IMEKO 2010 TC3, TC5 and TC22 Conferences
Metrology in Modern Context
November 22–25, 2010, Pattaya, Chonburi, Thailand

EVALUATION OF UNCERTAINTY OF CONTACT POINT AND CONTACT DEPTH BY USING NON-LINEAR FITTING

Koichiro Hattori, *Yutaka Seino*

National Metrology Institute of Japan, National Institute of Advanced Industrial Science and Technology (NMIJ-AIST), Tsukuba, Japan, hattori-k@aist.go.jp

Abstract – The one of difficulties is how to evaluate the origin of depth from force-indenter displacement curve in instrumented indentation test (IIT). One of determination technique it is that using the fitting using some model function. How to evaluate the uncertainty of the fitted coefficients is another problem in this case. We have been developing the uncertainty evaluation software and the uncertainty of fitting of power-law is evaluated by using the final Jacobian matrix of iteration process. In this paper, we explained the uncertainty calculation for the non-linear function. We applied it to the contact point estimation for force-depth curve obtained by the fused silica as a typical example. We also discussed the effect of model function and data length, which is obtained by the developed software.

Keywords: Instrumented indentation Test, uncertainty of fitting

1. INTRODUCTION

The difficulties of uncertainty evaluation in nano-indentation are which the better estimation method is and how to evaluate the uncertainty in contact point. The most common technique is estimate the contact by monitoring (or back calculating) the velocity and detected the change from before contact. We have been developing the back calculation method using power-law as a fitting function. To develop the software, we would like to know the details of the fitting and what is calculated from the actual force-depth curve. In this report, we developed the fitting software with calculation of uncertainty of fitted coefficients. And we applied it to the zero-point determination using power-law. And we tried to apply the effect of data length and selection of data set, which is used to the analysis. and we discussed the calculated uncertainty and the effect of used functions.

2. STANDARD ERROR FOR FITTING - PARAMETERS OF NON-LINEAR FITTING

The uncertainty calculation for the fitted parameter is well describes for the simple 1st order function, however, the information for the polynomial and non-liner function is not easy to find out. On the other hand, many analysis and graphing software calculates the "standard error" of the coefficients automatically. The detail of calculation is also not easy to find out. Therefore we would like to describe the details of the procedure to calculate standard error in the followings [1,2,3]. It is not new but, it necessary to developing the software apart from the ready for use sub-routings.

In this section, we use the word "standard error" instead of uncertainty of coefficients to easy to understand the previous information. In this section, we only take care about the diagonal element of the covariance that means we ignore the effect of correlation between the coefficients.

2.1. Standard error for linear fitting

We consider the fitting function, which is the liner combination of the function which have a single input variable, x, like a polynomial function,

$$y = a_1 f_1(x) + a_2 f_2(x) + ... + a_k f_k(x) \quad (1)$$

where the y, $f_k(x)$ and a_k are the output, input and the coefficients, respectively. We would like to determine the coefficients a_k using ordinary least-square fitting. Here, the measured values for x_i and y_i and (1) can be rewritten as,

$$\begin{cases} y_1 = a_1 f_1(x_1) + a_2 f_2(x_1) + ... + a_k f_k(x_1) \\ y_2 = ... \\ y_n = a_1 f_1(x_n) + a_2 f_2(x_n) + ... + a_k f_k(x_n) \end{cases} \quad (2)$$

Here we introduce, Jacobian matrix with respect to coefficients, X as

$$\mathbf{X} = \begin{bmatrix} \frac{\partial y_1}{\partial a_1} & \frac{\partial y_1}{\partial a_2} & \cdots & \frac{\partial y_1}{\partial a_k} \\ \frac{\partial y_2}{\partial a_1} & \frac{\partial y_2}{\partial a_2} & \cdots & \frac{\partial y_2}{\partial a_k} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_n}{\partial a_1} & \frac{\partial y_n}{\partial a_2} & \cdots & \frac{\partial y_n}{\partial a_k} \end{bmatrix} = \begin{bmatrix} f_1(x_1) & f_2(x_1) & \cdots & f_k(x_1) \\ f_1(x_2) & f_2(x_2) & \cdots & f_k(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ f_1(x_n) & f_1(x_n) & \cdots & f_k(x_n) \end{bmatrix} \quad (3)$$

Thus, the (2) can be rewritten in matrix form as,

$$\mathbf{y} = \mathbf{Xa} \quad (4)$$

where, **Y** and **a** denote output matrix and coefficient matrix, respectively. We would like to determine the best-fitted coefficients following this relation by using the least-square fit. Thus the best-fitted coefficients, $\hat{\mathbf{a}}$ are determined when the square-sum of error (SSE) is minimized. Following calculation we put the same weight for the all output data. That condition is given by,

$$\frac{\partial}{\partial a_k}\sum_{i=1}^{n}\left(y_i-\sum_{j=1}^{k}(\hat{a}_j f_j(x_i)\right)^2=0 \tag{5}$$

That is rewritten as in matrix form,

$$\mathbf{X^T y}=\left(\mathbf{X^T X}\right)\hat{\mathbf{a}}\,, \tag{6}$$

where the $\mathbf{X^T}$ denotes the transpose of matrix $\mathbf{X}$. The best fitted coefficients should be automatically determined using inverse matrix as,

$$\hat{\mathbf{a}}=\left(\mathbf{X^T X}\right)^{-1}\mathbf{X^T y}=\mathbf{Cy}\,. \tag{7}$$

Here the uncertainty of $\hat{\mathbf{a}}$ may be expected using,

$$\Delta\hat{\mathbf{a}}=\mathbf{C}\Delta\mathbf{y} \tag{8}$$

Then the uncertainty of the coefficients is defined

$$\begin{aligned}\mathbf{U}_{\hat{\mathbf{a}}}^2&=E(\Delta\hat{\mathbf{a}}\Delta\hat{\mathbf{a}})=\mathbf{C}\Delta\mathbf{y}(\mathbf{C}\Delta\mathbf{y})^{\mathbf{T}}\\&=\mathbf{C}E(\Delta\mathbf{y}\Delta\mathbf{y}^{\mathbf{T}})\mathbf{C}^{\mathbf{T}}=\mathbf{C}\mathbf{U}_{\mathbf{y}}^2\mathbf{C}^{\mathbf{T}}\end{aligned} \tag{9}$$

This eq. (8) denotes the covariance of the coefficients. We ignored the correlation between coefficients. Then the non-diagonal element should be zero and we also ignore the weighting between y_i. Thus the uncertainty of y is defined as using $n\times n$ matrix,

$$\mathbf{U}_{\mathbf{y}}^2=E(\Delta\mathbf{y}\Delta\mathbf{y}^{\mathbf{T}})=\begin{bmatrix}\sigma_y^2 & & 0\\ & \ddots & \\ 0 & & \sigma_y^2\end{bmatrix}. \tag{10}$$

The variation of y is calculated from the remained error of the fitting. The expected uncertainty for y may be given by the SSE and degree of freedom,

$$\hat{\sigma}_{y_i}^2=\mathrm{SSE/DoF}\,. \tag{11}$$

The uncertainty of fitted coefficients are determined by calculating (9) using (11) . The diagonal elements of the right hand-side of (9) gives the uncertainty of the fitted coefficients, the non-diagonal elements gives the uncertainty between the coefficients due to the correlation.

2.2. The standard error for non-linear fittings

We described the uncertainty for the function, which is linear combination of the single variable. This sub-section, we applied it to the non-linear function.

We define the initial coefficients as a first guess. The difference between the first estimated value, $y(x_i;\ \mathbf{a}^{(0)})$ and measured output, y_i may be calculated by using Taylor series, as

$$y_i-y_i^{(0)}=\sum_{j=1}^{k}\left(\left.\frac{\partial y(x,\mathbf{a})}{\partial a_j}\right|_{\mathbf{a}=\mathbf{a}^{(0)}}(a_j^{(1)}-a_j^{(0)})\right) \tag{12}$$

Where superscript (0) and (1) denote the initial and second coefficients, respectively. That is rewritten in matrix form using partial derivative for the fitting parameters as,

$$\mathbf{y}-\mathbf{y}^{(0)}=\begin{bmatrix}\frac{\partial y(x_1;\mathbf{a})}{\partial a_1} & \frac{\partial y(x_1;\mathbf{a})}{\partial a_2} & \cdots & \frac{\partial y(x_1;\mathbf{a})}{\partial a_k}\\ \frac{\partial y(x_2;\mathbf{a})}{\partial a_1} & \frac{\partial y(x_2;\mathbf{a})}{\partial a_2} & \cdots & \frac{\partial y(x_2;\mathbf{a})}{\partial a_k}\\ \vdots & \vdots & \ddots & \vdots\\ \frac{\partial y(x_n;\mathbf{a})}{\partial a_1} & \frac{\partial y(x_n;\mathbf{a})}{\partial a_2} & \cdots & \frac{\partial y(x_n;\mathbf{a})}{\partial a_k}\end{bmatrix}(\mathbf{a}^{(1)}-\mathbf{a}^{(0)}) \tag{13}$$

In more simple form, using Jacobian matrix,

$$\Delta\mathbf{y}=\mathbf{X}\Delta\mathbf{a} \tag{14}$$

This is the same formula of (3). That gives the next better fitting coefficients $\mathbf{a}^{(1)}$, as a difference between previous one, is automatically calculated by using (7). This procedure is iteratively applied until satisfying prescribed limit.

The uncertainty of the best-fitted coefficient was determined after the iteration, the Jacobian matrix gives directly a sensitivity of the fitted coefficient for variation of y. Then we can calculate the uncertainty by using (9), the same manner. The diagonal elements of (9) gives the uncertainty of each coefficient.

2.3. Fitting function for contact point determination

For the low-k materials, we explained the details of contact point determination procedure [4], in which we use the power-law with fixing the power as 1.5 to determine the contact point. Here, we also use inverse of the power-law as a model function. This is because the most indentation tester is used the force control type.

$$h=\left(\frac{F}{a}\right)^{1/m}+b \tag{15}$$

Where, h, F are the force applied and displacement of the indenter, respectively. a and b are the fitting parameters, the b denotes the expected correction for the contact point.

The partial derivatives of (15) with respect to a, b and m are used to calculation of Jacobian matrix (3). And the uncertainty of zero-point in the following results denotes the uncertainty of parameter b.

3. RESULTS AND DISCUSSION

3.1. Loading step effect for IIT

In generally, the indentation test is one of industrial test which value is easily varies with the test condition. The loading step is the one of typical example of it. We carried out IIT test on nano-indenter XP with simple loading-unloading test method. Two different loading step is used in the test mode. One is constant force step about 0.8mN, which is simply dividing the total test force by number of steps. The other one is strain rate constant, $\dot{h}/h=0.05\,\mathrm{s}^{-1}$. The difference between the loading steps for the Fused Silica sample is shown in fig. 1. The original (machine detected) zero-point is used in the graph. Here you can get the two different force-depth curves, by using the same machine and the same sample. In addition, you can see not-smooth gradient from 10–80 nm in depth of the curve obtained by the constant step. It may come from the characteristic time with respect to the deformation. For the determination of modulus and hardness, it gives us the completely different information, especially in the determination of zero point. That means, it is not easy to apply the back calculation using power-law or another function for the zero-point estimation to the constant force step data. The following experiment, we use constant strain rate data for the analysis.

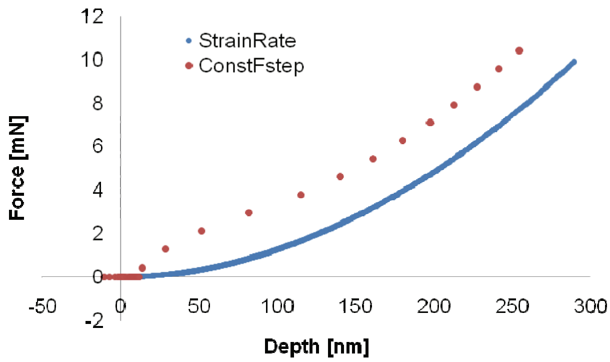

Fig. 1. The beginnings of the force-depth curve obtained by constant force step and strain rate constant.

3.2. Effect of length of the data and fitting functions

In the ISO ad-hoc WG meeting for ISO 14577 on Hannover 2008, the effect of data length was discussed which is used to the estimation of contact point. But, we found that we use the begging of the strain rate constant data analyzed by changing the power, *m* for *m*=1(linear), m=1.5 (elastic contact model), *m*=2, (suitable for perfect pyramid or cones) and *m* as a fitting parameter. Here the original data is obtained by a fused silica and the target force is 100 mN (about 950 nm in depth). We used the data up to 0.05 mN (up to about 20 nm in depth) for the analysis resulting in fig. 2. The result using data up to 5mN (up to 200 nm in depth) is shown in fig. 3. We expected that the tip rounding is effective shallower than 50 nm, which is roughly estimated as the deviating point from the linear line from square-root of the area and contact depth curve when the truncation approximation is used.

The data length was changed to investigate the data length effect of this analysis, where the upper bound of the analysed data is fixed to 0.05 mN and lower bound is changed to decrease the data length. For example, the data from 0.005 to 0.05 mN is used in the fitting when the "lower bound in force" indicating the 0.005 mN. The data from 0.02 to 0.05 mN is used when the "lower bound in force" indicating 0.02 mN, for another example. Therefore, the fitting is done with using much data point, when the left-hand side on the horizontal axis in the fig. 2. The vertical axis indicates the expected zero-point obtained by the fitting.

The fitting result obtained by the data up to 0.05 mN (up to 20 nm in depth) is shown in fig. 2, and used data is shown in fig.4. The estimated zero-point is varies from 4 to 7 nm for the *m*=1, that is at around 3 nm for *m*=1.5 and that is varies from 0 to -1 nm for *m*=2. The zero-point estimated is varies from 1.3 to near 5 nm when the *m* is treated as a fitting parameter. The uncertainty of the fitting for parameter *b* is within the 0.1 nm for the *m*=1, 1.5 and 2. When the *m* is fitted, the uncertainty of zero-point estimation is varies from 0.12 at much data points to 0.65 at less data points. The difference between the functions is about 4.5 nm when lower bound is 0.001 mN, however if we use much small portion very near the contact, the *m*=1 may gives much closer solution for the other one. The results from the *m*=1.5 and fitted m are similar and it may indicating that near *m*=1.5 gives the better solution at the region that tip rounding is significant. The uncertainty of the zero point, calculated from the fitting software, is small comparing the difference between the functions. The *m*=1 fit, for instance, the calculated line is not suitable for the analysis, which is shown in fig. 4.

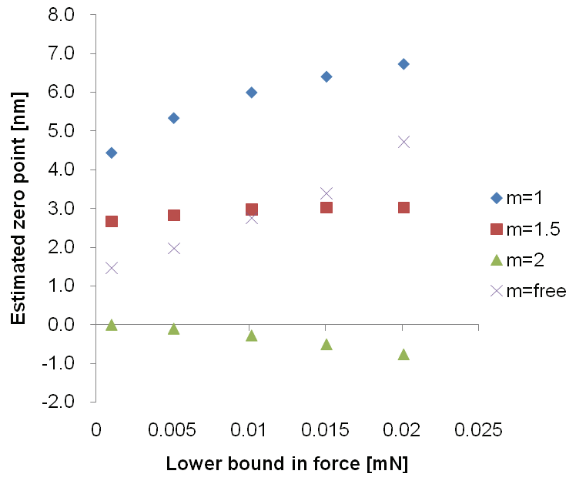

Fig. 2. Estimation of zero-point using power-law fit obtained by the data up to 0.05 mN (up to 20 nm). The tip rounding is significant in this region.

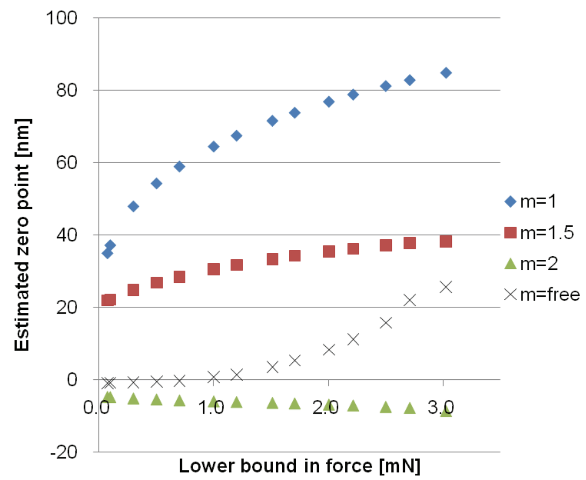

Fig. 3. Estimation of zero-point using power-law fit obtained by the data up to 5 mN (up to 200 nm). The tip rounding is significant in this region.

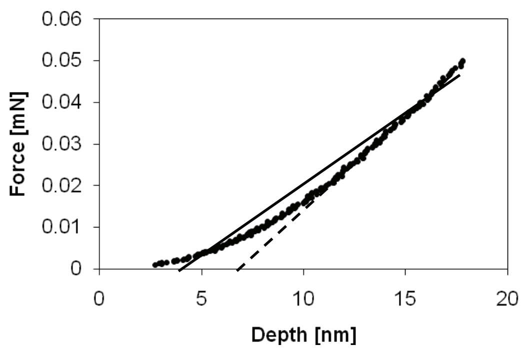

Fig. 4. The force-depth curve used the fitting in Fig. 2. The line and dashed line indicate the fitted results for m=1 and full data points and from 0.02-0.05 mN are used, respectively.

The solid line indicating full data length is used and the dashed-line indicating the shorter data length is used. But the uncertainty of cross section is about 0.1 nm for both case. That only indicates the degree of determination of parameter b, not reflecting the quality of fit. Quality of fit is detected by the (11) or χ^2 value.

The fitting result obtained by the data up to 5 mN (about 200 nm in depth) is shown in fig. 3. The estimated zero-point is varies from 35 to 85 nm for the m=1, that is from 20 to 40 nm for m=1.5 and that is from -5 to -9 nm for m=2, that is from -1 to 26 nm from the m as fitting parameter. Fitted m are also changed from 1.93 to 1.63. The zero-point uncertainties are less than 0.8 nm for m=1, 1.5 and 2. From 0.1 to 4.3 nm for the m is fitted. The difference between functions is about 40 nm at the lower bound in force is 0.08 mN. The depth used in this analysis is from about 20 nm to 200 nm, that region the tip area is regarded as proportional to the square of the contact depth. Thus the result from fitted m and m=2 gives the similar and stable results comparing the other.

The data used these analysis are 0.05% and 0.5% in force and 2% and 20% in depth from the test force 100 mN and 950 nm in maximum depth. It indicating that it is not suitable to fix the m for the analysis using the data that is selected by the some rate from the test force and total depth.

4. SUMMARY

We described the details of the non-linear fit, and standard error calculation procedure which is used to develop the software to calculated the contact point by using power-law. We applied it to the zero-points determination using two different types of the data, considering the tip rounding effect. The summary of this paper are: The data used to the contact point analysis should be determined with considering the tip shape and take care about the force increasing rate. It gives large difference in the fitted results. The uncertainty of the fitted coefficient or standard error not reflects the quality of fit. The IIT values with various zero-point analysis is may causes the difference with in the difference in model that indicated in this report.

ACKNOWLEDGMENTS

Authors thanks to Dr. M. Gripentrog, BAM for his fruitful discussions. One of author thanks to Dr. H. Tanaka for discussion in regression analysis in uncertainty analysis.

REFERENCES

[1] H. Tanaka, "Uncertainty estimation by means of regression analysis", Uncertainty Seminar, March 2008, in Japanese.

[2] T. Nakagawa and Y Koyanagi, "Data analysis using least-square method", The Univ. of Tokyo press, 1982, in Japanese.

[3] W. Press et al, "Numerical Recipes in C 2nd ed.", Cambridge University Press, chapter 15 .

[4] K. Hattori and Y.Seino, "A contact point detection or indentation test of low-k film", *XIX IMEKO World Congress*, pp. 1011, Lisbon, Portugal, 2009, Sept. 2009.