Direct and Inverse Calibration Curves of Measuring Instruments: Selection and Fitting

Valerii Granovskii[a] and Tatiana Siraia

Concern CSRI Elektropribor, JSC, 197046, 30, Malaya Posadskaya, St. Petersburg, Russia,

Abstract. In the paper the methodological problems of the measuring instrument calibration are presented. The key problems are connected with the choice of the form of calibration curve, and the methods of data processing. Two main cases are considered. The first one is when the least squares method is applicable. The second case is when this method is inconsistent, so the various confluent methods should be applied. The efficiency of method is described by the bias and second central moment of estimate. These problems and methods are specified for the particular case of linear calibration curve.

1 Introduction

Calibration curves are widely used as primary metrological characteristics of various measuring instruments, such as sensors, measuring transducers, and devices. The quality of calibration curves determines the accuracy of measuring instruments and measurements to a great extent, which conditions the importance of the methods for construction and testing of calibration curves, and also the methods for estimation of calibration [1].

Problems of calibration arise at various stages of measuring instrument's life cycle. The actual stage determines the matter of the problem.

So, the calibration of measuring instrument of the approved type is just a routine procedure with appointed parameters. Accordingly, the form of calibration curve, and the algorithm of data processing are well defined here.

Quite different is the case of the primary calibration, which is carried out while the making of the measuring instrument, just before the type approval tests. In this case the calibration problem arises in the most complete, interesting and complicated form.

Really, at the primary calibration, the initial information on the real calibration curve is only presented as the normative requirements, which are stated in technical assignment for the instrument construction. Usually these requirements are stated for the precision (or uncertainty) of the calibration curve, and also for the curve non-linearity (as in the most widespread linear case).

Thus, the problem of the primary calibration is formally presented as the problem of approximation for the real dependence, which is fuzzy defined.

In addition, as it will be shown below, the choice of the form of calibration curve should not be free or conditioned by merely algorithmic or computational reasons.

Taking into account all the stated above, we would further consider the construction of calibration curve in the primary calibration procedure.

2 General Formal Scheme

To construct a calibration curve, the joint measurements are usually performed for a sequence of input quantities X_i, and also for corresponding output quantities Y_i; so the set of experimental data is obtained:

$$\{ (x_i, y_i), i = 1 \ldots m \} \qquad (1)$$

Most often, a calibration curve is presented in analytical form, as a functional dependence of the output quantity Y upon the input quantity X:

$$Y = f(X). \qquad (2)$$

Here the function f is selected based on a priori data and practical requirements, including physical principles of measuring instrument operation.

Further, the constructed calibration curve is generally used in the inverse way, that is, to estimate the input quantities on basis of the measured output quantities:

$$X = f^{-1}(Y). \qquad (3)$$

---

[a] Corresponding author: vgranovsky@eprib.ru

For example, for a linear calibration curve one obtains the following estimates:

$$Y = a + b\,X \quad \Rightarrow \quad x = (y - a)/b \qquad (4)$$

Therefore in practice it was proposed to construct the inverse calibration curve immediately:

$$X = g\,(Y). \qquad (5)$$

Namely, for the linear function one obtains directly:

$$X = c + d\,Y \quad \Rightarrow \quad x = c + d\,y \qquad (6)$$

Commonly the selection of the functional form f of calibration curve (2) is given more attention in literature. However, it should be emphasized that beforehand just the form of the curve presentation, either direct type (2) or inverse (5), should be selected. Therefore, all the above mentioned problems should be solved as a whole, with the quality of the final estimates of input values serving as the main criterion.

This paper considers the rational selection of the curve form and the main construction methods. In particular, the following issues are discussed:

- selection of a rational way for presenting the calibration curve: direct or inverse;
- selection of the adequate functional form: linear, polynomial, or some others;
- selection of a fitting method: least squares method, or some confluent methods.

3 Classical Methods of Calibration

Assume that experimental data (1) used to construct the calibration curve (2) are described by the general model of the following structure [2]

$$x_{ij} = X_i + \theta_{xi} + \varepsilon_{xij}; \qquad i=1...m,$$

$$y_{ij} = f\,(X_i) + \theta_{yi} + \varepsilon_{yij}, \qquad j=1...n_i, \qquad (7)$$

where X_i, $Y_i = f\,(X_i)$ are the true input and output values; θ_{xi}, θ_{yi} are the systematic errors of measurements X_i, Y_i; ε_{xij}, ε_{yij} are the random errors of measurements X_i, Y_i.

The main scenarios of calibration curve construction are classified based on the relations of errors characteristics in the input and output values.

For these scenarios the basic construction methods may be described. Traditionally, the main technique is the classical least squares fitting, but it is relied on rather strict assumptions on error characteristics in general model (7). These are briefly as follows [2, 3]:

a) the measurements errors of input values X_i are insignificant (in comparison with the errors of output values Y_i);

b) the measurements errors of output values Y_i are random;

c) the measurements errors of Y_i have approximately Gaussian distributions.

The violation of these conditions leads to that least squares estimates lose their efficiency.

For example, with significant measurements errors of input values X_i the least squares estimates turn out statistically inconsistence. To be more precise, the random errors in X_i cause the systematic error in f, which is asymptotically unavoidable.

In particular, for the coefficient b of linear function (4) the least squares estimates are asymptotically biased (as $m \to \infty$):

$$b = \sum_1^m \left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right) \Big/ \sum_1^m \left(x_i - \bar{x}\right)^2 \cdot \to b_0 - \Delta, \qquad (8)$$

where $\Delta = b\,(m-3)\,\sigma_x^2 \Big/ \sum_1^m \left(x_i - \bar{x}\right)^2$,

σ_x^2 – variance of random errors of x_i.

Thus, if the classical condition (a) is fulfilled, then it is reasonable to construct the direct calibration curve (2), which would be statistically consistent, and then to invert the curve as (3). As a result, the estimated inverse curve (3) would be also consistent.

Besides, it is easy to see, that if one immediately constructs the inverse calibration curve as (5), it would lead to inconsistent results.

The presence of the systematic errors in data (contrary to condition b) is not too restrictive, and the least squares estimates are generally valid. But it is essential, that the uncertainties of the results should be estimated taking into account the systematic errors of data.

4 Confluent Methods of Calibration

Nowadays it is essential to consider the problem of calibration in more extensive statement, when the following characteristic properties take place:

A1) the measurements errors of input values X_i are significant;

C1) the measurements errors of Y_i have non-Gaussian distributions.

For these cases, when the least squares fitting cannot be straight applied, confluent or robust methods may be employed.

There are several groups of suitable confluent estimates, which may be used instead of least squares ones. These methods rely upon the supplementary information, which is either known a priori, or may be obtained from additional experiments.

For instance, in the case of linear calibration curves there are following main groups of confluent estimates:

α) generalized orthogonal regression estimates (when the ratio of error variances $\lambda = \sigma_y^2/\sigma_x^2$ is known);

β) corrected least squares, or variance analysis estimates (when the multiple observations in points X_i, Y_i are known, and the variances σ_y^2, σ_x^2 are estimated).

γ) homographic or linear fractional estimates (when the increasing order of true values X_i is a priori known).

Characteristics of various confluent estimates of the line coefficient b are presented in Tables 1 and 2, as compared with the least squares estimates.

Table 1. Confluent estimates for coefficient b of linear calibration curve

A priori information		Estimate b	Bias $B(b)$	Second moment M_2
known parameters	$\lambda = \sigma_y^2 / \sigma_x^2$, or estimate	$b_1 = v \pm \sqrt{v^2 + \lambda}$, $v = \left(\sum_y - \lambda \sum_x\right)/2\sum_{xy}$	$\dfrac{2\sigma_y^2 + b^2 \sigma_x^2}{b\sum_x}$	$\dfrac{\sigma_y^2 + b^2 \sigma_x^2}{\sum_x}$
	σ_x^2 or estimate	$b_2 = \sum_{xy} / \left(\sum_x - (m-1)\sigma_x^2\right)$	$2b\sigma_x^2 / \sum_x$	
	σ_y^2 or estimate	$b_3 = \left(\sum_y - (m-1)\sigma_y^2\right)/\sum_{xy}$	$\left(1 - \dfrac{\lambda}{b^2}\right) b\sigma_x^2 / \sum_x$	
estimates from additional experiments	estimate of λ, $n_i = kr_i$	$b_1 = v \pm \sqrt{v^2 + \lambda/k}$, $v = \left(\sum_y - \lambda\sum_x\right)/\left(2k\sum_{xy}\right)$	$B(b_1)$	$\dfrac{\sigma_y^2 + b\sigma_x^2}{\sum_x}$
	S_x^2	$b_2 = \sum_{xy} / \left(\sum_x - (m-1)S_x^2\right)$	$B(b_2)$	
	S_y^2	$b_3 = \left(\sum_y - (m-1)S_y^2\right)/\sum_{xy}$	$B(b_3)$	
	S_x^2 и S_y^2, $n_i = kr_i$	$b_4 = \dfrac{\pm\left(\sum_y - (m-1)S_y^2\right)}{\left\{k\left[\sum_x - (m-1)S_x^2\right]\right\}}$	$\left(1 - \dfrac{\lambda}{b^2}\right)\dfrac{b\sigma_x^2}{2k\sum_x}$	$\dfrac{\sigma_y^2 + b^2 \sigma_x^2}{k\sum_x}$
Least squares estimate		$b_0 = \sum_{xy} / \sum_x$	$-\dfrac{b\sigma_x^2(m-3)}{\sum_x}$	$\dfrac{\sigma_y^2 + b^2 \sigma_x^2}{\sum_x}$

Table2. Homographic, or linear fractional estimates for coefficient b of linear calibration curve

A priori information	Estimate b	Bias $B(b)$	Second moment M_2
Order of true values X_i	$b = \dfrac{\sum \omega_i y_i}{\sum \omega_i x_i}$	$\dfrac{b\sigma_x^2 W^2}{W(X)}$	$\dfrac{\sigma_0^2 W^2}{W^2(X)}$
Clustering	$b_g = \dfrac{\sum n_i(N - 2N_i + n_i)\bar{y}_i}{\sum n_i(N - 2N_i + n_i)\bar{x}_i}$	—	
Clustering for 2 or 3 groups	$b_W = \dfrac{y^{II} - y^{I}}{x^{II} - x^{I}}$	$\dfrac{mb\sigma_x^2}{\left(x^{II} - x^{I}\right)^2}$	$\dfrac{m\sigma_0^2}{\left(x^{II} - x^{I}\right)^2}$
Uniform location of values X_i	$b_H = \dfrac{\sum y_i(m + 1 - 2i)}{\sum x_i(m + 1 - 2i)}$	$\dfrac{12b\sigma_x^2}{h^2 m(m^2 - 1)}$	$\dfrac{12\sigma_0^2}{h^2 m(m^2 - 1)}$
Least squares estimate	$b_0 = \dfrac{\sum y_i(x_i - \bar{x})}{\sum x_i(x_i - \bar{x})^2}$	$-\dfrac{b\sigma^2(m-3)}{\sum_x}$	

The main characteristics of confluent estimates are the bias $B(b)$ and the second moment $M_2(b)$, which are defined as follows:

$$B(b) = M(b) - b_0, \qquad (9)$$

$$M_2(b) = M(b - b_0)^2 = \sigma^2(b) + B^2(b), \quad (10)$$

where b_0 is the true value of coefficient, and M is the symbol of expectation.

At the Table 1 two groups of confluent estimates are presented: orthogonal regression and the corrected least squares estimates. Here the sums are as follows:

$$\sum_y = \sum_1^m (y_i - \bar{y})^2 ; \quad \sum_{xy} = \sum_1^m (x_i - \bar{x})(y_i - \bar{y}).$$

It is easily seen, that in all the cases the confluent estimates b_1 - b_4 are consistent, as $B(b)$ tends to 0, but the least squares estimate b_0 is here inconsistent.

At the Table 2 the main groups of homographic, or linear fractional estimates are presented. They are quite different from the least squares estimates, and they are based on specific a priori information on the order or location of the true values X_i. These estimates are also consistent, and their characteristics are compared with those of least squares estimates.

Parameters are denoted as follows. Firstly, ω_i are weights, which are ascribed to the values X_i, with the sum $\Sigma \omega_i = 0$. Here the best weights would be $\omega_i = X_i - \bar{X}$, but these values are obviously unknown; and they are approximatedly used in least squares estimates.

a Corresponding author: vgranovsky@eprib.ru

Secondly, in case of clustering all the data are divided into several groups, which are independent from the errors of data. In practice, the most useful are estimates in two cases : small number - two or three groups, and uniformly located values X_i.

Thus, if condition (A1) is fulfilled, then it would be reasonable to construct the inverse calibration curve (5) straight away, using one of confluent methods. This curve would be consistent, and rather handy in practice.

The problem is that there is a wide choice of confluent methods, and a happy choice could significantly increase the efficiency of calibration curve. Some general recommendations are presented in [2, 4]; but further investigations are also needed for the specific cases of calibrations.

5 Calibration as Modelling Problem

The problem of measuring instruments calibration is generally presented as the problem of approximation for the fuzzy dependence. So the crucial point is the adequacy of the resulting calibration curve. As appears from the above, the vagueness of function may be partly overcomed by imposing restraints on the curve, resulting from the matter of the measurement problem, in particular, of the data processing.

The discussed problems of selection and construction of calibration curves can be also interpreted in terms of the theory of model adequacy. Then the calibration curve is considered as a special model of measuring instrument, constructed using both priori data and a posteriori (experimental) data. In particular, estimated uncertainties of calibration curves, which are constructed by different methods, make it possible to evaluate their adequacy within certain data models.

Generally, the total uncertainty of calibration curve may be presented as the sum of two components [5]:

$$\delta f(x) = \delta_c f(x) \ + \ \delta_a f(x) \ , \qquad (11)$$

where $\delta_c f(x) = \delta \ \{f(x) \ / \Omega_0 \ \}$ – the component, obtained under assumption of true model Ω_0 of the curve;
$\delta_a f(x)$ - the inadequacy component, due to the deviations of the real curve from the form, assumed by model Ω_0

The component $\delta_c f(x)$ may be further decomposed as:

$$\delta_c f(x) \ = \delta_m f(x) \ + \ \delta_t f(x) \qquad (12)$$

where $\delta_m f(x) \ = \delta_m \ \{ \ f(x) \ / \Omega_0 \}$ – methodical component, due to the accepted method (procedure) of calibration;
$\delta_t f(x) = \delta_t \ \{ \ f(x) \ / \Omega_0 \ \}$ - transformed component, due to properties of the data processing algorithms and the models of experimental data.

The latter components, δ_m and δ_t, are quite evaluated within the assumed model Ω_0 . On the other hand, the inadequacy component δ_a could be evaluated only by extending the initial model.

Therefore, it is essential to reasonably select the methods of the curve construction depending on available data. Further they provide a basis for creating techniques of testing calibration curves at different stages of their application.

6 Conclusions

As regards to the choice and construction of direct and inverse calibration curves, two significant conclusions can be drawn.

First, if the conditions for the least squares fitting, such as a) – c) stated above, are satisfied, then it is reasonable to construct the direct calibration curve (2), and then invert it as (3). Thus the estimated curve would be statistically consistent, and uncertainties of the results should be estimated with due regard for systematic errors of data.

Second, if the condition a) is not satisfied, then it is reasonable to construct immediately the inverse calibration curve as (5), and also estimate the curve uncertainties with regards for systematic errors.

References

1. L. Finkelstein, Measurement, **14**, 1 (1994)
2. V. Granovskii, T. Siraia, *Methods for Data Processing in Measurement* (Energoatomizdat, Leningrad, 1990, in Russian).
3. D. J. Hudson, *Statistics. Lectures on Elementary Statistics and Probability* (Geneva, 1964).
4. E. Demidenko, *Linear and Non-linear Regression* (Moscow, 1981, in Russian).
5. V. Granovskii, T. Siraia, AMCTM VIII – Ser. Adv. Math. for Appl. Sci., World Scientific (2009.