Statistical Analysis of Interlaboratory Studies with Linear Trends

N. F. Zhang, H.K Liu, N. Sedransk, and W. Strawderman

National Institute of Standards & Technology
Gaithersburg, MD 20899
Outlines

1. Introduction
2. Models and parameter estimation
3. KCRV and its uncertainty
4. Degrees of equivalence
5. An example
6. Conclusions
1. Introduction

• Background

• General practices for Key Comparisons

• Key Comparisons with trends

• Existing methods to deal with trends
Background

• International comparisons of measurements – Key Comparisons.
  – Previously conducted for the mutual information
  – With the signing of MRA in 1999, NMIs and RMOs have committed to establishing the equivalence of measurement standards for selected measurement standards in each field – CIPM Key Comparisons.
Traceability & Comparability

NMI: National Measurement Institute
NMI participating in BIPM/CC Key Comparisons and in RMO Key Comparisons

NMI participating in neither BIPM/CC nor regional Key Comparisons but making bilateral comparisons directly with BIPM or with NMIs
General Practices for Key Comparisons

• The Pilot NMI helped by several participants organizes and draws up the technical protocol for the KC.

• The pilot NMI is responsible for organizing the circulation and transport of the standards (single or multiple).

• The participating NMIs report the measurement results of a KC together with the uncertainties.

• The pilot NMI is responsible for statistical analysis and preparation of the report on a KC.

• Key Comparison Reference Value (KCRV) is an estimate of the measurand. KCRV can be calculated using simple mean, or weighted mean or median or other methods.

• Establish degrees of equivalence.
Key Comparisons with trends

In some KCs, measurements of the transport standards will show a trend or a drift. The figure shows in CCEM-K2, the measurements of resistance standard S/N HR7550 by all participants. Linear regression line is based on the pilot NMI measurements.
Existing methods to deal with trends

The pilot lab data sets were used to determine a trend for the transport standard.

A simple linear regression line was fit to the pilot data using the mean date of each of \( n \) periods as the independent variable in regression.

\[
x_{1,k} = \alpha + \beta t_{1,k} + \varepsilon_k
\]

The difference between the measurement of \( i \)th NMI and the prediction based on a regression at \( t = t_i \) is

\[
D_i = x_i - x_{i,p}
\]

The KCRV in this case is a weighted mean of prediction errors or a key comparison difference. The major disadvantage of this approach is that the results cannot be applied to the trivial case when the trend reduces to zero.
2. Models and parameter estimation

We assume that for each non-pilot lab, a measurement of the traveling standard can be expressed as

\[ X_i = X_{i,A} + X_{i,B}. \]

The two components are independent with each other and with uncertainties \( \sigma_{i,A} \) and \( u_{i,B} \). The uncertainty for the ith lab is given by

\[ u_i^2 = \sigma_{i,A}^2 + u_{i,B}^2. \]

For the pilot lab, each of the K measurements

\[ X_{1k} = X_{1k,A} + X_{1,B} \]

\( X_1 \) is the average and its uncertainty is given by

\[ u_1^2 = \frac{\sigma_{1,A}^2}{K} + u_{1,B}^2. \]
We assume that a simple linear regression model holds for the measurements by the pilot lab

\[ X_{1k} = \alpha_1 + \beta t_{1k} + \epsilon_{1k} \]

For a non-pilot lab, we assume that

\[ X_i = \alpha_i + \beta t_i + \epsilon_i \]

The parameters are estimated based on the pilot lab data and

\[
\hat{\beta} = \frac{\sum_{k=1}^{K} (t_{1k} - t_1)(X_{1k} - X_1)}{\sum_{k=1}^{K} (t_{1k} - t_1)^2}, \quad \hat{\alpha}_1 = X_1 - \hat{\beta}t_1.
\]

For a non-pilot lab, the intercept is estimated by

\[ \hat{\alpha}_i = X_i - \hat{\beta}t_i \]

The uncertainty of \( \hat{\alpha}_1 \) is

\[
u_{\hat{\alpha}_1}^2 = u_{1,B}^2 + \frac{\sigma_{1,A}^2}{K} + \frac{t_1^2 \sigma_{1,A}^2}{\sum_{k=1}^{K} (t_{1k} - t_1)^2} = u_1^2 + \frac{t_1^2 \sigma_{1,A}^2}{\sum_{k=1}^{K} (t_{1k} - t_1)^2}.
\]
For a non-pilot lab,

\[ u_{\hat{\alpha}_i}^2 = u_i^2 + \frac{t_i^2 \sigma_{1,A}^2}{\sum_{i=1}^{K} (t_{1k} - t_1)^2}. \]

The uncertainty for the prediction \( \hat{\alpha}_i + \hat{\beta}t \)

\[ u_{\hat{\alpha}_i + \hat{\beta}t}^2 = u_i^2 + \frac{(t_i - t)^2 \sigma_{1,A}^2}{\sum_{k=1}^{K} (t_{1k} - t_1)^2}. \]

An unbiased estimator of \( \sigma_{1,A}^2 \) is given by

\[ \hat{\sigma}_{1,A}^2 = \frac{\sum_{k=1}^{K} (X_{1k} - \hat{\alpha}_1 - \hat{\beta}t_{1k})^2}{K - 2}. \]
3. Key comparison reference value

One approach to calculating the KCRV at anytime $t$ is to use a weighted mean of $\hat{\alpha}_i + \hat{\beta}t$ over the laboratories,

$$KCRV_t(w) = \sum_{i=1}^{I} w_i (\hat{\alpha}_i + \hat{\beta}t) = \sum_{i=1}^{I} w_i X_i - \hat{\beta} \sum_{i=1}^{I} w_i (t_i - t)$$

The KCRV depends on $t$ and the weights. The corresponding uncertainty of KCRV is given by

$$u^2_{KCRV_t(w)} = \sum_{i=1}^{I} w_i^2 u_i^2 + \frac{\sigma^2_{1,A} \left[ \sum_{i=1}^{I} w_i (t_i - t) \right]^2}{\sum_{k=1}^{K} (t_{1k} - t_1)^2}.$$
For any given set of weights, the minimum value of the uncertainty of \( KCRV_i(w) \) occurs when

\[
t_w = \sum_{i=1}^{I} w_i t_i.
\]

In this case, the KCRV has the same form as in the no-trend case and is given by

\[
KCRV_{i,w}(w) = \sum_{i=1}^{I} w_i X_i.
\]

As in the no-trend case, \( u_{KCRV_{i,w}(w)}^2 \) is minimized when

\[
w_i(1) = \frac{1}{\sum_{k=1}^{I} \frac{1}{u_k^2}}.
\]

The corresponding KCRV is

\[
KCRV_{i,w}(w(1)) = \sum_{i=1}^{I} w_i(1) X_i, \quad (*)
\]

where

\[
t^* = \sum_{i=1}^{I} w_i(1) t_i. \quad (**)
\]
For this KCRV, the uncertainty is the same as in the no-trend case

\[ u_{KCRV_t^*}^2(w(1)) = \frac{1}{\sum_{i=1}^{I} \frac{1}{u_i^2}} \] (***)

assuming that \( u_i \) in the weights are the standard deviations. Otherwise, e.g., \( u_i = S_i \)

the sample sample standard deviation, the weights equal

\[ \frac{1}{S_i^2} \times \frac{1}{\sum_{k=1}^{I} \frac{1}{S_k^2}}. \]

If these weights are used, then (***) does not hold.

**We recommend** \( KCRV_{t^*}(w(1)) \) **with** \( t^* \)

and the weights given by (*) and (**) to be the KCRV.
4. Degrees of equivalence

4.1 Degrees of equivalence of the national measurement standards w.r.t KCRV

It is defined as

\[ D_{i,KCRV(w)} = \hat{\alpha}_i + \hat{\beta} t - KCRV_t(w) \]

\[ = (1 - w_i)\hat{\alpha}_i + \sum_{j \neq i, j=1}^{I} w_j \hat{\alpha}_j. \]

The corresponding uncertainty for the recommended KCRV is

\[ u^2_{D_{i,KCRV(w1)}} = [1 - 2w_i(1)]u_i^2 + \frac{1}{\sum_{k=1}^{K} u_k^2} + \frac{(t_i - t^*)^2 \sigma_{1,A}^2}{\sum_{k=1}^{K} (t_{1k} - t_1)^2}. \]
4.2. Degrees of equivalence of pairs of national measurement standards

It is defined as

\[ D_{i,k} = (\hat{\alpha}_i + \hat{\beta}t) - (\hat{\alpha}_k + \hat{\beta}t) = \hat{\alpha}_i - \hat{\alpha}_k \]

This is independent of t. The corresponding uncertainty is given by

\[ u_{D_{i,k}}^2 = u_i^2 + u_k^2 + \frac{(t_i - t_k)^2 \sigma_{1,A}^2}{K \sum_{k=1}^{K} (t_{1k} - t_1)^2} \]
5. An example

The approach is applied to CCEM-K2, an international comparison of dc resistance at 10MΩ. During the comparison, the transport standard, S/NHR7551 was measured at the pilot NMI, NIST, for 7 separate periods. Each non-pilot NMI (total # 14) reported a mean value (corrected by the nominal value in the unit of μΩ/Ω) and a mean date.

We calculated \( \hat{\alpha}_1 = 6.03 \), \( \hat{\beta} = 1.05 \).

The recommended KCRV at \( t^* = 1998.23 \) with weights \( w(1) \) is 8.03.

The degrees of equivalence were also calculated.
<table>
<thead>
<tr>
<th>Lab</th>
<th>Mean date of Measurement</th>
<th>Measurement (x10^{-6})</th>
<th>Type A uncertainty (x10^{-6})</th>
<th>Type B uncertainty (x10^{-6})</th>
</tr>
</thead>
<tbody>
<tr>
<td>NIST</td>
<td>1996.65</td>
<td>4.6</td>
<td>0.2 *</td>
<td>1.51*</td>
</tr>
<tr>
<td>NRC</td>
<td>1996.80</td>
<td>5</td>
<td>1.88</td>
<td>2.29</td>
</tr>
<tr>
<td>NIST</td>
<td>1996.94</td>
<td>6.7</td>
<td>0.2 *</td>
<td>1.51*</td>
</tr>
<tr>
<td>BNM-LCIE</td>
<td>1997.17</td>
<td>6.97</td>
<td>0.50</td>
<td>0.35</td>
</tr>
<tr>
<td>NPL</td>
<td>1997.35</td>
<td>7.1</td>
<td>0.52</td>
<td>0.61</td>
</tr>
<tr>
<td>PTB</td>
<td>1997.50</td>
<td>7.5</td>
<td>1.0</td>
<td>2.19</td>
</tr>
<tr>
<td>NIST</td>
<td>1997.62</td>
<td>8.1</td>
<td>0.2 *</td>
<td>1.51*</td>
</tr>
<tr>
<td>CSIRO-NML</td>
<td>1997.82</td>
<td>7.3</td>
<td>0.07</td>
<td>2.56</td>
</tr>
<tr>
<td>MSL</td>
<td>1998.03</td>
<td>7.3</td>
<td>0.04</td>
<td>0.59</td>
</tr>
<tr>
<td>CSIR-NML</td>
<td>1998.13</td>
<td>-20</td>
<td>50</td>
<td>13.52</td>
</tr>
<tr>
<td>NIST</td>
<td>1998.33</td>
<td>8.9</td>
<td>0.2 *</td>
<td>1.51*</td>
</tr>
<tr>
<td>SP</td>
<td>1998.49</td>
<td>8.7</td>
<td>0.17</td>
<td>1.79</td>
</tr>
<tr>
<td>OFMET</td>
<td>1998.62</td>
<td>8.9</td>
<td>0.39</td>
<td>0.58</td>
</tr>
<tr>
<td>IEN</td>
<td>1998.74</td>
<td>9.4</td>
<td>0.79</td>
<td>2.53</td>
</tr>
<tr>
<td>NMi-VSL</td>
<td>1998.98</td>
<td>9.1</td>
<td>0.80</td>
<td>3.04</td>
</tr>
<tr>
<td>NIST</td>
<td>1999.15</td>
<td>7.8</td>
<td>0.2 *</td>
<td>1.51*</td>
</tr>
<tr>
<td>KRIST</td>
<td>1999.39</td>
<td>7.1</td>
<td>0.30</td>
<td>3</td>
</tr>
<tr>
<td>NIST</td>
<td>1999.60</td>
<td>8.5</td>
<td>0.2 *</td>
<td>1.51*</td>
</tr>
<tr>
<td>NIM</td>
<td>1999.87</td>
<td>10.2</td>
<td>0.10</td>
<td>0.83</td>
</tr>
<tr>
<td>VNIIM</td>
<td>2000.03</td>
<td>10</td>
<td>0.25</td>
<td>1.03</td>
</tr>
<tr>
<td>NIST</td>
<td>2000.20</td>
<td>10</td>
<td>0.2 *</td>
<td>1.51*</td>
</tr>
</tbody>
</table>
Click Here
6. Conclusions

We propose a new statistical analysis for Key Comparisons with linear trends. The calculation of the KCRV is consistent with the case in which there is no trend. The corresponding uncertainties for the KCRV and the degrees of equivalence are also provided.