INTERNATIONAL COLLABORATION

The Editorial Board continues the publication of the materials of seminar on “Mathematical, statistical, and computer support to measurement quality.” For the beginning see No. 4 (2003).

IDENTIFICATION AND HANDLING OF DISCREPANT MEASUREMENTS IN KEY COMPARISONS

L. Nielsen

In a key comparison one or several measurement objects are circulated among a number of laboratories, each of which measures the quantities represented by the objects. In order to compare the measurement results obtained by the participating laboratories, the values represented by the circulated objects have to be established. These values, known as the key comparison reference values, and their associated uncertainties can easily be calculated by the method of least squares from the measurement results provided by the participating laboratories. Since this method requires that the measurement results be mutually consistent, a hypothesis that sometimes has to be rejected at a reasonable level of significance, a procedure for identification and handling of discrepant measurements is needed. In this paper such a procedure is suggested. It is demonstrated that although a key comparison reference value is attracted to a discrepant measurement result that has an uncertainty much smaller than the remaining results, the suggested procedure is able to identify this discrepant result. It is also demonstrated that the exclusion of a discrepant measurement result from the calculation of the reference values does not amplify the discrepancy of that result. As the discrepant result is not excluded from the comparison itself, the exclusion of the result in the calculation of the reference values should therefore be uncontroversial.

Key words: identification, handling, key comparisons, discrepant measurement results, reference values.

1. Introduction. The purpose of a key comparison is to test that the measurements performed by national measurement institutes are consistent with the definition of the relevant SI unit taking into account the uncertainties claimed by the institutes. In order to perform such a test, it is necessary to establish a reference value to which the results of the individual participating institute can be compared. One of the methods suggested for establishing reference values is the method of least squares [1], which in the simplest case leads to a reference value equal to the weighted average of the presented results. According to the multilateral recognition agreement (MRA) the established reference value should be a close approximation to the SI value of the object circulated in the key comparison [2]. Since the weighted average and other least squares estimates are based on the a priori assumption that the results are mutually consistent taking into account the claimed uncertainties, discrepant measurements should be identified and excluded from the calculation of the final reference values in case the hypothesis of consistency has to be rejected.
Since the reference value will be attracted by a result to which a small uncertainty has been assigned, the procedure for identification of discrepant measurements should not favor such results. In order to avoid too many discussions on whether or not a discrepant result should be excluded from the calculation of the reference value, the measure of the discrepancy of a single laboratory from the reference value should not change (to the worse) when the result of this laboratory is excluded. The normalized deviations introduced in Section 5 provide a measure of discrepancy that has these two desirable properties. In order to facilitate the understanding of the suggested method, a comprehensive summary of the evaluation of key comparisons by the method of least squares is given.

2. Modeling Measurement Comparisons. Assume that the participants in a measurement comparison provide a total of \( n \) measured values \( y_1, \ldots, y_n \) on one or several circulated measurement objects. For the expectations \( E(y_1), \ldots, E(y_n) \), i.e., the values that would have been measured in the absence of measurement uncertainty and random instability of the measurement objects, a linear model in \( k \geq 1 \) parameters \( a_1, \ldots, a_k \) is assumed:

\[
\begin{align*}
E(y_1) &= x_{11} a_1 + \cdots + x_{1k} a_k \\
E(y_2) &= x_{21} a_1 + \cdots + x_{2k} a_k \\
&\vdots \\
E(y_n) &= x_{n1} a_1 + \cdots + x_{nk} a_k
\end{align*}
\]

or in matrix notation:

\[
E(y) = Xa. \quad (1)
\]

The matrix \( X \) is the design matrix of the intercomparison. The elements of the matrix are known a priori, in principle with zero uncertainty. The parameters \( a \) are unknown a priori and are estimated from the \( n \) measurement results \( y \) provided by the participants and the associated covariance matrix \( \Sigma \):

\[
y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \quad \Sigma = V(y) = \begin{pmatrix} u^2(y_1) & u(y_1, y_2) & \cdots & u(y_1, y_n) \\ u(y_1, y_2) & u^2(y_2) & \cdots & u(y_2, y_n) \\ \vdots & \vdots & \ddots & \vdots \\ u(y_1, y_n) & u(y_2, y_n) & \cdots & u^2(y_n) \end{pmatrix}. \quad (3)
\]

In general there are two sources to the variation observed in an intercomparison: 1) The measurement uncertainties claimed by the participants taking into account the correlations between the provided measurement results, and 2) random instability of the circulated objects.

The first source is described by a covariance matrix \( \Sigma_{\text{meas}} \) with diagonal elements equal to the squares of the claimed standard uncertainties and with off-diagonal elements equal to the recognized covariances between the provided measurement results. Such a covariance would exist, for example, if a laboratory measures on two circulated objects and use the same reference standard in both measurement, or if one participant is using a reference standard calibrated by another participant.

The second source is described by a diagonal covariance matrix \( \Sigma_{\text{ob}} \) with diagonal elements equal to the estimated variance of the value of the measurand due to the random instability of the circulated objects.

The covariance matrix \( \Sigma \), on which the following analysis is based, is given by

\[
\Sigma = V(y) = \Sigma_{\text{meas}} + \Sigma_{\text{ob}}. \quad (4)
\]

3. Estimation of the Unknown Parameters in the Model. The least squares estimates \( \hat{a} \) of the unknown parameters are found by solving the equation

\[
\varphi(\hat{a}) = (y - X\hat{a})^T \Sigma^{-1} (y - X\hat{a}) = \min.
\]

The solution is

\[
\hat{a} = CX^T \Sigma^{-1} y; \quad C = (X^T \Sigma^{-1} X)^{-1}.
\]

514