
steel rod	Length: 12m	12.000	0.006	0.005	0.008
	Diameter: 50mm	50.000	0.057	0.001	0.057
	Mass: (KG)	187.300	0.548	0.1	0.557
	moment ofInertia: (m^4)	306796.200	68.673	0.001	68.673
	Elastic Modulus: (GPa)	170.000	9.076		9.076
foam	Diameter: 156mm	156.000	0.030	0.005	0.031

Step Seven: For the physical quantities elastic modulus E and moment of inertia I, which cannot be directly measured, their relationships with other physical quantities

$\delta I = I\delta\left(\frac{\pi D^4}{64}\right) = \frac{\pi}{16} I \left|\frac{\delta D}{D}\right|$ and $\delta E = \frac{E}{3}\left(\left(\frac{\delta F}{F}\right) + 3\left(\frac{\delta L}{L}\right) + \left(\frac{\delta\delta}{\delta}\right) + \left(\frac{\pi}{16}\frac{\delta D}{D}\right)\right)$ are used to determine the corresponding uncertainties. The results are shown in Table 2.

Step Eight: Similarly $\delta\omega_0 = \delta\left((\frac{\pi}{l})^2 \sqrt{\frac{EI}{m_0}}\right) = \omega_0 \pi \left[2\frac{\delta L}{L} + \frac{1}{2}(\frac{\delta E}{E} + \frac{\delta m_0}{m_0} + \frac{\delta I}{I})\right]$, the corresponding uncertainties are determined accordingly. The results are shown in Table 3.

Steps Nine to Eleven: For the physical quantities of interest in the experiment, find the corresponding testing instrument, the accelerometer. Repeat steps three to six to obtain the measured value range from the experiment. Based on the Central Limit Theorem and the uncertainty intervals of theoretical and measured values, assess the accuracy and reproducibility. The results are shown in Table 3.

From the table, it can be seen that the errors and reproducibility for simply supported and fixed support conditions are 5.1%,4.9% and 75.1%, 75.3%, respectively. The errors and reproducibility for simply supported and fixed support conditions are both close, indicating that the boundary conditions have a small influence on errors and reproducibility.

Table 3. Frequency Test Results

Boundary Condition	Mean of meas-urem-ent	Type A Un-certa-inty	Type B Uncer-tai-nty	com-bined uncer-taint-y	Re-pro-duci-bility	Theo-retical Values	Uncer-tainty of Theoret-ical Values	Error
Simply Supported	6.4358	0.895	0.00005	0.895	75.1%	6.1104	0.550	5.1%
Fixed Support	9.3223	1.301	0.00005	1.301	75.3%	8.8678	0.798	4.9%

5 Conclusion

To address the issue of the lack of clear error propagation functions between experimental steps and between experimental steps and results, which impedes the assessment of accuracy and reproducibility, this paper proposes a method for evaluating errors and reproducibility based on uncertainty. The main conclusions obtained are as follows:

(1) Based on the general principles of uncertainty and error propagation, a method for calculating the uncertainty of experimental theoretical values is provided.

(2) Through the probability distribution patterns between theoretical and measured values and relying on the Central Limit Theorem, the errors and reproducibility are assessed. This is further validated through experimentation.

(3) The results of suspended tunnel experiments under both simply supported and fixed support boundary conditions indicate low experimental errors and high reproducibility, suggesting high reliability of the experimental results. The proposed algorithm demonstrates feasibility and applicability.

Reference

1. Li Dongying, Li Mengqi, Zhang Genbao, et al. Study on Error Sources and Error Transmission Models of Basic Action Assembly Units[J]. Journal of Mechanical Engineering, 2015, 51(17):10.DOI:10.3901/JME.2015.17.146.
2. Norkute T C G .Instrumental variable estimation of dynamic linear panel data models with defactored regressors and a multifactor error structure[J].Journal of Econometrics, 2021, 220(2).https://doi.org/10.1016/j.jeconom.2020.04.008.
3. Zhang Weili, Li Yonghong, Dang Jinqian, Wang Fei. Research on Improving the Accuracy of Free Swelling Rate Test of Expansive Soil[J]. Journal of Northwest A&F University: Natural Science Edition, 2015, 43(9):7.DOI:10.13207/j.cnki.jnwafu.2015.09.029.
4. Hasan M N , Haris M , Qin S .Vibration suppression and fault-tolerant attitude control for flexible spacecraft with actuator faults and malalignments[J].Aerospace Science and Technology, 2022, 120:107290-.DOI:10.1016/j.ast.2021.107290.
5. Wang Gaoli, Tang Jiyun. Error Analysis and Optimization of Constraint Point Loads in Full-Scale Aircraft Structural Tests[J]. Engineering and Experiment, 2014, 54(2):5. DOI:CNKI:CDMD:1.1012.021096. DOI:CNKI:CDMD:2.1016.914558.
6. Tong Lin. Study on Stamping Forming and Precision Control of Arc-shaped Plate[D]. Harbin Institute of Technology. DOI:CNKI:CDMD:2.1016.914558.
7. Chang Shuai. Study on Data Processing Methods for Engine Static Noise Experiment[D]. Civil Aviation University of China. DOI:CNKI:CDMD:2.1016.917572.
8. Kmmer J E, Schauber S K, Hautz S C, et al. Differential diagnosis checklists reduce diagnostic error differentially: a randomized experiment[J]. Medical Education, 2021. DOl: 10.1111/medu.14596.
9. Hummel T, Coatantiec C, Gnata X, et al.A fibre-based 2D-slit homogenizer concept for high-precision space-based spectrometer missions[J].CEAS Space Journal, 2022:1-14.DOI:10.1007/s12567-021-00419-8.
10. Tatarczak A. LogForum A Decision Making Support System In Logistics Cooperation USING a Modified Vikor Method Under an Intuituinistic Fuzzy Environment[J]. Logforum, 2020, 16(2):251-260.DOI:10.17270/J.LOG.2020.436.
11. Zhang Weiwei. Study on Temperature and Strain Sensing Characteristics and Reproducibility of Fiber Bragg Grating in Low Temperature Environment[D]. Wuhan University of Technology, 2022. DOI:10.27381/d.cnki.gwlgu.2022.000632.
12. Meng Xiangyuan. Study on Reproducibility Problem of Redundant Manipulator Based on Neural Network Dynamics[D]. Harbin Engineering University, 2019. DOI: 10. 27060/ d.cnki.ghbcu.2019.000866.