CrossMark

Weighting by Cross-Validation: A Calibration Method for Force Measurements via Transient Response Analysis

C. Luo[1] · Y. Wang[1] · Z. Hu[1,2] · J. Li[1] · Z. Jiang[1,2]

Abstract

Transient response analysis is an indirect method for the measurement of the aerodynamic force experienced by hypersonic vehicles in impulse wind-tunnel facilities; a transfer function is identified by a calibration experiment in advance and is then used to recover the target aerodynamic force. Theoretically, the transfer function is unique for a given measurement system. However, the calibration experiment may involve unpredictable factors and noise that are inevitable in practical applications, which will result in systematic errors. In this paper, a new calibration method, weighting by cross-validation (WCV), is proposed to reduce the systematic errors. In WCV, a series of on-site calibration experiments are carried out to obtain a set of transfer functions prior to the wind tunnel test. The transfer functions are cross-validated against each other to create a cross-validation table of relative measurement errors. The aerodynamic force is then calibrated using the average and standard deviation of the cross-validation errors. The working mechanism of the WCV method is demonstrated by analogy of measurements with a set of non-standard rulers. The effectiveness of the WCV method has been verified by tests in the JF-12 shock tunnel. Studies show that the WCV method improves the measurement accuracy significantly. In addition, the concept of the WCV method is general and can also be applied to other indirect measurement problems to reduce systematic errors, especially when no exact/standard measurement tools are available.

Introduction

Impulse facilities such as shock tunnels are playing important roles in developing new hypersonic vehicles. However,

✉ C. Luo
luo@imech.ac.cn

Y. Wang
wangyp@imech.ac.cn

Z. Hu
huzm@imech.ac.cn

J. Li
lijinping@imech.ac.cn

Z. Jiang
zljiang@imech.ac.cn

[1] State Key Laboratory of High Temperature Gas Dynamics, Institute of Mechanics, Chinese Academy of Sciences, Beijing 100190, China

[2] School of Engineering Sciences, University of Chinese Academy of Sciences, Beijing 100049, China

the measurement of the aerodynamic forces of test models in impulse facilities remains a challenge [1–3] because the test time is very short, usually in the order of milliseconds. Sudden aerodynamic loading causes vibrations of the force measurement system, and the test time runs out before the system reaches a state of force equilibrium. Early research has relied on damping mechanisms and/or inertial compensation to reduce the influence of vibrations. For example, Jessen and Groenig suggested strengthening the stiffness of the force measurement system [4]. Joshi and Reddy [5] and Raju and Reddy [6] used accelerometers to compensate for the vibrations, and Naumann et al. proposed a free flight technique to reduce the vibration effect [7, 8]. More recently, transient response analysis has been widely used to recover the aerodynamic forces from impacted balance signals [9–14], in which the force measurement is considered a dynamic process. A transfer function is identified by a calibration experiment prior to the wind tunnel test and is then used to recover the target aerodynamic force. The calibration experiment may be a hammer knocking or a wire cutting.

Theoretically speaking, the transfer function is unique for a given measurement system. However, the calibration

SEM

experiment might involve *unpredictable* factors and noise in practical applications (see Section 2). Here, an unpredictable factor is defined as an influence factor that is known but the specifics are unpredictable. For example, only some (i.e., NOT all) low-frequency vibrations may be excited in a single knocking test [16]. Unpredictable factors result in systematic errors.

In other words, different calibration experiments may result in different transfer functions. This means it is difficult to obtain an ideal transfer function for force recovery. Note that the role of the **transfer function** in a transient response analysis is similar to a **ruler** in length measurements. A force measurement with an exact transfer function is just similar to a length measurement with a non-standard ruler. The error due to the measurement tool is a systematic error that affects the final measurement result.

In this study, we suggest conducting a series of on-site calibration experiments to obtain a set of transfer functions (as a set of rulers) prior to every wind tunnel shot. The transfer functions are cross-validated against each other to determine the weight of each transfer function. A cross-validation table of the relative measurement errors is developed. The aerodynamic force is then calibrated using the average and standard deviation of the cross-validation errors. This calibration method is referred to as weighting by cross-validation (WCV).

The working mechanism of the WCV method is demonstrated by using the analogy of obtaining measurements with a set of non-standard rulers. The effectiveness of WCV is verified by tests in the JF-12 shock tunnel. The results of the study show that WCV improves the measurement accuracy significantly. The method is also a promising calibration method for other indirect measurements when no exact/standard measurement tools are available.

Dynamics of Force Measurement System

Usually, the force measurement system used in impulse facilities consists of a test model, a stress/gauge balance, and a crescent sting (see Fig. 8 in reference [15]). All parts of the model-balance-sting system are made of solid metal and are connected tightly to form a single object. This system is usually considered to be a linear dynamic system. The measurement signals are analyzed as follows. Since the measurement system is linear, the balance outputs y(t) and the aerodynamic inputs f(t) have the following relations:

$$a_n \frac{d^n y(t)}{dt^n} + a_{n-1} \frac{d^{n-1} y(t)}{dt^{n-1}} + \cdots + a_0 y(t) = b_m \frac{d^m f(t)}{dt^m}$$
$$+ b_{m-1} \frac{d^{m-1} f(t)}{dt^{m-1}} + \cdots b_0 f(t).$$

Thus the transfer function in the frequency domain H(s) satisfies

$$H(s) = \frac{Y(s)}{F(s)} = \frac{b_m s^m + \cdots + b_1 s + b_0}{a_n s^n + \cdots + a_1 s + a_0}, \qquad (1)$$

where $Y(s) = \mathcal{L}(y(t))$, $F(s) = \mathcal{L}(f(t))$, and $n \geq m$. By using the convolution formula of the Laplace transform, we obtain the following formula under the assumption of zero initial conditions in the time domain:

$$y(t) = \int_0^t h(t - \tau) \cdot f(\tau) d\tau.$$

Once the test model is installed with a given force balance and crescent sting in the wind tunnel, the transfer function should be invariant. It can be identified using calibration experiments such as a knocking test with an impact hammer (to simulate an impulse input), or a wire-cutting experiment with a mass (to simulate a step input) (see [11] for an example). Subsequently, the aerodynamic input is obtained by a deconvolution with the balance signals of the wind tunnel test.

The Calibration Method

The aim of the calibration is to determine the transfer function $H(s)$ of the force measurement system by observing the inputs $f_c(t)$ and outputs $y_c(t)$ of the calibration experiments using the formula

$$H(s) = \frac{\mathcal{L}(y_c(t))}{\mathcal{L}(f_c(t))}. \qquad (2)$$

Theoretically speaking, the transfer function is unique for a given measurement system. Only one calibration experiment is sufficient to determine the transfer function. This has been a common approach in the past, and has been described in the literature. However, practical applications indicate that different calibration experiments may result in different transfer functions, although the differences might not be very large. In fact, the calibration experiment might involve unpredictable factors and noise. For example, finite element analyses (FEA) and knocking experiments have shown that the vibrations of a model-balance-sting system involve a series of natural frequencies but only some of them were excited in a single knock (refer to [16] for details).

Note that the role of a transfer function in a transient response analysis is similar to using a ruler for length measurements. That is, a transfer function may be considered a *ruler* of the force measurement but the measurement tool is not exact and all *rulers* contain errors. Clearly, the error will be passed on to the final measurement results in the force recovery process. This type of error is systematic.

The problem is that several rulers exists, and any of them may contain errors. A natural and easy method is to recover the aerodynamic force from the transfer functions one by one and use the average as the final result. However, averaging only removes random errors and not the systematic errors. This is an interesting finding and is further discussed in the following section (particularly, see Table 2 in Section 2). Note that this situation is different from the situation in which several measurements are made using an ideal ruler with random errors. In that case, the error can be almost eliminated if we use the average of the measurement results.

Therefore, in this study, we suggest conducting a systematic on-site calibration for every wind-tunnel run and use a new method, referred to as WCV, to reduce the systematic errors and improve the measurement accuracy.

Note that calibration experiments are much cheaper than wind-tunnel tests; therefore, the WCV method will not result in much additional cost.

The cross-validation here means that all calibration results (i.e., transfer functions) need to be validated against each other. For the i-th calibration experiment, the inputs $f_c{}^{(i)}(t)$ and outputs $y_c{}^{(i)}(t)$ determine the transfer function $H^{(i)}(s)$ (see Formula (2)). For another experiment (e.g., j-th), the inputs $f_c{}^{(j)}(t)$ can be derived from the outputs $y_c{}^{(j)}(t)$ and the transfer function determined by the i-th experiment $H^{(i)}(s)$. We denote the average of the derived force load as $f_j^{(i)}$. Although the input of the j-th experiment is now derived, in the case of wire-cutting experiment, the input is considered known because it is easy to obtain from the weight of the suspended mass, denoted by f_j. Therefore, the difference between the true load f_j and the derived load $f_j^{(i)}$ can be considered the prediction error. Usually, we use the relative value $\varepsilon_{ij} = \frac{f_j^{(i)} - f_j}{f_j}$ for further calibration. Thus we can create a calibration table (see Table 1) that contains the relative errors of each cross-validation.

The average error identified by the i-th transfer function μ is defined as $\varepsilon_{i\cdot} = \frac{1}{n}\sum_{j=1}^{n}\varepsilon_{ij}$ and the standard deviation of the i-th transfer function is defined as $\sigma_i = \sqrt{\frac{1}{n-1}\sum_{j=1}^{n}(\varepsilon_{ij} - \varepsilon_{i\cdot})^2}$.

The total average error ε_{tot} is further defined as $\varepsilon_{\text{tot}} = \frac{1}{n}|\sum_{i}^{n}\varepsilon_{i\cdot}|$. Note that the total average error ε_{tot} is a key parameter and it should not be large if this set of transfer functions is self-consistent (i.e., the results do not conflict with each other).

After the calibration experiments, the wind tunnel test is carried out. The aerodynamic force loads, denoted by $\hat{f}_i$, can be recovered from the output signals of the force balance and the previously identified transfer functions $H^{(i)}(s)$, $i = 1, 2, \cdots, n$.

In a simple way, one can use the average value of $\hat{f}_i$ ($i = 1, 2, \cdots, n$) as the final measured force load, i.e., $\hat{f} = \frac{1}{n}\sum_{i=1}^{n}\hat{f}_i$. However, as mentioned above and demonstrated in the next section, this method does not reduce the systematic errors. Therefore, a calibration equation based on the cross-validation table is defined as follows.

$$\hat{f} = \sum_{i=1}^{n}\hat{f}_i * W_{gi} * W_{li} \tag{3}$$

where $W_{gi} = \frac{1}{1+\varepsilon_{i\cdot}}$ is called the weight of global calibration and $W_{li} = \frac{\ln(\sigma_i)}{\sum_{j=1}^{n}\ln(\sigma_j)}$ is called the weight of local fine-tuning.

In summary, the steps of WCV are as follows.

Step 1. Prior to every wind-tunnel run, perform a series of wire-cutting calibration experiments to obtain a set of transfer functions $H^{(i)}(s)$ as rulers.

Step 2. Create a cross-validation table (see Table 1) using the transfer functions obtained in step 1 and check its self-consistency according to the total average of the relative errors ε_{tot}. Continue if the error is consistent; otherwise, perform additional calibration experiments and return to step 2.

Step 3. Determine the calibration weights, W_{gi} and W_{li}, for each wire-cutting experiment.

Step 4. Run the wind tunnel test and obtain the signals of the aerodynamic loading response $y(t)$.

Step 5. Calculate the recovered force load $\hat{f}_i$ for every transfer function.

Step 6. Ensemble the force loads using equation (3) with the calibration weights obtained in step 3.

The working mechanism behind WCV is presented in the following section.

The Concept of WCV

Weight of Global Calibration

Suppose we have a set of non-standard (i.e., not ideal) rulers, and need to measure the distance between two points with these rulers. Without loss of generality, suppose we have

Table 1 Cross-validation table for calibration

	①	$\cdots$	①	$\cdots$	⓷	μ	σ
①⇒	ε_{11}	$\cdots$	ε_{1j}	$\cdots$	ε_{1n}	$\varepsilon_{1\cdot}$	σ_1
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
①⇒	ε_{i1}		ε_{ij}		ε_{in}	$\varepsilon_{i\cdot}$	σ_i
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
⓷⇒	ε_{n1}	$\cdots$	ε_{nj}	$\cdots$	ε_{nn}	$\varepsilon_{n\cdot}$	σ_n

five rulers; only the third can provide the exact result (but it is not known before the measurement) and the first and second rulers always provide lower measurement values, e.g., $x_{\text{tick1}} = 0.8 * x_{\text{actual}}$ and $x_{\text{tick2}} = 0.9 * x_{\text{actual}}$ respectively. The fourth and fifth rulers always provide higher measurement values, e.g., $x_{\text{tick4}} = 1.1 * x_{\text{actual}}$, and $x_{\text{tick5}} = 1.2 * x_{\text{actual}}$ respectively. Let the rulers measure each other for a nominal length of 1 m and we obtain 25 measured values x_{ij}. These values are displayed in a cross-measurement table (see Table 2). In addition, suppose the actual distance between the two given points is 0.95 m. The distances measured with the five rulers are quite different; the results are listed in the last column of Table 2.

Table 2 shows that the measured results may vary significantly with non-standard rulers. The last column indicates that averaging the measured values provides a reasonable result (i.e., 0.9697 in this case). However, the average distance still has a relative error (2.07%) given that the actual value is 0.95. Note that the bias of this set of rulers is symmetric about zero, but the average measurement value still has a considerable bias.

One can verify that this bias does not change much even if more rulers are used. For example, the average distance is 0.964 when measured with a set of 41 non-standard rulers: $x_{\text{tick}} = k * x_{\text{actual}}$, $k = 0.80, 0.81, 0.82, \cdots, 1.19, 1.20$. It is apparent that the error is larger if the bias of the rulers is asymmetric about zero. All these facts indicate that this bias is a systematic error, and cannot be removed by repeating the measurement process. Recall that increasing the number of calibration experiments results in a larger number of transfer functions $H^{(i)}(s)$, similar to the larger number of rulers in this example. This does not help to improve the measurement accuracy because the average value still has systematic errors. This motivates us to design a method to reduce the systematic errors.

First, the measured values in the cross-validation (Table 2) are normalized to relative errors (see Table 3), where $\varepsilon_{ij} = \frac{x_{ij} - x_j^*}{x_j^*}$ and $x_j^* = 1.0$. Then the average of each row $\varepsilon_{i.}$ is used for the calibration. If $\varepsilon_{i.} > 0$, the values measured with the i-th ruler are likely (in the sense of

average) larger than the actual values. For example, in Table 3, the errors in the first row are all non-negative and the average (i.e., 0.25) is positive. This means that the measurement obtained by the first ruler is always larger. Therefore, the first ruler $\boxed{0.8}$ can be regarded as a *shortened ruler*, i.e., its unit width is larger than that of a standard ruler. Similarly, the fifth ruler $\boxed{1.2}$ is a *lengthened ruler*.

In Table 3, the average of the relative errors 0.0207 is a small number. This means that the cross-validation results are self-consistent and this set of rulers is suitable for the measurements.

Therefore, we suggest using a scaling factor $W_{gi} = \frac{1}{1+\varepsilon_{i.}}$ to reduce the difference and let the calibrated value be:

$$\hat{x}_{\text{calibrated}} = W_{gi} * x_{\text{measured}}.$$

The scaling factor W_{gi} reflects the global property of an individual ruler in the sense of the average. Thus it is called the weight of global calibration. With the calibration coefficients $W_{gi}, i = 1, 2, \cdots, 5$, the measured distances in the last column of Table 2 can all be calibrated to the actual value 0.95 in this case. The result indicates that WCV completely remove systematic errors if the ticks of the rulers are uniform and the bias of the rulers is symmetric about zero.

Weight of Local Fine-Tuning

In practical force measurement applications, the situation might be even worse: the identified transfer function may not be scaled linearly. This is comparable to an unequal distribution of the ruler's tick marks. It is evident that a nonlinear distribution of the tick marks will result in larger errors.

If the degree of nonlinearity of a ruler is large, the relative errors of cross-validation are more scattered and the measured values are more unreliable. We can use the standard deviation of the relative errors ($\sigma_i = \sqrt{\frac{1}{n-1}\sum_j^n (\varepsilon_{ij} - \varepsilon_{i.})^2}$) to represent the nonlinearity degree of the i-th ruler. It is apparent that the standard deviation satisfies $0 < \sigma_i < 1$ and usually $0 < \sigma_i < 0.2$. In fact, a large nonlinearity degree means that the measurement results have more uncertainties and the ruler is unreliable and should be abandoned. In practical applications, one can use a small number σ_0 (0.2 for instance) as the tolerance limit so that $\sigma_i < \sigma_0$. Then, if $\sigma_i \geq \sigma_0$, the ruler (or the transfer function) is considered invalid and is discarded.

The nonlinearity degree σ can be used to define another scaling factor W_l to balance the weight of each ruler's measured value in the final result. We define

$$W_{li} = \frac{\ln(\sigma_i)}{\sum_{j=1}^{n} \ln(\sigma_j)} \tag{4}$$

Table 2 Measured values of a nominal distance of 1.0 m for each ruler and an actual length of 0.95 m with a set of non-standard rulers

	0.8	0.9	1.0	1.1	1.2	[0.95]
0.8 ⇒	1.0000	1.1250	1.2500	1.3750	1.5000	1.1875
0.9 ⇒	0.8889	1.0000	1.1111	1.2222	1.3333	1.0556
1.0 ⇒	0.8000	0.9000	1.0000	1.1000	1.2000	0.9500
1.1 ⇒	0.7273	0.8182	0.9091	1.0000	1.0909	0.8636
1.2 ⇒	0.6667	0.7500	0.8333	0.9167	1.0000	0.7917
Average						0.9697

Table 3 Relative errors when using the non-standard rulers and the calibrated values of the actual distance of 0.95 m

	0.8	0.9	1.0	1.1	1.2	μ	Calibrated
0.8 $\Rightarrow$	0.0000	0.1250	0.2500	0.3750	0.5000	0.2500	0.95
0.9 $\Rightarrow$	−0.1111	0.0000	0.1111	0.2222	0.3333	0.1111	0.95
1.0 $\Rightarrow$	−0.2000	−0.1000	0.0000	0.1000	0.2000	0.0000	0.95
1.1 $\Rightarrow$	−0.2727	−0.1818	−0.0909	0.0000	0.0909	−0.0909	0.95
1.2 $\Rightarrow$	−0.3333	−0.2500	−0.1667	−0.0833	0.0000	−0.1667	0.95
Average						0.0207	0.95

in our applications. Of course, the user may define W_{li} by another decreasing function of σ_i, such that $0 \leq W_{li} \leq 1$ and $\sum_{i=1}^{n} W_{li} = 1$.

The definition of W_{li} involves a weight scaling function $s(\sigma)$ that determines how quickly the weighting should decrease with the increase in σ, which is $s(\sigma) = -\ln(\sigma)$ in equation (4). In fact, equation (4) can be rewritten as $W_{li} = \frac{-\ln(\sigma_i)}{\sum_{j=1}^{n} -\ln(\sigma_j)}$. The scaling factor W_l is called the weight of local fine tuning because the nonlinear distribution of the tick marks is a local property of the ruler.
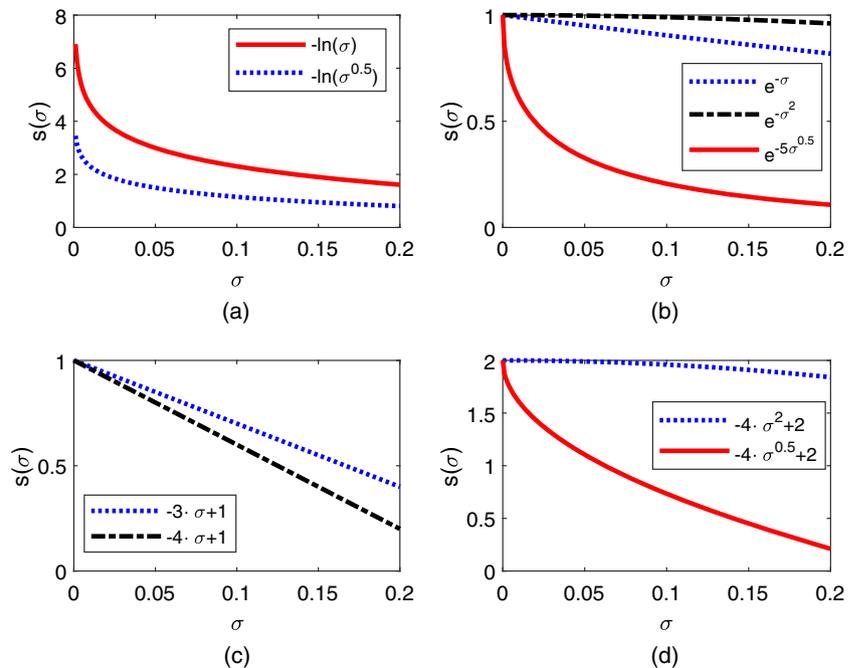
The weight scaling function should be carefully designed to provide larger weight the better *ruler*. Several types of functions including logarithmic, exponential, linear, and polynomial functions have been considered (Fig. 1). Note that the parameter σ_i satisfies $0 < \sigma_i < 0.2$ in most cases. Therefore, the scaling function should decrease convexly, have a significant (e.g., 80%) decrement from 0 to 0.2, and steadily approach a positive value as $\sigma_i \rightarrow 0.2$. On the other hand, keep in mind that W_{li} is defined only for local fine tuning, the decrement factor should not be too large (e.g., $s(0.2)/s(0.001) < 5$) to avoid over-weighting. Figure 1 shows that $s(\sigma) = -\ln(\sigma)$ and $s(\sigma) = e^{-5\sigma^{0.5}}$ are good function candidates. The logarithmic function $-\ln(\sigma)$ is applied in this study for its simplicity (i.e., it does not involve additional parameters and the minus sign can also be eliminated).

Suppose the measurement value is a random variable distributed normally with mean μ and variance σ^2. According to the empirical rule in statistics, for the measurement with the i-th ruler (or the i-th transfer function), the observation lies in the interval $[\mu - 2\sigma, \mu + 2\sigma]$ with 95.45% confidence. Thus the radius of the measurement confidence interval (of 95.45%) without local tuning is $r = 2\sum_{j=1}^{n} \sigma_i$. With the weight scaling function $s(\sigma)$, the scaled radius satisfies

$$r' = 2\sum_{i=1}^{n} \sigma_i \cdot s(\sigma_i) / \sum_{j=1}^{n} s(\sigma_j).$$

Fig. 1 Candidates for the scaling function

Fig. 2 Division of a sample of balance signals
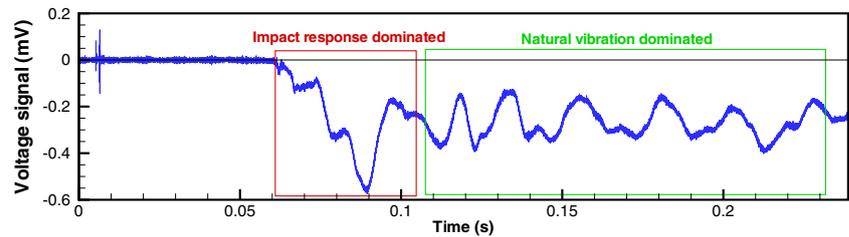
It is apparent that $r' \leq r$ (and $r' < r$ in practice) because the scaling function $s(\sigma)$ is a decreasing function. This means that the weight scaling results in measurement results with a narrower confidence interval.

Experimental Verifications

Calibration methods are usually verified with historical data or theoretical formulas of standard test models in the literature. However, for a high enthalpy wind tunnel, this verification method may experience difficulties. In fact, the working conditions (such as the parameters of free stream, boundary conditions, and measurement system) may differ from wind tunnel to wind tunnel. Even for the same test model, test data from different wind tunnels cannot be directly compared and correlated. For example, it is common that the actual Mach number and unit Reynolds number could be quite different in practice although the nominal Mach number of test flows in two wind tunnels is the same. Moreover, theoretical formulas of standard test models are derived using simplifications and assuming the use of an ideal gas. Actual factors such as the real gas and viscosity effects are ignored.

In this study, the proposed method is verified by comparing it with wave system fitting (WSF) [16], which is a verified method. WSF decomposes the balance signal to separate the vibration waves using the following model.

$$x(t) = f + \sum_{i=1}^{m} A_i \sin(2\pi \omega_i t + \varphi_i) + \varepsilon. \tag{5}$$

where the first term f (aerodynamic force) is a constant and represents the direct current (DC) component of the signal, the second part represents the low-frequency signals, and the last term ε contains high-frequency signals and random noise. The parameter m in equation (5) is a choice for the number of low frequency waves to be removed. The amplitudes A_i, frequencies ω_i, and initial phases φ_i are computed by a global optimization algorithm. WSF separates and removes low-frequency vibration waves without influencing the DC component and it not affected by the completeness of the cycle of the sample signal.

Note that both WCV and WSF compute the target aerodynamic force with balance signals from the same wind tunnel tests. Thus the working conditions are the same and the results have better comparability than those from different wind tunnels.

More specifically, the proposed method is verified with experiments conducted in the JF-12 shock tunnel [17]. This is a detonation-driven wind tunnel, capable of duplicating pure airflow conditions of Mach numbers 5 to 9 at altitudes from 25 to 50 km with a test duration of more than 100 ms; this shock tunnel was developed by Jiang et al. [17]. Due to the relatively long test duration, in addition to the impact-response dominated signals, the test signal curve of the JF-12 shock tunnel includes an extra part of the periodic signals of natural vibration (see Fig. 2). The periodic signals are more informative and thus easier to analyze, and were investigated in our previous studies [15, 16]. The results are reliable and the accuracy is within 2%. This gives us the opportunity to evaluate the performance of the proposed WCV method.

In this study, only the first part of the balance signal is used to calculate the target force using the WCV method. The second part is used for the verification of the WCV's performance (see Fig. 2).

The test model is a sharp cone of 10 degree half-angle with length 750 mm. The support system and the strain gauge balance are the same as in References [16] and [15]. For detailed schematic of the test model and sensor, we refer the readers to [15].

Three wind tunnel experiments were carried out; the working conditions include the angle of attack α, total temperature T_0, total pressure P_0, and Mach number M are listed in Table 4.

Linearity Verification of the Force Measurement System

Before applying the proposed WCV method, it must be confirmed that the dynamics of the measurement system

Table 4 Working conditions of the three wind tunnel tests

Shot no.	α (degree)	T_0 (K)	P_0 (MPa)	M
#20150428	5.2	1889	2.38	7.05
#20150430	5.2	2336	2.34	7.06
#20150507	13.8	2338	2.44	7.04

is linear. The structural joints and boundary supports may affect its linearity [19]. Therefore, further investigations are required to confirm the linearity of the dynamics.

Two kinds of experiments were carried out to perform the transient response analysis: hammer knocking and wire cutting (Fig. 3). The hammer knocking tests are used for a qualitative analysis to verify the linearity of the force measurement system and the wire-cutting experiments are used for a quantitative analysis to determine the transfer function of the system for the force measurement.

The hammer knocking test is used to simulate an impulse input signal. However, it is far from a perfect impulse, which makes it difficult to determine the transfer function. In this study, a global optimization algorithm, i.e., low-dimensional simplex evolution [18], is used to determine the coefficients of the transfer functions. The optimization model is as follows.

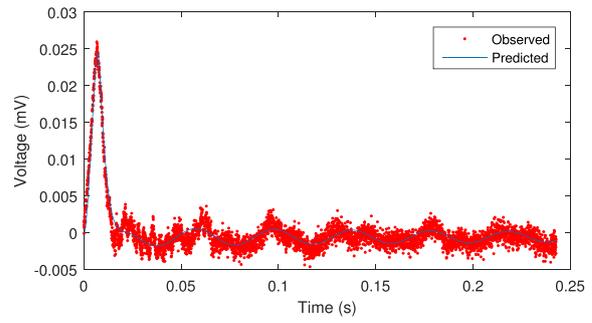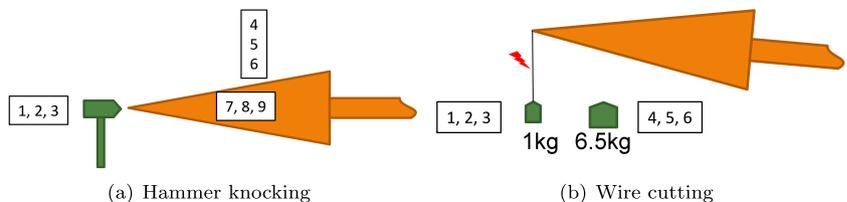$$\min_{a_i, b_j, n} \left\| y(t) - y_{\hat{H}(s)}(t) \right\|$$

where a_i, b_j, and n are the coefficients of the transfer function $\hat{H}(s)$ to be determined as in equation (1). A series of knocking tests were performed. The first 3 knocks are in the axial direction, the next 3 knocks are in the normal direction, and the last 3 knocks are in lateral direction (Fig. 3(a)). The optimization results show that a fifth-order transfer function fits the axial response, and a sixth-order function fits the normal and lateral responses. Increasing or reducing the order of the transfer function results in larger fitting errors. The cross-prediction results (Fig. 4) also confirm these findings. Thus the conclusion is that the dynamics of the force measurement system in the JF-12 is linear, and has a fifth-order function in axis direction, and sixth-order functions in normal and side directions.

Calibration Experiments in Wind Tunnels

Note that the results of the hammer knocking inevitably involve human factors. For example, it is very difficult for a person to execute a perfect knock normal to the test model in exactly the same manner for each test. Therefore, hammer knocking is used only for a qualitative analysis. In contrast, the results of the wire-cutting experiments are more

(a) Output prediction of knocking test number 2 with the transfer function determined from knocking test number 1
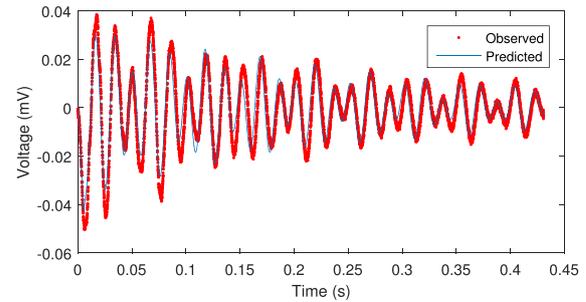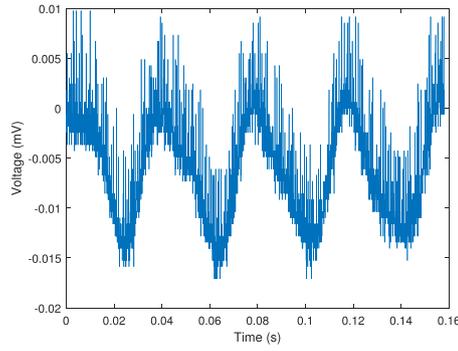

(b) Output prediction of knocking test number 6 with the transfer function determined from knocking test number 5

Fig. 4 Cross-prediction results of knocking tests

objective and we use them for the quantitative measurements and calibrations. Theoretically, the proposed WCV method can be used to calibrate the measurement of six-degree-of-freedom forces. Note that the aim is to test the effectiveness of WCV. Therefore, the most reliable result, i.e., the normal force, is used for method verification. Only normal forces are considered in this section for method verification.

Prior to each wind tunnel experiment, six wire-cutting experiments are conducted with two different suspended weights (1 kg and 6.5 kg, see Fig. 3(b)). The resulting signals of the balance outputs are shown in Fig. 5(a). Six transfer functions were identified by integral transformation with the step inputs and balance outputs of the wire-cutting experiments. The relative errors ε_{ij} are obtained by cross-validation and are listed in Table 5, where the vertical component of the weights f_j^* is known and $\varepsilon_{ij} = \dfrac{f_j^{(i)} - f_j^*}{f_j^*}$. In

Fig. 3 Experimental settings for the linearity verification. The Arabic numerals denote the test numbers and the knocking in different directions (left) or using different mass weights (right)

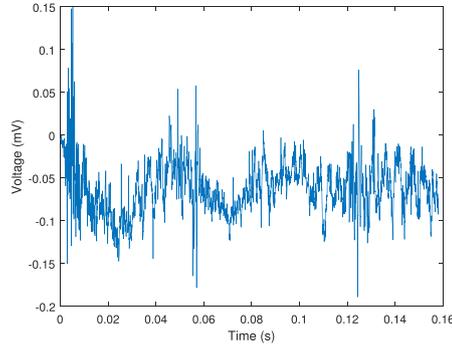(a) Hammer knocking

(b) Wire cutting

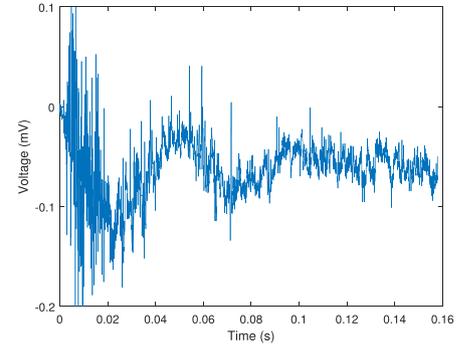**Fig. 5** Balance signals of the impact response

(a) Balance signals of the wire-cutting experiment
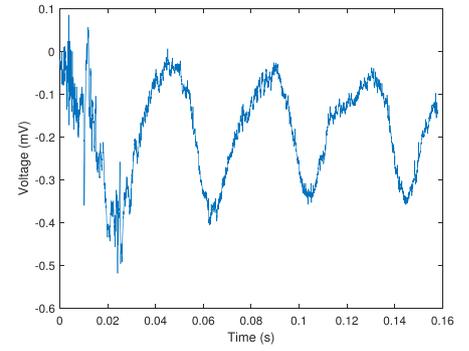
(b) Balance signals of shot #20150507

(c) Balance signals of shot #20150428

(d) Balance signals of shot #20150430

Table 5, the errors in the first row are all non-positive and the average (-0.1042) is negative. This means the prediction based on the transfer function obtained from the first wire-cutting experiment always results in a lower value than the actual value. Therefore, the transfer function $H^{(1)}(s)$ can be regarded as a *lengthened ruler*. Similarly, the transfer function determined by the third wire-cutting experiment $H^{(3)}(s)$ can be regarded as a *shortened ruler*.

The key parameter, the total average of relative errors is 0.0049 (Table 5). It is small enough to indicate that the cross-validation results are self-consistent and this set of *rulers* (i.e., transfer functions) can be used for the wind-tunnel measurement.

For the first two shots (i.e., #20150428 and #20150430), the angles of attack are the same ($\alpha = 5.2$ degree) and the model-balance-sting system remains unchanged. Therefore, the transfer functions of the impact response should be identical and the two shots can share the same calibration table. For the third wind-tunnel run (#20150430), the angle of attack is changed to 13.8 degrees. The calibration experiments are conducted separately and the calibration table is reconstructed (omitted here).

After the determination of the transfer functions, the wind tunnel experiment is conducted. The balance signal is divided into two parts as shown in Fig. 2; the first part (0 to 50 ms in this study) is impact-response dominated and

Table 5 Table of cross-validation for calibration

	①	②	③	④	⑤	⑥	μ	σ
①⇒	0	−0.1193	−0.182	−0.1335	−0.1536	−0.0368	−0.1042	0.0707
②⇒	0.1345	0	−0.0708	−0.016	−0.0388	0.0939	0.0171	0.0799
③⇒	0.2214	0.0758	0	0.0591	0.0346	0.1773	0.0947	0.0861
④⇒	0.1538	0.016	−0.0561	0	−0.0232	0.1117	0.0337	0.0816
⑤⇒	0.1811	0.0401	−0.0337	0.0236	0	0.1378	0.0582	0.0834
⑥⇒	0.0379	−0.086	−0.1508	−0.1004	−0.1212	0	−0.0701	0.0733
Average	0.1215	−0.0122	−0.0822	−0.0279	−0.0504	0.0807	0.0049	0.0792

Table 6 Predicted normal forces of wind-tunnel runs with the transfer functions obtained from wire-cutting experiments (Unit: Newton)

	Shot #20150428	Shot #20150430	Shot #20150507	Average
$\hat{f}_1$	96.82	98.69	288.79	
$\hat{f}_2$	109.11	101.03	291.43	
$\hat{f}_3$	97.1	88.78	334.68	
$\hat{f}_4$	99.89	85.52	319.00	
$\hat{f}_5$	95.28	97.74	300.55	
$\hat{f}_6$	102.54	104.53	300.78	
$\hat{f}$	100.38	96.66	305.41	
f_{ref}	107.5	102	288.94	
Error	−6.63%	−5.23%	5.70%	5.85%

second part (50 to 160 ms) is natural-vibration dominated. They are analyzed separately as follows.

The first part is used to calculate the aerodynamic forces using WCV. With the transfer functions obtained from the wire-cutting experiments $H^{(i)}(s)$ ($i = 1, 2, \cdots, 6$), it is easy to compute the average of the force inputs $\hat{f}_i$ by deconvolution. The results are listed in Table 6. The calibrated force $\hat{f}$ is then calculated using equation (3). In Table 6, it is observed that the recovered force ranges from 95.28 N to 109.11 N for Shot #20150428, from 85.52 N to 104.53 N for Shot #20150430, and from 288.79 to 334.68 for Shot #20150507 if we use a conventional transient response analysis method. The uncertainty relative errors are 12.9%, 18.6%, and 15.0% respectively.

The second part contains periodic signals and is more informative. It is analyzed by the WSF method (for details, see Ref. [16]) and the fitted aerodynamic force is used as the reference value f_{ref} to evaluate the performance of WCV (shown in Table 6).

The errors of the WCV calibrated values are −6.63%, −5.23%, and 5.70% respectively. The average of the absolute errors is 5.85% (see the last row of Table 6). The accuracy is acceptable for practical engineering applications and is much smaller than the uncertainty relative errors.

Discussion

The proposed WCV method is designed to calibrate force measurements for general impulse facilities; this method is suitable for situations when neither steady signals nor periodic signals are available. The shock tunnel JF-12 with a relatively long test time was used to test the performance of the WCV method. It is worth noting that the WCV method is also useful for the JF-12 shock tunnel or other impulse facilities with relatively long test times. In fact, the JF-12

data contain an additional part of the periodic signals of natural vibration only for normal-sized test models. For large/heavy models (e.g., larger than 4 m and heavier than 500 kg), the periodic signals are not available and the impact signals must be used to recover the aerodynamic forces.

Although the WCV method is originally proposed for aerodynamic force measurements of hypersonic vehicles in impulse facilities, the concept of the WCV method is general and is applicable to other indirect measurement problems that require the reduction of systematic errors, especially when no exact/standard measurement tools are available; examples include the indirect measurement of thermal conductivity in silicon nanowires [20] and soot size distributions in compression ignition engines [21]. The workflow is the same: obtain a set of *rulers* (not necessarily transfer functions) by conducting a series of calibration experiments, create a cross-validation table of relative errors, and assign the measurement result of each ruler a proper weight based on the cross-validation table (see equation (3)).

Conclusion

A calibration method to reduce the systematic errors, weighting by cross-validation (WCV), for force measurements via transient response analysis has been proposed. In the WCV method, a series of on-site calibration experiments are performed, and a cross-validation table of relative measurement errors is created. The aerodynamic force is then calibrated using the average and standard deviation of the cross-validation errors. The identified transfer functions are cross-validated against each other to determine the weight of each calibration experiment. The calibration experiments are much cheaper than wind-tunnel tests and do not result in much additional cost.

The working mechanism of the WCV method is demonstrated by using the analogy of obtaining measurements with a set of non-standard rulers. This differs from the situation in which several measurements are obtained using an ideal ruler with random errors. In the latter case, the error can be nearly eliminated if the average of the measurements is used.

The effectiveness of WCV has been verified by the signal analysis of tests conducted in the JF-12 shock tunnel. Due to the relatively long test duration, the test signal curve includes an additional set of periodic signals of natural vibration in addition to the impact-response dominated signals. The periodic signals are much more informative when the WSF is used for the analysis. This provides us an opportunity to evaluate the performance of the proposed WCV method. Our results indicate that WCV improves the measurement accuracy significantly.

Although the WCV method is originally proposed for aerodynamic force measurements of hypersonic vehicles in impulse facilities, the concept of the WCV method is general and is applicable to other indirect measurement problems to reduce systematic errors. Research is underway to use WCV in other applications.

References

1. Laurence SJ, Butler CS, Martinez Schramm J, Hannemann K (2018) Force and moment measurements on a free-flying capsule in a shock tunnel. J Spacecr Rocket 55(2):403–414
2. Meng B, Han G, Luo C, Jiang Z (2018) Numerical investigation of the axial impulse load during the startup in the shock tunnel. Aerosp Sci Technol 73:332–342
3. Qiu H, Min F, Zhong S, Song X, Yang Y (2018) Hypersonic force measurements using internal balance based on optical micromachined Fabry-Perot interferometry. Rev Sci Instrum 89(5):035004
4. Jessen C, Groenig H (1989) A new principle for a short-duration six component balance. Exp Fluids 8(3–4):231–233
5. Joshi MV, Reddy NM (1986) Aerodynamic force measurements over missile configurations in IISc shock tunnel at $M_\infty = 5.5$. Exp Fluids 4(6):338–340
6. Raju C, Reddy NM (1990) Aerodynamic force measurements over missile configurations in IISc shock tunnel at $M_\infty = 3.85$ and 9.15. Exp Fluids 10(2):175–177
7. Naumann KW, Ende H, Mathieu G, George A (1993) Millisecond aerodynamic force measurement with side-jet model in the ISL shock tunnel. AIAA J 31(6):1068–1074
8. Tanno H, Komuro T, Sato K, Fujita K, Laurence SJ (2014) Free-flight measurement technique in the free-piston high-enthalpy shock tunnel. Rev Sci Instrum 85:045112
9. Simmons JM, Sanderson SR (1991) Drag balance for hypervelocity impulse facilities. AIAA J 29(12):2185–2191
10. Smith AL, Mee DJ, Daniel WJT, Shimoda T (2001) Design, modelling and analysis of a six component force balance for hypervelocity wind tunnel testing. Comput Struct 79(11):1077–1088
11. Mee DJ (2003) Dynamic calibration of force balances for impulse hypersonic facilities. Shock Waves 12(6):443–455
12. Satheesh K, Jagadeesh G (2009) Analysis of an internally mountable accelerometer balance system for use with non-isotropic models in shock tunnels. Measurement 42(6):856–862
13. Vadassery P, Joshi DD, Rolim TC, Lu FK (2013) Design and testing of an external drag balance for a hypersonic shock tunnel. Measurement 46(7):2110–2117
14. Pallekonda RB, Nanda SR, Dwivedy SK, Kulkarni V, Menezes V (2018) Soft computing based force recovery technique for hypersonic shock tunnel tests. Int J Struct Stab Dyn 18(5):1871004
15. Wang Y, Liu Y, Luo C, Jiang Z (2016) Force measurement using strain-gauge balance in a shock tunnel with long test duration. Rev Sci Instrum 87:055108
16. Luo C, Wang Y, Wang C, Jiang Z (2015) Wave system fitting: a new method for force measurements in shock tunnels with long test duration. Mech Syst Signal Process 62–63:296–304
17. Jiang Z, Yu H Experiments and development of Long-test-duration Hyper velocity Detonation-driven Shock Tunnel (LHDst), AIAA Paper, 2014-1012
18. Luo C, Yu B (2012) Low dimensional simplex evolution: a new heuristic for global optimization. J Glob Optim 52(1):45–55
19. Lin RM, Ng TY (2018) A new method for the accurate measurement of higher-order frequency response functions of nonlinear structural systems. ISA Trans. Online first. https://doi.org/10.1016/j.isatra.2018.05.015
20. Pennelli G, Nannini A, Macucci M (2014) Indirect measurement of thermal conductivity in silicon nanowires. J Appl Phys 115:084507
21. Martos FJ, Martín-González G, Herreros JM (2018) Semi-empirical model for indirect measurement of soot size distributions in compression ignition engines. Measurement 124:32–39