Data Combination from Multiple Sources Under Measurement Error

Hugo Gasca-Aragon
University of Massachusetts Amherst, hugo_gasca_aragon@hotmail.com

Follow this and additional works at: https://scholarworks.umass.edu/open_access_dissertations
Part of the Mathematics Commons, and the Statistics and Probability Commons

Recommended Citation

This Open Access Dissertation is brought to you for free and open access by ScholarWorks@UMass Amherst. It has been accepted for inclusion in Open Access Dissertations by an authorized administrator of ScholarWorks@UMass Amherst. For more information, please contact scholarworks@library.umass.edu.
DATA COMBINATION FROM MULTIPLE SOURCES
UNDER MEASUREMENT ERROR

A Dissertation Presented
by
HUGO GASCA-ARAGON

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

February 2013

Department of Mathematics and Statistics
DATA COMBINATION FROM MULTIPLE SOURCES UNDER MEASUREMENT ERROR

A Dissertation Presented

by

HUGO GASCA-ARAGON

Approved as to style and content by:

__________________________
John P. Buonaccorsi, Chair

__________________________
David L. Duewer, Member

__________________________
John Staudenmayer, Member

__________________________
Anna Liu, Member

__________________________
Edward Stanek, Member

__________________________
Michael Lavine, Department Chair
Department of Mathematics and Statistics
To my lovely family, Mariana, Iris & Emily
To my outstanding parents, Esther & Sóstenes
To my brother and sisters, Raúl, Judith, Alma Alicia & Jacqueline
ACKNOWLEDGMENTS

Is has been a great pleasure and honor to achieve this collaborative research through two excellent Institutes: University of Massachusetts Amherst (UMASS) and National Institute of Standards and Technology (NIST). Their respective, advisors, staff and friends were a great team that provided me with passionate support during this commitment. Thanks to CONACYT, MEXICO, for their doctoral grant support.

Special thanks to my committee members. Thanks to the unconditional and delighted mentoring provided by Dr. John P. Buonaccorsi, my UMASS advisor and committee chair. Thanks also to Dr. David L. Duewer, my NIST committee advisor and UMASS external committee member, for his patient and unsurpassed guidance in this project. My gratitude also to Dr. John Standermayer, Dr. Anna Liu, UMASS committee members, for your guidance, who always answered my questions and Dr. Edward Stanek, UMASS external committee member.

Thanks to NIST for providing the facilities, funding and staff for this project, Dr. Stephen Wise, Dr. Willie May, who made this commitment possible. Dr. Michele Schantz, Dr. Katrice Lippa, NIST Analytical Chemistry Division, for their technical support, for making detailed data accessible for this work and for answering my questions. Dr. Andrew Rukhin, Dr. Blaza Toman and Dr. William Guthrie, NIST Statistical Engineering Division, for their statistical support and conversations along this work. My acknowledge to NIST Analytical Chemistry Division, for their unconditional support and their friendship.

My acknowledge and thanks to my friends and project partners at UMASS: David Resendes and Meng Shiou Shieh.
ABSTRACT

DATA COMBINATION FROM MULTIPLE SOURCES UNDER MEASUREMENT ERROR

FEBRUARY 2013

HUGO GASCA-ARAGON

B.Sc., INSTITUTO POLITÉCNICO NACIONAL
Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor John P. Buonaccorsi

Regulatory Agencies are responsible for monitoring the performance of particular measurement communities. In order to achieve their objectives, they sponsor Intercomparison exercises between the members of these communities.

The Intercomparison Exercise Program for Organic Contaminants in the Marine Environment is an ongoing NIST/NOAA program. It was started in 1986 and there have been 19 studies to date. Using this data as a motivation we review the theory and practices applied to its analysis.

It is a common practice to apply some kind of filter to the comparison study data. These filters go from outliers detection and exclusion to exclusion of the entire data from a participant when its measurements are very “different”. When the measurements are not so “different” the usual assumption is that the laboratories are unbiased then the simple mean, the weighted mean or the one way random effects model are applied to obtain estimates of the true value.
Instead we explore methods to analyze these data under weaker assumptions and apply them to some of the available data. More specifically we explore estimation of models assessing the laboratories performance and way to use those fitted models in estimating a consensus value for new study material. This is done in various ways starting with models that allow a separate bias for each lab with each compound at each point in time and then considering generalizations of that. This is done first by exploiting models where, for a particular compound, the bias may be shared over labs or over time and then by modeling systematic biases (which depend on the concentration) by combining data from different labs. As seen in the analyses, the latter models may be more realistic.

Due to uncertainty in the certified reference material analyzing systematic biases leads to a measurement error in linear regression problem. This work has two differences from the standard work in this area. First, it allows heterogeneity in the material being delivered to the lab, whether it be control or study material. Secondly, we make use of Fieller’s method for estimation which has not been used in the context before, although others have suggested it. One challenge in using Fieller’s method is that explicit expressions for the variance and covariance of the sample variance and covariance of independent but non-identically distributed random variables are needed. These are developed.

Simulations are used to compare the performance of moment/Wald, Fieller and bootstrap methods for getting confidence intervals for the slope in the measurement model. These suggest that the Fieller’s method performs better than the bootstrap technique. We also explore four estimators for the variance of the error in the equation in this context and determine that the estimator based on the modified squared residuals outperforms the others.

Homogeneity is a desirable property in control and study samples. Special experiments with nested designs must be conducted for homogeneity analysis and as-
essment purposes. However, simulation shows that heterogeneity has low impact on the performance of the studied estimators. This work shows that a biased but consistent estimator for the heterogeneity variance can be obtained from the current experimental design.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>ACKNOWLEDGMENTS</th>
<th>vii</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>vi</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>xiii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>xvii</td>
</tr>
</tbody>
</table>

## CHAPTER

### 1. INTRODUCTION

1. Context .................................................. 1
2. Specific Study Data ........................................... 3
   1.2.1 About the Treatments .................................. 3
   1.2.2 About the Participating Laboratories ................. 4
   1.2.3 About the Samples .................................... 5
   1.2.4 About the Data Composition ............................ 6
3. Motivation .................................................. 6
4. Organization Outline ......................................... 8

### 2. MODELS AND NOTATION

2.1 Laboratory Measurement Error Model (LMEM) .................. 11
   2.1.1 Special Cases ......................................... 14
   2.1.2 Bias of the Method .................................... 15
2.2 The Study Data .............................................. 16
   2.2.1 Pairing of Data ....................................... 16
2.3 The Certified and Reference Values ............................ 17
3. THE UNBIASED AND CONSTANT BIAS MODEL .............. 22

3.1 One laboratory, one compound at one point in time ................. 22
   3.1.1 The Basic Model ........................................... 22
   3.1.2 Working with the Control Data Only ......................... 23
      3.1.2.1 Parameter Estimation ................................. 23
      3.1.2.2 Inferences about the Parameters ....................... 26
      3.1.2.3 Performance Evaluation ............................... 26
      3.1.2.4 Examples ............................................ 29
   3.1.3 Exporting the Model to the Study Data ....................... 35
      3.1.3.1 Parameter Estimation ................................. 35
      3.1.3.2 Examples ............................................ 39

3.2 Multiple laboratories, one compound at one point in time ............ 43
   3.2.1 The Basic Model ........................................... 43
   3.2.2 Working with the Study Data Only .......................... 43
      3.2.2.1 Parameter Estimation ................................. 44
   3.2.3 Working with the Control and Study Data .................... 48
   3.2.4 Performance Evaluation ................................... 52
   3.2.5 Examples ................................................. 54

3.3 One laboratory, multiple compounds at one point in time ............. 62
   3.3.1 The Basic Model ........................................... 62
   3.3.2 Working with the Control Data Only ......................... 63
      3.3.2.1 Parameter Estimation ................................. 63
      3.3.2.2 Inferences about the Parameters ....................... 67
      3.3.2.3 Performance Evaluation ............................... 69
      3.3.2.4 Examples ............................................ 69
   3.3.3 Exporting the Model to the Study Data ....................... 74
      3.3.3.1 Parameter Estimation ................................. 74
      3.3.3.2 Examples ............................................ 76

3.4 One laboratory, one compound at several points in time .............. 79
<table>
<thead>
<tr>
<th>3.4.1</th>
<th>The Basic Model</th>
<th>79</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.4.2</td>
<td>Working with the Control Data Only</td>
<td>81</td>
</tr>
<tr>
<td>3.4.2.1</td>
<td>Estimation and inferences</td>
<td>82</td>
</tr>
<tr>
<td>3.4.2.2</td>
<td>Examples</td>
<td>87</td>
</tr>
<tr>
<td>3.4.3</td>
<td>Exporting the Model to the Study Data</td>
<td>92</td>
</tr>
<tr>
<td>3.4.3.1</td>
<td>Parameter Estimation</td>
<td>92</td>
</tr>
<tr>
<td>3.4.3.2</td>
<td>Examples</td>
<td>94</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>4.1</th>
<th>One laboratory, one compound at one point in time</th>
<th>98</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1.1</td>
<td>The Basic Model</td>
<td>98</td>
</tr>
<tr>
<td>4.1.2</td>
<td>Working with the Control Data Only</td>
<td>99</td>
</tr>
<tr>
<td>4.1.2.1</td>
<td>Parameter Estimation</td>
<td>99</td>
</tr>
<tr>
<td>4.1.2.2</td>
<td>Inferences about the Parameters</td>
<td>103</td>
</tr>
<tr>
<td>4.1.2.3</td>
<td>Performance Evaluation</td>
<td>104</td>
</tr>
<tr>
<td>4.1.2.4</td>
<td>Examples</td>
<td>104</td>
</tr>
<tr>
<td>4.1.3</td>
<td>Exporting the Model to the Study Data</td>
<td>105</td>
</tr>
<tr>
<td>4.1.3.1</td>
<td>Parameter Estimation</td>
<td>105</td>
</tr>
<tr>
<td>4.1.3.2</td>
<td>Example</td>
<td>113</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>4.2</th>
<th>One lab, multiple compounds at one point in time</th>
<th>121</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.2.1</td>
<td>Working with the Control Data Only</td>
<td>121</td>
</tr>
<tr>
<td>4.2.1.1</td>
<td>The Model</td>
<td>121</td>
</tr>
<tr>
<td>4.2.1.2</td>
<td>Coefficient Estimation</td>
<td>123</td>
</tr>
<tr>
<td>4.2.1.3</td>
<td>Interval Estimators and Estimators for the Coefficient Covariance</td>
<td>125</td>
</tr>
<tr>
<td>4.2.1.4</td>
<td>Testing for unbiasedness of the method</td>
<td>127</td>
</tr>
<tr>
<td>4.2.1.5</td>
<td>Estimation of the Variance of the Error in Equation</td>
<td>127</td>
</tr>
<tr>
<td>4.2.1.6</td>
<td>Modified Estimators</td>
<td>130</td>
</tr>
<tr>
<td>4.2.1.7</td>
<td>Bootstrapping for inference</td>
<td>131</td>
</tr>
<tr>
<td>4.2.1.8</td>
<td>Performance Evaluation</td>
<td>135</td>
</tr>
<tr>
<td>4.2.1.9</td>
<td>Residual Analysis</td>
<td>136</td>
</tr>
<tr>
<td>4.2.1.10</td>
<td>Fieller’s confidence intervals</td>
<td>138</td>
</tr>
<tr>
<td>4.2.1.11</td>
<td>Simulation Study</td>
<td>142</td>
</tr>
<tr>
<td>4.2.1.12</td>
<td>Weighted estimators</td>
<td>158</td>
</tr>
</tbody>
</table>
4.2.1.13 Example ........................................ 162
4.2.1.14 Small Range Example .......................... 179

4.2.2 Exporting the Model to the Study Data ............. 181
  4.2.2.1 Parameter Estimation ............................ 181
  4.2.2.2 Inferences for $d_a$ ............................. 183
  4.2.2.3 Inferences for $\sigma^2_V$ ........................ 186
  4.2.2.4 Inferences Under Correlated Errors ............ 187
  4.2.2.5 Example ........................................ 192

5. CHANGING MEASUREMENT VARIANCE .................... 198
  5.1 Modeling the variance as a function of concentration..... 198
  5.2 Variance proportional to the squared concentration .... 198
     5.2.1 Parameter Estimation ............................ 199
     5.2.2 Performance Evaluation .......................... 200
     5.2.3 Exporting the Variance Model ..................... 202
        5.2.3.1 Parameter Estimation ....................... 202
  5.3 Example ............................................ 205

6. CONCLUSIONS AND FUTURE WORK ....................... 209
  6.1 Conclusions ......................................... 209
  6.2 Future Work ........................................ 211

APPENDICES

A. THE WELCH-SATTERTHWAITE APPROXIMATION ............ 213
B. LEVENE TEST UNDER HETEROGENEITY ..................... 217
C. THE VARIANCE AND COVARIANCE OF THE SAMPLE
   VARIANCE AND SAMPLE COVARIANCE OF
   INDEPENDENT RANDOM VARIABLES ..................... 218

BIBLIOGRAPHY .............................................. 226
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Example of detailed reported data and the related CRM reference values.</td>
<td>10</td>
</tr>
<tr>
<td>2.1 Data structure for the control material.</td>
<td>12</td>
</tr>
<tr>
<td>2.2 Data structure for the study material.</td>
<td>16</td>
</tr>
<tr>
<td>2.3 Data Notation.</td>
<td>20</td>
</tr>
<tr>
<td>3.1 Example of summary data and reference values.</td>
<td>30</td>
</tr>
<tr>
<td>3.2 Example of point and interval estimation of a method constant bias.</td>
<td>31</td>
</tr>
<tr>
<td>3.3 Example of point and interval estimates of the method within uncertainty.</td>
<td>32</td>
</tr>
<tr>
<td>3.4 Evaluation scores under constant method bias and constant method variance.</td>
<td>33</td>
</tr>
<tr>
<td>3.5 Example of study values and variances corrected for bias using control data when errors are independent.</td>
<td>41</td>
</tr>
<tr>
<td>3.6 Example of study values corrected for bias using control data when errors are correlated.</td>
<td>42</td>
</tr>
<tr>
<td>3.7 Example of MLE under the one way random effects model using the raw study data.</td>
<td>57</td>
</tr>
<tr>
<td>3.8 Example of simple mean and weighted mean of study values corrected for constant bias, under independent errors.</td>
<td>58</td>
</tr>
<tr>
<td>3.9 Example of simple mean and weighted mean of study values corrected for constant bias, under correlated errors.</td>
<td>59</td>
</tr>
</tbody>
</table>
3.10 Parameter pooled estimates assuming constant bias and constant within-method variance. ................................. 70

3.11 Example of study values corrected for bias using control data assuming independent errors and correlated errors, assuming a constant bias across compounds. ................................. 77

3.12 Example of adjusted study values when errors are correlated, constant bias and constant correlation across compounds. .......... 78

3.13 Example of summary data reported over time. ....................... 88

3.14 Method bias and within variance estimated separately over time using the control data. ................................. 89

3.15 Example of adjusted study values when errors are independent and constant bias at each point in time. ................................. 95

3.16 Example of study values corrected for bias with correlated errors, assuming constant correlation over time and constant bias at each point in time. ................................. 96

4.1 Example of point and interval estimates using model with proportional bias. ................................................. 106

4.2 Example of study values and variances corrected for proportional bias using control data when errors are independent. ................. 116

4.3 Example of study values and variances corrected for proportional bias using control data when errors are correlated. ................. 117

4.4 Example of study values corrected for proportional bias using control data, contrasting when errors are assumed independent or correlated. ................................. 118

4.5 Example of estimated bias of the study values corrected for proportional bias using control data, contrasting when errors are assumed independent or correlated. ................................. 119

4.6 Example of estimated within material variance for study data, under proportional bias. ................................. 120

4.7 Summary of parameter values used during the simulation study. ...... 143
4.8 Summary for the mean and standard deviation of the estimated coverage of the different CIs for $\beta_1$. ......................................................... 153

4.9 Summary for the mean and standard deviation of the estimated CI length, while varying $K =$ the number of replicates and $\sigma^2_U =$ variance due to heterogeneity of the material. The Fieller’s CI estimates are based on the finite length CIs only when possible. ................................................................. 154

4.10 Summary for the mean and standard deviation of the proportion of finite and infinite length Fieller’s CI for $\beta_1$......................................................... 156

4.11 Results for the estimated variance of the error in the equation......... 159

4.12 Results for the ratio of the estimated standard deviation (CV) of the error in the equation over the true standard deviation (CV), for $A=10, K=3, \beta_0 = 20, \beta_1 = 0.9$. ......................................................... 160

4.13 Example of summary data and reference values for laboratory 1. ...... 163

4.14 Example of summary data and reference values for laboratory 2. ...... 164

4.15 Coefficients estimate assuming linear bias and constant variance. ...... 165

4.16 Parameter confidence intervals assuming linear bias and constant variance. ................................................................. 172

4.17 Test for unbiasedness of the method assuming linear bias model and constant variance. ................................................................. 172

4.18 Estimated variance of the error in equation assuming linear bias. ...... 172

4.19 Parameter confidence interval estimates using bootstrap and assuming linear bias and constant variance for the error in the equation. ................................................................. 173

4.20 Parameter bias and standard error estimates using bootstrap and assuming linear bias and constant variance for the error in the equation. ................................................................. 173

4.21 Test for unbiasedness of the method assuming linear bias model and constant variance. ................................................................. 173

4.22 Bootstrapped p-values for testing unbiasedness of the method assuming linear bias model. ................................................................. 173
4.23 Parameter weighted estimates assuming linear bias and changing variance of the error in the equation. .......................... 178

4.24 Parameter weighted confidence intervals assuming linear bias and changing variance of the error in the equation. ................. 178

4.25 Parameter pooled estimates using the means data only, assuming linear bias and constant variance of the error in the equation. ....... 181

4.26 Parameter pooled weighted estimates, assuming linear bias. .......... 182

4.27 Coefficients estimate assuming linear bias and constant variance for a subset with small range of the predictor. ......................... 182

4.28 Example of study values and variances corrected for linear bias using control data when errors are independent. ......................... 195

4.29 Example of study values and variances corrected for linear bias using control data when errors are correlated. ......................... 196

4.30 Example of study within material variances for proportional bias. ....... 197

5.1 Variance modeling separate estimates for laboratory 1. ................. 205

5.2 Variance modeling assuming constant coefficient for laboratory 1. ....... 206

5.3 Example of point and interval estimates of study sample within-material variance using the separate variance models from the control sample. ............................................... 207

5.4 Example of point and interval estimates of study sample within-material variance using the pooled estimate for the variance model from the control sample. ................................. 208

5.5 Variance modeling assuming homogeneous materials and constant coefficient for laboratory 1, using both control and study data simultaneously. ................................................. 208
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Example of an evaluation scores.</td>
</tr>
<tr>
<td>3.2</td>
<td>Estimated biases for the different laboratories for each compound, assuming constant bias across samples.</td>
</tr>
<tr>
<td>3.3</td>
<td>Estimated CI lengths for each compound, assuming constant bias across samples.</td>
</tr>
<tr>
<td>3.4</td>
<td>Example of performance evaluation using the adjusted study data.</td>
</tr>
<tr>
<td>3.5</td>
<td>Point and interval estimation of the method constant bias.</td>
</tr>
<tr>
<td>3.6</td>
<td>Point and interval estimation of the method constant variance.</td>
</tr>
<tr>
<td>3.7</td>
<td>Point and interval estimates of the method bias over time.</td>
</tr>
<tr>
<td>3.8</td>
<td>Point and interval estimates of the within-method variance over time.</td>
</tr>
<tr>
<td>4.1</td>
<td>Estimated coverage of the Wald’s CI for $\beta_1$.</td>
</tr>
<tr>
<td>4.2</td>
<td>Estimated coverage of the Bootstrap CI for $\beta_1$.</td>
</tr>
<tr>
<td>4.3</td>
<td>Estimated coverage of the Fieller’s CI for $\beta_1$.</td>
</tr>
<tr>
<td>4.4</td>
<td>Estimated coverage of the Wald’s CI for $\beta_1$ using replicates.</td>
</tr>
<tr>
<td>4.5</td>
<td>Estimated coverage of the Bootstrap CI for $\beta_1$ using replicates.</td>
</tr>
<tr>
<td>4.6</td>
<td>Estimated coverage of the Fieller’s CI for $\beta_1$ using replicates.</td>
</tr>
<tr>
<td>4.7</td>
<td>Observed data for laboratory 1.</td>
</tr>
<tr>
<td>4.8</td>
<td>Observed data for laboratory 2.</td>
</tr>
</tbody>
</table>
4.9 Confidence regions for the linear coefficients for laboratory 1. 169
4.10 Confidence regions for the linear coefficients for laboratory 2. 170
4.11 Confidence regions for the linear coefficients for laboratory 6. 171
4.12 Example of residual versus adjusted true value estimate. 174
4.13 Example of residual versus adjusted true value estimate. 175
4.14 Example of the estimated variance of the error in the equation versus adjusted true value estimate. 176
4.15 Example of the modified squared residual versus adjusted true value estimate. 177
4.16 Example of performance evaluation under the linear bias model assuming constant variance of error in the equation for laboratory 1. 179
4.17 Example of performance evaluation under the linear bias model assuming constant variance of error in the equation for laboratory 2. 180
CHAPTER 1
INTRODUCTION

1.1 Context

In the globalizing world, there is an evolving need to assure the quantities involved in all trading are what they are meant to be. In order to face this challenge, governments around the world have designed a complex structure supported on three organizational pillars:

- The National Metrology Institutes (NMIs) which provide main technical and scientific research and development and higher-metrological order reference materials and procedures;
- The regulatory agencies, which write standards and protocols on how to make official measurements. They may also grant or revoke the rights of the producers or service providers to continue operating in the market; and
- Independent accreditation societies (private and public) which audit their associates compliance with the standards and protocols.

In order to achieve their objectives, NMIs and accreditation societies have developed multiple source measurement exercises/experiments. These exercises vary widely in structure, rules and goals.

- Key Comparisons are exercises among NMIs where the participants publicly demonstrate their measurement competencies with full attribution. Key Comparisons are typically focused on testing the highest measurement capability at
one point in time for specific measurements for specific materials. In late 1999 the Mutual Recognition Arrangement of the Comit International des Poids et Mesures (CIPM MRA) for national measurement standards and for calibration and measurement certificates issued by NMIs was signed. The number and diversity of non-routine multiple laboratory studies involving NMIs has dramatically increased since then.

• Exercises sponsored by a regulatory agency and sometimes run by a NMI are called inter-comparison exercises where the participants have the opportunity to share their experiences and learn from the others and their own participation. While intended to monitor the performance of particular measurement communities, these studies are generally open to other interested participants. The results of these studies are often publicly available, generally in somewhat coded form.

• Accreditation society exercises sponsored by their own members are called proficiency tests (PTs) where the participants have the opportunity to demonstrate their capabilities. These are run under strict communication rules due to commercial and economic implications. Such rules include confidential communication and results shared only between each participant and the coordinator. Hence information for group analysis or trend analysis is generally only available to the coordinator. Regulatory agencies often require participation in PTs, which explains the increasing trend of these PTs around the world.

With the increasing number and variety of interlaboratory comparisons, there is a need for consistent methods of data evaluation and presentation. In order to improve these practices it is necessary to analyze the long-term behavior of these exercises. It is desirable to have automated tools to support these activities both numerically and visually.
Several questions arise about the benefits of participating in such an exercise. This study addresses some of the questions about the qualitative and quantitative benefits of participation. It develops methods for data analysis, measurement performance analysis. A set of automated tools that could support these exercises in a consistent and interoperative way was developed as a result but is out of the scope of the present work to detail these tools.

1.2 Specific Study Data

The Intercomparison Exercise Program for Organic Contaminants in the Marine Environment (IEPOCME), initially sponsored by the National Oceanic and Atmospheric Administration (NOAA), is an ongoing program coordinated by the National Institute of Standards and Technology (NIST). It was started in 1986 and there have been 19 studies to date.

In these exercises all participants are invited to follow their regular procedures knowing that these exercises and their results are intended for sharing experience and learning from the others and their own participation. Thus biases are not due to financial or managerial pressures but mainly to technical differences including methods, training, instruments, materials and measurement procedures.

The design of the IEPOCME is fortunately similar to that of many proficiency tests, thus methods developed for these data are expected to be more generally useful. This specific data will be used throughout the document to illustrate the use of models developed for their analysis. The results of the analysis are compared against those from the current practices.

1.2.1 About the Treatments

There are four main families of compounds to be analyzed: polycyclic aromatic hydrocarbons (PAH), chlorinated pesticides (PES), polychlorinated biphenyl congeners
(PCB), and brominated diphenyl ethers (BDE). The compounds of different families are extracted using different procedures so the compound family is a main factor to be considered in the data analysis.

There are three different types of samples: sediment, mussel tissue, and fish tissue. The sample type refers to an environmental material in which the compounds to be analyzed are embedded, so each sample type requires a different handling and measurement process. The sample type imposes different conditions on the measurement, so the sample type is a main factor to be considered in the data analysis.

The treatment definition is the combination of these two main factors, but not all combinations are available. In particular, the PAH family of compounds is not measured for the fish tissue type sample, since the PAH compounds are metabolized in fish.

1.2.2 About the Participating Laboratories

From 1986 to 2006 there were 79 different participants. The number of participants varies on each exercise. Mainly three groups are recognized:

- The core group: A few laboratories who have been participating from the beginning of the program,
- The recurring group: The laboratories who have been participating with interruptions, and
- The occasional group: The laboratories that have participated less than three times per specific treatment (combination of type of sample and family of compounds).

Sequences of participations can be built for each laboratory while measuring the same family of compounds on the same type of sample. The distribution of these sequences is:
<table>
<thead>
<tr>
<th>Group</th>
<th>Compound Family</th>
<th>Sediments</th>
<th>Mussel Tissue</th>
<th>Fish Tissue</th>
<th>Total Sequences</th>
</tr>
</thead>
<tbody>
<tr>
<td>core</td>
<td>PAH</td>
<td>8</td>
<td>6</td>
<td>NA</td>
<td>42</td>
</tr>
<tr>
<td></td>
<td>PES</td>
<td>10</td>
<td>6</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PCB</td>
<td>7</td>
<td>5</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>BDE</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>recurring</td>
<td>PAH</td>
<td>23</td>
<td>15</td>
<td>NA</td>
<td>142</td>
</tr>
<tr>
<td></td>
<td>PES</td>
<td>20</td>
<td>17</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PCB</td>
<td>27</td>
<td>17</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td></td>
<td>BDE</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>occasional</td>
<td>PAH</td>
<td>27</td>
<td>29</td>
<td>NA</td>
<td>258</td>
</tr>
<tr>
<td></td>
<td>PES</td>
<td>24</td>
<td>27</td>
<td>38</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PCB</td>
<td>23</td>
<td>34</td>
<td>43</td>
<td></td>
</tr>
<tr>
<td></td>
<td>BDE</td>
<td>6</td>
<td>7</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

The BDE family of compounds and the fish tissue samples were included more recently, that is the reason there are fewer sequences recorded. The 258 sequences from the participants in the occasional group can only be evaluated with point estimators. The 142 sequences from the participants in the recurring group can in addition be evaluated with interval estimators. Finally, the 42 sequences from the participants in the core group may be evaluated for trend estimation with some risk involved since the largest sequences are just 16 events long.

1.2.3 About the Samples

The intercomparison events have evolved to include an unknown study sample and a control sample, where subsamples of both are distributed to all participants. The control samples used in this program are Certified Reference Materials (CRM) produced by the NIST or some other NMI (CRMs produced by NIST are named “Standard Reference Material ®” and are called “SRMs”). CRMs are prepared using strict quality control procedures and evaluations of material homogeneity and stability as well as analyte quantity. Two or more independent analytical techniques are typically used to measure the concentration of the chemical compounds. When statistical agreement on the results from the several techniques is reached, a consensus
value and its related uncertainty are assigned as a certified value. Most known or suspected sources of bias are investigated or accounted for by the NMI. The unknown samples are prepared and bottled with the same procedures as the CRM, but no certified values are available for its composition. In fact, one of the goals is to estimate the concentration of the selected chemical compounds in the unknown sample.

1.2.4 About the Data Composition

The protocol of the IEPOCME exercises requests three or more quantitative readings for each selected chemical compound from each participant at each point in time. Due to the complexity of the chemical analysis process (resources, environmental and human constraints) and the low concentration of the chemical compounds (sometimes at ultra-trace levels), it is not always feasible to successfully obtain the intended number of replicates as designed.

The reported measurements vary from qualitative values such as "Not Analyzed" (NA) or "Below Detection Limit" (DL), to censored values such as "< 1", to mixtures of qualitative and quantitative values, to finally pure quantitative replicates. This results in an unbalanced and incomplete design of the experiment. Table 1.1 shows an example of the data reported on an intercomparison exercise by one of the participants.

1.3 Motivation

This thesis investigates bias models and variance models that may occur in the combination of data from multiple sources. Key comparisons, agency-sponsored intercomparisons, and PTs are just some of the exercise examples where method bias and method variance analysis must be considered. Assuming that all measurement methods are unbiased with low variability can lead to imprecise, inadequate and misleading estimates. Estimators are imprecise when the true variance terms are large.
They are inadequate when the bias terms are large and not included in the model or simply assumed zero without testing for their presence. They are misleading because they might predict a certain consensus/score to be satisfactory when in fact it is unsatisfactory. Therefore, to fully grasp and anticipate the behavior of a multiple sources study, one needs to consider the influences of both kinds of error inherent in the system.

Unless the bias and variance terms are fully understood and accounted for in analyzing the experiments, unpredictable results may occur. Consider the following scenarios:

- The most notorious case is when a participant reports data with large variability and low relative bias. In general this scenario will produce imprecise estimates (i.e., estimates with undesirable large uncertainty). Depending on the evaluation scheme it can be evaluated as satisfactory if the large variability is ignored or used to normalize the values as part of a larger divisor uncertainty, or unsatisfactory if the large variability is evaluated against the variability of the consensus value (i.e., the coverage region of the reported value is simply too wide).

- A second case is a participant reporting data with moderate variability and large relative bias. In general this scenario will produce inadequate consensus value if the bias terms are not taken into account. Also depending on the evaluation scheme it can be evaluated as satisfactory if the moderate variability is considered, such that the coverage region includes the true value, or it can be evaluated as unsatisfactory if the moderate variability is ignored and the consensus coverage region does not include the large biased estimate of the mean.
• A third case is a participant reporting data with unrealistically low variance. This data will be given too much weight and this generates a push-pull effect on all other participants. In general this will produce misleading estimates due to the bias terms of the participants with lower variances.

These generic incidents demonstrate the importance of considering bias and variance terms, whether occurring within the method itself or between methods. The purpose of this work is to investigate such terms in combining data from multiple sources.

1.4 Organization Outline

The control materials used in the IEPOCME have a certified value along with a measure of uncertainty and there is no strong evidence of among sample heterogeneity.

Using the IEPOCME data as motivation, in this thesis we explore fitting measurement models for the laboratories.

• A. This is done at individual points in time for each compound and laboratory, allowing a constant bias for a given laboratory and compound, but allowing this bias and the variance to change over compounds and laboratories. Also an among sample heterogeneity is allowed.

• B. These models are used to make inferences on study samples. First, for a single laboratory and compound at a point in time and then by combining information over many laboratories at a point in time.

• C. Then the work is expanded to allow for the possibility that the bias may be related to the underlying concentration for similar compounds. This is done using a linear bias model and by combining data over different compounds for each laboratory and point in time. In principle, this could also be used for
a single laboratory and compound with different values of CRM over time. However, for this data there was little or no variation in the CRMs’ values over time.

- D. Throughout, we allow the replicate measures from the control and study material to be either independent or for the replicates to be paired and correlated.

Chapter 2 presents the basic error models, and the data and notation definition.

Chapter 3 presents the unbiased and the constant bias models. First using the control data only and secondly exporting the models to the study data to estimate the true values of the study material. This chapter also presents ways to evaluate the performance of the participants. In addition, it is shown how to combine the individual results from the laboratories into a consensus value for the study sample. When extending the analysis over time, the use of the same reference values introduces some correlation that needs to be accounted for.

In Chapter 4, it is assumed the bias model is a function of concentration and a bias model is obtained by combining data over compounds for a particular laboratory. Since the values of the predictor are known with uncertainty, it turns into a measurement error in linear regression problem. Parameters are estimated using method of moments and inferences carried out using the delta method and a two stage parametric bootstrap. We also introduce the use of Fieller’s method since the problem can be stated as a ratio of random variables. This approach has not been explored for this problem. A simulation was conducted to evaluate these options and results are presented. Later the three methods are applied to make inferences on the study values.

Chapter 5 presents a brief description of additional problems including variance modeling as a function of concentration, bias modeling as a function of time only, bias modeling as a function of time and concentration (interactions) and trend analysis.

Chapter 6 presents the conclusions and recommendations for future work.
### Table 1.1. Example of detailed reported data and the related CRM reference values.

<table>
<thead>
<tr>
<th>Sample</th>
<th>CASE</th>
<th></th>
<th></th>
<th>CONTROL</th>
<th></th>
<th></th>
<th></th>
<th>SRM1944</th>
<th></th>
<th>mean</th>
<th>s.d.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Repeat</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td></td>
<td>2.00</td>
<td>0.153</td>
<td></td>
<td></td>
</tr>
<tr>
<td>alpha-HCH</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hexachlorobenzene</td>
<td>5.46</td>
<td>5.21</td>
<td>4.84</td>
<td>4.29</td>
<td>4.22</td>
<td>6.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gamma-HCH</td>
<td>&lt;21.2</td>
<td>&lt;25.7</td>
<td>&lt;23.5</td>
<td>&lt;10.9</td>
<td>&lt;13.5</td>
<td>&lt;13.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>beta-HCH</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>heptachlor</td>
<td>&lt;21.2</td>
<td>&lt;25.7</td>
<td>&lt;23.5</td>
<td>&lt;10.9</td>
<td>&lt;13.5</td>
<td>&lt;13.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>aldrin</td>
<td>&lt;21.2</td>
<td>&lt;25.7</td>
<td>&lt;23.5</td>
<td>&lt;10.9</td>
<td>&lt;13.5</td>
<td>&lt;13.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>heptachlor epoxide</td>
<td>&lt;21.2</td>
<td>&lt;25.7</td>
<td>&lt;23.5</td>
<td>&lt;10.9</td>
<td>&lt;13.5</td>
<td>&lt;13.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>oxychlorodane</td>
<td>&lt;21.2</td>
<td>&lt;25.7</td>
<td>&lt;23.5</td>
<td>&lt;10.9</td>
<td>&lt;13.5</td>
<td>&lt;13.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gamma-chlordane</td>
<td>43.6</td>
<td>35.4</td>
<td>47.3</td>
<td>21.7</td>
<td>21.8</td>
<td>32.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2,4'-DDE</td>
<td>40.0</td>
<td>25.6</td>
<td>35.0</td>
<td>8.07</td>
<td>8.47</td>
<td>9.12</td>
<td></td>
<td>19</td>
<td>1.53</td>
<td></td>
<td></td>
</tr>
<tr>
<td>endosulfan I</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cis-chlordane</td>
<td>29.1</td>
<td>31.7</td>
<td>37.0</td>
<td>15.3</td>
<td>16.4</td>
<td>25.9</td>
<td></td>
<td>16.51</td>
<td>0.423</td>
<td></td>
<td></td>
</tr>
<tr>
<td>trans-nonachlor</td>
<td>19.4</td>
<td>15.9</td>
<td>19.9</td>
<td>10.4</td>
<td>8.50</td>
<td>13.6</td>
<td></td>
<td>8.19</td>
<td>0.260</td>
<td></td>
<td></td>
</tr>
<tr>
<td>dieldrin</td>
<td>&lt;21.2</td>
<td>&lt;25.7</td>
<td>&lt;23.5</td>
<td>&lt;10.9</td>
<td>&lt;13.5</td>
<td>&lt;13.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4,4'-DDE</td>
<td>190</td>
<td>160</td>
<td>181</td>
<td>94.6</td>
<td>70.1</td>
<td>114</td>
<td></td>
<td>86</td>
<td>6.12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2,4'-DDD</td>
<td>126</td>
<td>98.3</td>
<td>119</td>
<td>54.8</td>
<td>33.2</td>
<td>56.8</td>
<td></td>
<td>38</td>
<td>4.08</td>
<td></td>
<td></td>
</tr>
<tr>
<td>endrin</td>
<td>&lt;21.2</td>
<td>&lt;25.7</td>
<td>&lt;23.5</td>
<td>&lt;10.9</td>
<td>&lt;13.5</td>
<td>&lt;13.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>endosulfan II</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4,4'-DDD</td>
<td>390</td>
<td>330</td>
<td>404</td>
<td>215</td>
<td>164</td>
<td>276</td>
<td></td>
<td>108</td>
<td>8.163</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2,4'-DDT</td>
<td>&lt;21.2</td>
<td>&lt;25.7</td>
<td>&lt;23.5</td>
<td>&lt;10.9</td>
<td>&lt;13.5</td>
<td>&lt;13.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cis-nonachlor</td>
<td>8.24</td>
<td>6.72</td>
<td>8.02</td>
<td>4.31</td>
<td>3.97</td>
<td>6.10</td>
<td></td>
<td>3.70</td>
<td>0.357</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4,4'-DDT</td>
<td>733</td>
<td>612</td>
<td>758</td>
<td>222</td>
<td>161</td>
<td>274</td>
<td></td>
<td>119</td>
<td>5.61</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mirex</td>
<td>&lt;21.2</td>
<td>&lt;25.7</td>
<td>&lt;23.5</td>
<td>&lt;10.9</td>
<td>&lt;13.5</td>
<td>&lt;13.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>endosulfan sulfate</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>chlorpyrifos</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Consider measuring the quantity of a compound within a material by different methods. For the purpose of this document, this case will be referred as a multiple source measurement exercise (MSME).

The differences between the methods can be due to (i) different measurement procedures, (ii) different facilities and equipment where the materials are analyzed, or (iii) different participating laboratories/analysts making the measurements.

This document describes the estimation of parameters involved in characterizing the quantity of a compound within a material by combining information from a MSME.

2.1 Laboratory Measurement Error Model (LMEM)

For simplicity, consider first measurements for one compound, from one lab, made over a short period of time by one analyst (i.e., repeatability conditions VIM (2006) [34]). This can be extended to handle multivariate measures. When necessary, the convention of underlined letters for a multivariate vector and raw letters for a univariate value are used.

Let the participating laboratory receive $K$ specimens of the same material. The quantity of the compound in the material has a true value $x$ with a variance $\sigma_U^2$, both characteristic of the material. If the material is sufficiently homogeneous and stable then it can be used for calibration or quality assurance purposes (i.e., reference
material VIM (2006) [34]). For the purpose of this document this kind of material is called a control material.

In some places it is necessary to distinguish a random variable from its realized value. When necessary, the usual convention of a capital letter for the random variable and a small letter for the realized value is used.

The quantity of the compound in the $k^{th}$ allocated specimen has a true value $X_k = x + U_k$, where $U_k$ is assumed to be an independent and additive random error with mean $E[U_k] = 0$ and variance $V[U_k] = \sigma_U^2$. The data structure is laid out in 2.1.

**Table 2.1.** Data structure for the control material.

<table>
<thead>
<tr>
<th>observation</th>
<th>1</th>
<th>2</th>
<th>...</th>
<th>$k$</th>
</tr>
</thead>
<tbody>
<tr>
<td>true allocated value (unobserved)</td>
<td>$X_1$</td>
<td>$X_2$</td>
<td>...</td>
<td>$X_k$</td>
</tr>
<tr>
<td>single observed value</td>
<td>$Y_1$</td>
<td>$Y_2$</td>
<td>...</td>
<td>$Y_k$</td>
</tr>
</tbody>
</table>

The laboratory makes its experiments and gets $K$ values, one for each specimen. For the $k^{th}$ specimen the measured value is $Y_k$, where it is assumed that

$$Y_k = g(X_k, \beta) + \epsilon_k$$

(2.1)

where $g$ is a known function and $\beta$ is a vector of parameters. Some models for $g$ will be detailed later. The $\epsilon_k$ term is an independent and additive random error characteristic of the laboratory’s measurement process, occurred while measuring the quantity of the compound in the $k^{th}$ specimen, with mean $E[\epsilon_k] = 0$ and variance $\sigma^2_\epsilon$. The setting described above with $\sigma^2_U = 0$ is sometimes referred to as a functional model while the setting with $\sigma^2_U \neq 0$ is known as a structural model. For our purpose $k$ will denote the order of analysis or measuring of an allocated specimen.

The true value in the specimens allocated to each participant is independent of any participant’s measurement method, so the error terms $U_k$ and $\epsilon_k$ are taken to be independent, then

$$E[Y_k] = E[g(X_k, \beta)]$$

and

$$E[Y_k] = E[g(X_k, \beta)]$$

(2.2)
\[ V[Y_k] = V[g(X_k, \beta)] + \sigma^2 \varepsilon. \]  

(2.3)

The conditional expectation and variance of \( Y_k \) on \( X_k = x_k \) are:

\[ E[Y_k|x_k] = g(x_k, \beta), \]  

(2.4)

\[ V[Y_k|x_k] = V[g(x_k, \beta)] + \sigma^2_\varepsilon, \]  

(2.5)

where \( |x_k \) is a shorthand for the more precise ”given \( X_k = x_k \).” In the case of independent error terms \( U_k, \epsilon_k \) and homoscedasticity over the compound’s level of concentration (2.5) becomes simply \( V[\epsilon_k|x_k] = V[\epsilon_k] = \sigma^2_\varepsilon \). The \( g \) term vanishes from the variance since it is conditionally constant and the new expression strictly depends only on the true value and \( \sigma^2_\varepsilon|x_k = h(x, \theta) \), where the \( \theta \) parameters are not necessarily related to the \( \beta \) parameters. Then \( g \) is the mean measured value given the specimen true value and \( h \) is the variance of the measured value given the specimen true value.

Using the conditional variance identity (e.g. Casella and Berger (2002) [11], Theorem 4.4.7.), (2.3) becomes:

\[ V[Y_k] = V[E[Y_k|X_k]] + E[V[Y_k|X_k]] = V[g(X_k, \beta)] + E[h(X_k, \theta)]. \]  

(2.6)

This is a general model for the variance and generalizes some of the previous approaches in two ways: 1) it is a structural extension of the functional case and 2) it is an extension of the functional dependency of the variance with respect to the true value. Some of these approaches are listed below, where the specimen’s index is omitted for clarity:

- **Anscombe (1961) [3]** stated as \( \sigma^2_\varepsilon|x = \theta_0^2 e^{p(\theta_1 g(x, \beta))} \)

- **Zitter and God (1971) [53]** stated as \( \sigma^2_\varepsilon|x = (\theta_0 + \theta_1 g(x, \beta))^2 \)

- **Amemiya (1973) [1]** stated as \( \sigma^2_\varepsilon|x = \theta_0^2 (g(x, \beta))^2 \)

- **Box and Hill (1974) [4], Horwitz (1982) [25]** stated as \( \sigma^2_\varepsilon|x = \theta_0^2 (g(x, \beta))^{\theta_1} \)
Jobson and Fuller (1980) [33], Thompson et al (2008) [50] stated as
\[ \sigma^2_{e|x} = \theta_0^2 + \theta_1^2 (g(x, \beta))^2 \]

Carroll and Ruppert (1982) [10] stated as \( \sigma^2_{e|x} = h(g(x, \beta), \theta) \)

Duewer et al (1997) [18] stated as \( \sigma^2_{e|x} = \theta_0^2 + \theta_1^2 g(x, \beta)^\theta_2 \)

Buonaccorsi (2006) [8] stated as \( \sigma^2_{e|x} = E[h(X, \Theta)|x] \), where \( \Theta \) is random and can depend on the specific specimen,

By taking \( \sigma^2_{U} \to 0 \) the structural case is reduced to the functional case. Note that \( h(g(X, \beta), \theta) \) allows the conditional variance to depend on \( x \) only through the conditional mean measured value, while \( h(X, \theta) \) allows the conditional variance to depend directly on the true value possibly in ways that cannot be expressed in terms of the measured value. By taking \( g(X, \beta) = X \) (the measured value is conditionally unbiased for the true value) the case is reduced to the two-stage model with fixed parameters described by Buonaccorsi (2006) [8].

2.1.1 Special Cases

Some models for the conditional mean \( g \) are

- Conditionally unbiased model: \( g(X, \beta) = X \)
- Conditionally constant bias model: \( g(X, \beta) = X + \beta_0 \)
- Conditionally linear bias model: \( g(X, \beta) = \beta_0 + \beta_1 X \)
- Conditionally proportional bias model: \( g(X, \beta) = \beta_1 X \)
- Conditionally quadratic bias model: \( g(X, \beta) = \beta_0 + \beta_1 X + \beta_2 X^2 \)
- Conditionally power bias model: \( g(X, \beta) = \beta_0 X^{\beta_1} \)

Some models for the conditional variance \( h \) are
• Variance conditionally constant:
  \[ h(X, \theta) = \theta^2 \]

• Variance conditionally proportional to the squared level:
  \[ h(X, \theta) = \theta^2 X^2 \]

• Variance conditionally linear on the squared level:
  \[ h(X, \theta) = \theta_0^2 + \theta_1^2 X^2 \]

• Variance conditionally quadratic on the level:
  \[ h(X, \theta) = \theta_0^2 + \theta_1 X + \theta_2^2 X^2 \]

• Variance conditionally proportional to a power of the level:
  \[ h(X, \theta) = \theta_0^2 X^{\theta_1} \]

• Variance conditionally exponential on the level:
  \[ h(X, \theta) = \theta_0^2 exp(\theta_1 X) \]

Later these models will be extended to allow changes over time, compounds and laboratories.

It is reasonable to start to model each LMEM in a separate way since each laboratory’s operation is independent from any other one.

2.1.2 Bias of the Method

Recalling (2.1), \( Y_k = g(X_k, \beta) + \epsilon_k \), in general the bias of \( Y_k \) as an estimator of \( x \) is

\[ b = \text{bias}[Y_k] = E[Y_k] - x = E[g(X_k, \beta)] - x \] (2.7)

where \( b \) is possibly a function of \( x, \beta, \sigma_{\epsilon}^2 \) and additional parameters.
2.2 The Study Data

Let the participating laboratory receive $M$ specimens of another material simultaneously with the control material specimens, namely a study material. The study material is similar to the control material but with unknown true composition. That is, each allocated specimen has a true value $D_m = d + V_m$, where $d$ is an unknown constant representing the mean value for the quantity of the compound in the study material and $V_m$ is assumed to be independent with zero mean, $E[V_m] = 0$, and unknown variance $V[V_m] = \sigma^2_V$. The data structure of the study material is laid out in Table 2.2.

<table>
<thead>
<tr>
<th>Observation</th>
<th>$D_1$</th>
<th>$D_2$</th>
<th>...</th>
<th>$D_m$</th>
</tr>
</thead>
<tbody>
<tr>
<td>true allocated value (unobserved)</td>
<td>$D_1$</td>
<td>$D_2$</td>
<td>...</td>
<td>$D_m$</td>
</tr>
<tr>
<td>single observed value</td>
<td>$W_1$</td>
<td>$W_2$</td>
<td>...</td>
<td>$W_m$</td>
</tr>
</tbody>
</table>

The lab measures the study material using the same method used for the control material and gets $M$ measurements, one for each of the specimens, the $m^{th}$ measured value is denoted by $W_m$ with

$$W_m = g(D_m, \beta) + \delta_m$$

where the $\delta_m$ term is a random error corresponding to the $m^{th}$ specimen and has zero mean $E[\delta_m] = 0$ and variance $V[\delta_m] = \sigma^2_\delta$. Under repeatability conditions (VIM (2006) [34]) it can be assumed that $(\epsilon_k, \delta_m)$ share the same distribution within the same laboratory, but in general $\sigma^2_k$ may not equal $\sigma^2_\delta$ if the compound level is very different between the control material and the study material. In principle $\epsilon_k$ and $\delta_m$ are independent if $k \neq m$ (i.e., $Corr[\epsilon_k, \delta_m] = \rho_{k,m} = 0$), but they can be correlated if paired (i.e., $Corr[\epsilon_k, \delta_k] = \rho$).

2.2.1 Pairing of Data

Readings from the control and study materials are intended to be taken in a pairwise way. If for some reason the whole experiment cannot be performed within
the same conditions, the readings are planned to be taken on different days but always pairing one control material reading and one study material reading. This helps to detect large deviations from what is expected and also allows for systematic error to affect both paired readings the same way. This implies the assumption of independent \( \epsilon_k \) and \( \delta_m \) may not hold, but gives credibility to the assertion that the transformation \( g \) (and its \( \beta \) parameters) is shared among both materials while the measurements are conducted. This also implies that ideally \( M = K \) pairs of readings (control, study) are available.

### 2.3 The Certified and Reference Values

The control material is ideally a Certified Reference Material (CRM). These materials are produced by National Metrology Institutes (NMIs), such as NIST, and other organizations with a proven history of higher order measurement capabilities. The certification process is itself an MSME. It is assumed that the quantity of the compound within the material is completely described by some true value \( x \) and its variance \( \sigma^2_U \) which represents variability among subsamples. Both \( x \) and \( \sigma^2_U \) are unknown, instead some estimates are available.

During the certification process two or more critically evaluated independent methods (see May et al. 2000 [39]) are typically used to measure the quantity of the compounds within the material. The measured values are then combined to produce a “conventional value”. The combination of different methods is believed to be superior to any individual method available. Conventional values are used for both reference values and certified values. However a certified value is a reference value with fully characterized uncertainty. Let \( \hat{x} \) and \( \hat{\sigma}^2_x \) be the conventional value and its related variance. Note that \( \hat{x} \) is an estimator of the true value \( x \), and \( \hat{\sigma}^2_x \) is an estimator of the variance of \( \hat{x} \) itself, hence \( \hat{\sigma}^2_x \) can have several variance components.
and is not necessarily an estimate of $\sigma^2_U$ alone. An approximate confidence interval for the true value can be built by using this data as:

$$100(1 - \alpha)\% CI(x) \approx [\hat{x} - 2\hat{\sigma}_x, \hat{x} + 2\hat{\sigma}_x].$$ (2.9)

In order to obtain an estimate of $\sigma^2_U$ we need to consider replication within each allocated unit/specimen in the design of the experiment. Heterogeneity studies are a kind of nested designed experiments with units/splits/replicates levels to estimate if a between units effect is present for each compound (see ISO GUIDE 34 (2009) [31], ISO GUIDE 35 (2006) [29] and references therein). We assume this study is conducted first to assess the within material variance $\sigma^2_U$.

Assuming the method of each participating laboratory during the certification process is unbiased then for the $j^{th}$ participant in the certification:

$$Y_{cjk} = X_{cjk} + \epsilon_{cjk}, j = 1..J, k = 1..K_j$$ (2.10)

where $E[\epsilon_{cjk}] = 0$, and $V[\epsilon_{cjk}|X_{cjk}] = \sigma^2_{c|X_{cjk}} = h(X_{cjk}, \eta_j)$, the $c$ suffix is added to indicate the data is related to the certification. Therefore

$$E[Y_{cjk}] = x_{cj}$$ (2.11)

and

$$V[Y_{cjk}] = \sigma^2_{c|X_{cjk}} = E[h(X_{cjk}, \eta_j)].$$ (2.12)

Under this scenario we can obtain estimators for the true value $x$ and for $\sigma^2_x = V[\hat{x}]$ using a simple weighted mean (see Graybill and Deal (1959) [23]).

The observed values from different participants often appear to be dissimilar, although we still assume the methods are unbiased. In those cases model (2.1) must be extended to include a random effect term to account for the variability between methods, referred to as the variance under reproducibility conditions (VIM (2006) [34]).
Assuming the methods of the participating laboratories during the certification process have a random bias with zero mean, then for the \( j^{\text{th}} \) participant in the certification:

\[
Y_{cjk} = X_{cjk} + b_{cj} + \epsilon_{cjk}, \quad j = 1..J, \quad k = 1..K_j
\]  

(2.13)

where \( b_{cj} \) is the random bias associated to the \( j^{\text{th}} \) laboratory, \( E[b_{cj}] = 0, \quad V[b_{cj}] = \sigma_b^2, \) \( b_{cj}, \) \( E[\epsilon_{cjk}] = 0, \quad V[\epsilon_{cjk}|X_{cjk}] = \sigma_{\text{e|X}}^2, \) \( h(X_{cjk}, \eta_j), \) and \( \epsilon_{cjk} \) are independent. This model is also known as a random bias model or random effects model.

Under this scenario and assuming normality we can obtain estimators for the true value \( x, \) an estimator of \( \sigma_x^2 = V[\hat{x}], \) and the between method variance \( \sigma_b^2 \) by applying the maximum likelihood technique (see Searle et al (1992) [47], Vangel and Rukhin (1999) [51]).

By construction the estimated variance of the conventional value (\( \hat{\sigma}_x^2 \)) can have several variance components (see Searle et al (1992) [47], Buonaccorsi (2006) [8]). The number of variance components is determined by the specific design of the experiment for the certification, while the specific formula for the variance of the conventional value is determined by the way the data are combined to produce the conventional value and the sampling effort. Depending on the producer of the reference material, separate estimates of these variance components may be available, or just one single overall variance estimate. These issues will be addressed in the following chapters as required. Later some variance structures resulting from a certification process will be detailed.

A common way to proceed is to assume the conventional values (\( \hat{x}, \hat{\sigma}_x^2 \)) are the true values (\( x, \sigma_U^2 \)). The consequences of using conventional values instead of the true values on doing inferences about the participants’ method are explored in this work.

### 2.4 The Performance Evaluation

MSME are also conducted to evaluate the performance of the participants. What is an acceptable way to evaluate a method? There exist a set of scores to express and
measure the performance of the participating methods (see ISO GUIDE 13528 (2005) [28]) with different criteria to evaluate them. The existence of multiple schemes and criteria for performance evaluation is confusing and raise some doubts about their validity.

A formal evaluation of the participating method must be based on the statistics resulting from its data, such as: the reported quantity of each compound and the reported variance of the quantity of each compound.

2.5 Data Notation

Table 2.3 summarizes the data notation involved in a MSME.

<table>
<thead>
<tr>
<th>Notation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Y_1, Y_2, ..., Y_k$</td>
<td>Control specimen’s measured values (observed)</td>
</tr>
<tr>
<td>$X_1, X_2, ..., X_k$</td>
<td>Control specimen’s allocated true values (unobserved)</td>
</tr>
<tr>
<td>$\sigma^2_c$</td>
<td>Method/laboratory variance for the control material (unknown)</td>
</tr>
<tr>
<td>$\sigma^2_U$</td>
<td>Heterogeneity variance of the control material (unknown)</td>
</tr>
<tr>
<td>$\hat{x}$</td>
<td>Conventional value (estimate of $x$)</td>
</tr>
<tr>
<td>$\sigma^2_x$</td>
<td>Variance associated with $\hat{x}$</td>
</tr>
<tr>
<td>$\hat{\sigma}^2_b$</td>
<td>Between methods/laboratories estimated variance (unknown)</td>
</tr>
<tr>
<td>$W_1, W_2, ..., W_m$</td>
<td>Study specimen’s measured value (observed)</td>
</tr>
<tr>
<td>$D_1, D_2, ..., D_m$</td>
<td>Study specimen’s allocated true values (unobserved)</td>
</tr>
<tr>
<td>$\sigma^2_d$</td>
<td>Method/laboratory variance for the study material (unknown)</td>
</tr>
<tr>
<td>$\sigma^2_V$</td>
<td>Heterogeneity variance of the study material (unknown)</td>
</tr>
</tbody>
</table>

Additional Parameters

$Corr[\epsilon_k, \delta_k] = \rho$  Method/laboratory error correlation (unknown)
When necessary, the usual convention of a capital letter for the random variable and a small letter for the realized value is used.

The usual convention $X \sim F(\theta)$ is used to state that the random variable $X$ follows the distribution $F(\theta)$ and $X \sim F(\theta)$ to express that the random variable $X$ follows the distribution $F(\theta)$ only approximately.

When necessary, additional subscripted modifiers are explicitly written to identify independent random variables $X_i$ following the distribution $F(\theta)$ as $X_i \sim_{ind} F(\theta)$. Another usual convention is $X \sim_{iid} F(\theta)$ to express that the random variables $X_i$ are independent and identically distributed as $F(\theta)$. 
CHAPTER 3
THE UNBIASED AND CONSTANT BIAS MODEL

The first section of this chapter explores the model for one laboratory, one compound, and one point in time, using only the control data. The model found for the control data is then exported for use with the study data. Finally the model is built by using the control and study data simultaneously when possible. This three-part structure is used in the subsequent sections. Section 3.2 considers the combination of multiple sources at one point in time. Section 3.3 expands the model to multiple compounds at one point in time. Section 3.4 expands the model to vary over time.

3.1 One laboratory, one compound at one point in time

While working with one laboratory at one point in time, the notation is simplified by dropping the related indexes for the corresponding labs and times.

3.1.1 The Basic Model

The model for the observed concentration of the particular compound in the $k^{th}$ allocated unit of the control material is:

$$Y_k = \beta_0 + x + U_k + \epsilon_k \quad (3.1)$$

where $k = 1$ to $K$, $E[U_k] = 0$, $V[U_k] = \sigma_U^2$, $E[\epsilon_k] = 0$, $V[\epsilon_k] = \sigma_\epsilon^2$, and $\beta_0, x, \sigma_U^2$ and $\sigma_\epsilon^2$ are unknown parameters. The $U_k$ and $\epsilon_k$ are independent for a given $k$ and are independent among $k$. Assume that $\hat{x}, \hat{\sigma}_U^2$ and $\hat{\sigma}_\epsilon^2$ are unbiased estimators for $x, \sigma_U^2$ and $V[\hat{x}]$ and we are given values of these (these come from the certification process as discussed in Chapter 2).
The model for the observed concentration of the particular compound in the study material is:

\[ W_k = \beta_0 + d + V_k + \delta_k \]  \hspace{1cm} (3.2)

where \( k = 1 \) to \( K \), \( E[\delta_k] = 0 \), \( V[\delta_k] = \sigma_\delta^2 \), \( E[V_k] = 0 \), \( Var[V_k] = \sigma_V^2 \), and \( d \) and \( \sigma_V^2 \) are unknown parameters. The bias term \( \beta_0 \) is shared by both models, \( V_k \) and \( \delta_k \) are independent for a given \( k \) and are independent among \( k \). In addition, assume that \( \delta_k, V_k \) are independent of \( \epsilon_k, U_k \); then

\[ E[Y_k] = \beta_0 + x, \]
\[ V[Y_k] = \sigma_U^2 + \sigma_\epsilon^2, \]  \hspace{1cm} (3.3)
\[ E[W_k] = \beta_0 + d \]
\[ V[W_k] = \sigma_V^2 + \sigma_\delta^2. \]

3.1.2 Working with the Control Data Only

The main goal is to obtain estimates of each individual LMEM, hence there is no interest on pooling information across different participants. However there is interest in testing whether a participant’s method is biased. When working with the control data only, the performance of the participants can be assessed directly provided that the conventional value of the control material and its related variance components are known or estimated.

3.1.2.1 Parameter Estimation

In practice \( E[Y_k] \) and \( V[Y_k] \) are unknown and estimates must be used instead. Define the sample mean \( \bar{Y} = \frac{1}{K} \sum_{k=1}^{K} Y_k \) which is unbiased for \( E[Y_k] \) and the sample variance \( s_Y^2 = \frac{1}{K-1} \sum_{k=1}^{K} (Y_k - \bar{Y})^2 \) which is unbiased for \( V[Y_k] \). Let

\[ \hat{\beta}_0 = \bar{Y} - \hat{x} \]  \hspace{1cm} (3.4)

and
\[ \hat{\sigma}^2 = s_Y^2 - \tilde{\sigma}_U^2 \] (3.5)

then

\[ E[\hat{\beta}_0] = E[\hat{Y} - \hat{x}] = \beta_0 + x - E[\hat{x}] = \beta_0, \] (3.6)

\[ E[\hat{\sigma}_\epsilon^2] = E[s_Y^2 - \tilde{\sigma}_U^2] = \sigma^2_\epsilon + \tilde{\sigma}_U^2 - E[\tilde{\sigma}_U^2] = \sigma^2_\epsilon \] and

\[ V[\hat{\beta}_0] = V[\hat{Y} - \hat{x}] = E\left[ \frac{s_Y^2}{K} + \hat{\sigma}_x^2 \right]. \] (3.8)

Then \( \hat{\beta}_0, \hat{\sigma}^2_\epsilon, \) and \( \hat{\sigma}^2_{\hat{\beta}_0} = \frac{s_Y^2}{K} + \hat{\sigma}_x^2 \) are unbiased estimators of \( \beta_0, \sigma^2_\epsilon, \) and \( V[\hat{\beta}_0] \) respectively. However, this definition for \( \hat{\sigma}^2_\epsilon \) allows a negative value of a non-negative quantity. In such a case the interpretation is that \( \sigma^2_\epsilon \) is zero. Sometimes it is preferable to obtain a confidence interval for \( \sigma^2_\epsilon \), as described in (3.15) or (3.16).

Some additional distributional assumptions on the error terms are required in order to get small sample confidence intervals for \( \beta_0 \) and \( \sigma^2_\epsilon \). Assuming normality of \( U_k \) and \( \epsilon_k \), and assuming \( \hat{\sigma}_x^2 \) is distributed as \( \chi^2 \) with \( \nu_x \) degrees of freedom, the distribution of \( \hat{\beta}_0 \) is

\[ \frac{\hat{\beta}_0 - \beta_0}{\hat{\sigma}_{\hat{\beta}_0}} \sim T_{\nu_{\beta_0}}. \] (3.9)

Hence, an approximate 100(1 - \( \alpha \))% confidence interval (CI) for \( \beta_0 \) is given by

\[ \hat{\beta}_0 \pm t(\alpha/2, \nu_{\beta_0})\hat{\sigma}_{\hat{\beta}_0}, \] (3.10)

where \( \nu_{\beta_0} \) is estimated by using the Welch-Satterthwaite approximation (see Appendix A)

\[ \nu_{\beta_0} = \frac{(s_Y^2 + \hat{\sigma}_x^2)^2}{\frac{s_Y^2}{K^2(K-1)} + \hat{\sigma}_x^2}. \] (3.11)

When \( K \) and \( \nu_x \) are large the distribution is

\[ \frac{\hat{\beta}_0 - \beta_0}{\hat{\sigma}_{\hat{\beta}_0}} \sim Z \] (3.12)

and an approximate CI for \( \beta_0 \) is given by
\[
\hat{\beta}_0 \pm z(\alpha/2)\hat{\sigma}_{\hat{\beta}_0}.
\] (3.13)

Similarly, the exact distribution of \( s^2_Y \) can be obtained from
\[
\frac{(K - 1)s^2_Y}{\sigma^2_U + \sigma^2_{\epsilon}} \sim \chi^2_{K-1}
\] (3.14)

and an approximate CI for \( \sigma^2_{\epsilon} \) is
\[
\left(\frac{(K - 1)s^2_Y}{\chi^2_{(1-\alpha/2;K-1)}} - \hat{\sigma}^2_{\epsilon}, \frac{(K - 1)s^2_Y}{\chi^2_{(\alpha/2;K-1)}} - \hat{\sigma}^2_{\epsilon}\right)
\] (3.15)

assuming the estimator \( \hat{\sigma}^2_{\epsilon} \) equals the true parameter \( \sigma^2_{\epsilon} \).

There is potential for getting estimates outside of the parameter space for \( \sigma^2_{\epsilon} \). The interval should be truncated so that it contains only non-negative values. However the coverage of the estimated confidence interval would decrease. If the upper bound of (3.15) is less than zero then the interpretation would be that \( \sigma^2_{\epsilon} \) is zero.

Another option is to use additional information about \( U_k \) if it is available. Under normality suppose \( \nu_U \hat{\sigma}^2_U/\sigma^2_U \sim \chi^2(\nu_U) \) then a 100(1-\alpha)\% CI for \( \sigma^2_U \), say \([ lcl(\sigma^2_U), ucl(\sigma^2_U) ]\) can be constructed. Using Bonferroni’s method a CI with confidence level 1 - 2\alpha is given by
\[
\left(\frac{(K - 1)s^2_Y}{\chi^2_{(1-\alpha/2;K-1)}} - ucl(\sigma^2_U), \frac{(K - 1)s^2_Y}{\chi^2_{(\alpha/2;K-1)}} - lcl(\sigma^2_U)\right)
\] (3.16)

Thus even in the simplest case a CRM must provide at least 3 unbiased estimates \((\hat{x}, \hat{\sigma}^2_{\hat{x}}, \hat{\sigma}^2_U)\) for each compound within the material. Sometimes the CRM certificate just reports that evidence of heterogeneity was not found as the result of conducted studies on the material. This implies \( \hat{\sigma}^2_U = 0 \) for all the reported compounds within the material.
3.1.2.2 Inferences about the Parameters

*Testing for Zero Bias*

The hypothesis of unbiasedness \((H_0 : \beta_0 = 0)\) can be tested approximately using (3.9) or (3.12). For small \(K\), \(H_0\) can be rejected and \((\beta_0 \neq 0)\) can be concluded with a significance level of \(\alpha\) if

\[
\left|\frac{\hat{\beta}_0}{\hat{\sigma}_{\hat{\beta}_0}}\right| > t_{(1-\alpha/2; \nu_{\hat{\beta}_0})}.
\]  
(3.17)

Likewise for large \(K\), if

\[
z = \frac{|\hat{\beta}_0|}{\hat{\sigma}_{\hat{\beta}_0}} > z_{(1-\alpha/2)}.
\]  
(3.18)

While these tests allow testing the bias for zero equality, they do not address testing for bias constancy. Two or more points are required across compounds or time in order to test for constancy of the bias or constancy of the within variance. This is addressed in Section 3.3.

3.1.2.3 Performance Evaluation

The statistic \(z\) in (3.18) is called the \(z' - score\) and it is often used to evaluate the performance of the participants in MSME. There are several proposed statistics to evaluate a method’s performance (ISO GUIDE 13528 (2005) \([28]\)).

Participant performance can be assessed directly by using only the control data and the conventional values of the reference material. Using (3.17) and taking advantage of the symmetry of the \(T\) distribution

\[
t_{\text{score}} = \frac{\hat{\beta}_0}{\hat{\sigma}_{\hat{\beta}_0}} \text{ and } |t_{\text{score}}| < \text{ some critical value}.
\]  
(3.19)

- For small \(K\), the usual evaluation criteria are:
  - \(t_{\text{score}}\) is satisfactory if \(|t_{\text{score}}| \leq t_{(0.977; \nu_{\hat{\beta}_0})} = t_{\text{sat}}\),

  - \(t_{\text{score}}\) is unsatisfactory if \(|t_{\text{score}}| > t_{(0.999; \nu_{\hat{\beta}_0})} = t_{\text{que}}\) and
\( t_{\text{score}} \) is questionable if \( t_{\text{sat}} < |t_{\text{score}}| \leq t_{\text{que}} \).

- For large \( K \), the evaluation equivalent criteria are:

  - \( t_{\text{score}} \) is satisfactory if \( |t_{\text{score}}| \leq 2.0 = z_{\text{sat}} \approx z(0.977) \),
  - \( t_{\text{score}} \) is unsatisfactory if \( |t_{\text{score}}| > 3.0 = z_{\text{que}} \approx z(0.999) \) and
  - \( t_{\text{score}} \) is questionable if \( z_{\text{sat}} < |t_{\text{score}}| \leq z_{\text{que}} \).

The satisfactory and unsatisfactory results are interpreted as evidence of unbiasedness and biasedness of the method respectively. The percentiles 0.977 and 0.999 are chosen to match the conventional critical values 2 and 3 standard deviations from the mean of 0 respectively under normality.

**Assessing the variance**

The above is a partial evaluation centered on the location bias, estimated by \( \hat{\beta}_0 \) as in (3.4). There is additional information available about the variance of the participant’s method in the \( h \) model, estimated by \( \hat{\sigma}_2^2 \) as in (3.5). What is a satisfactory or unsatisfactory method variance?

The evaluation of the participant’s method variability can be done by comparing it against the variability of the methods used during the certification. This requires that the variance component associated to the within variability of the methods used in the CRM certification be readily available (see Section 2.3). We use method and laboratory interchangeably here. The within-method variance is the expected variance of the error in the equation (Fuller (1987) [22]) of a method selected randomly from the population of methods and it is also called the repeatability variance of the method in a metrology context (ISO GUIDE 5725 (1998) [27]). Let \( \sigma_{ec}^2 \) be the true within method variance of the population of methods at the certification time. This is the expected within variance of a single method chosen randomly from the population of methods. Assume the methods participating in the certification are a
representative sample of the methods’ space. Let $\hat{\sigma}_{ec}^2$ and $\nu_{ec}$ be an unbiased estimator for $\sigma_{ec}^2$ and its related degrees of freedom obtained during the certification.

**Point-wise evaluation:** Under normality the within variance of the conventional value

$$\nu_{ec} \frac{\hat{\sigma}_{ec}^2}{\sigma_{ec}^2} \sim \chi^2(\nu_{ec}) \tag{3.20}$$

and the within variance of any other method can be seen as a random variable from the distribution

$$(K - 1) \frac{s_Y^2 - \hat{\sigma}_U^2}{\hat{\sigma}_{ec}^2} \sim \chi^2(K - 1), \tag{3.21}$$

hence if $\sigma_e^2 = \sigma_{ec}^2$ then

$$f_{score} = \frac{s_Y^2 - \hat{\sigma}_U^2}{\hat{\sigma}_{ec}^2} \sim F(K - 1, \nu_{ec}). \tag{3.22}$$

The participant’s variance performance can be assessed directly by using only the control data and the conventional values of the reference material. Using (3.22) and considering the asymmetry of the F distribution

$$f_{score} \in \text{some acceptance region.} \tag{3.23}$$

- For arbitrary $K$, $\nu_{ec} = \text{dof}[\hat{\sigma}_{ec}^2]$, the evaluation criteria are:

  - $f_{score}$ is satisfactory if $f_{score} \in [f(0.023;K - 1, \nu_{ec}), f(0.977;K - 1, \nu_{ec})] = S$;
  - $f_{score}$ is unsatisfactory if $f_{score} \notin [f(0.001;K - 1, \nu_{ec}), f(0.999;K - 1, \nu_{ec})] = Q$ and
  - $f_{score}$ is questionable if $f_{score} \in Q \cap S^c$, where $S^c$ is the complement of $S$.

The satisfactory and unsatisfactory results are typically interpreted as evidence of unbiasedness and biasedness (underestimation or overestimation) of the within variance of the participating method respectively.
3.1.2.4 Examples

As an example, Table 3.1 lists the conventional values for the control material SRM 1941b and the observed values as reported by one of the participants about the PAH family of compounds in sediments from the IEPOCME 2003 inter-comparison exercise.

Applying (3.4), (3.8), (3.10) and (3.13) we obtain the results in Table 3.2. Similarly applying (3.5) and (3.15) we obtain the results shown in Table 3.3.

Applying (3.19), (3.23) and using the data in Table 3.1 we obtain the evaluation of the within variance in Table 3.4, for convenience of display the $\sqrt{f_{\text{score}}}$ are listed, all the $t_{\text{scores}}$ are satisfactory, eight $f_{\text{scores}}$ appear questionably small and 18 are satisfactory. In addition in the Figure 3.1 we show a target plot combining both scores for each compound applying the square root on the $f_{\text{score}}$ in order to display comparable quantities.
<table>
<thead>
<tr>
<th>Compound</th>
<th>$\bar{Y}$</th>
<th>$s_Y$</th>
<th>$K$</th>
<th>$\bar{W}$</th>
<th>$s_W$</th>
<th>$M$</th>
<th>$\hat{\rho}_{YW}$</th>
<th>$\hat{x}$</th>
<th>$\hat{\sigma}_x$</th>
<th>$\nu_x$</th>
<th>$\hat{\sigma}_{\epsilon c}$</th>
<th>$\nu_{\epsilon c}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2,3,5-trimethylnaphthalene</td>
<td>22.7</td>
<td>1.05</td>
<td>3</td>
<td>25.7</td>
<td>0.56</td>
<td>3</td>
<td>0.69</td>
<td>25.5</td>
<td>2.60</td>
<td>29</td>
<td>3.18</td>
<td>9</td>
</tr>
<tr>
<td>acenaphthene</td>
<td>36.0</td>
<td>2.19</td>
<td>3</td>
<td>34.7</td>
<td>0.46</td>
<td>3</td>
<td>-0.89</td>
<td>38.4</td>
<td>2.65</td>
<td>30</td>
<td>3.23</td>
<td>13</td>
</tr>
</tbody>
</table>
dibenzo[a,h]anthracene   | 56.3      | 0.69  | 3   | 112.0     | 4.16  | 3   | 0.28                | 53.0      | 5.10             | 10       | 3.20              | 4                |
|acenaphthylene            | 56.8      | 2.07  | 3   | 29.6      | 1.36  | 3   | -0.52               | 53.3      | 3.27             | 28       | 4.80              | 12               |
|1-methylphenanthrene      | 71.8      | 0.66  | 3   | 69.4      | 0.62  | 3   | 0.92                | 73.2      | 3.01             | 26       | 5.38              | 12               |
|biphenyl                  | 70.0      | 2.03  | 3   | 29.2      | 0.71  | 3   | -1.00               | 74.0      | 4.08             | 28       | 6.29              | 13               |
|2,6-dimethylnaphthalene   | 74.0      | 1.98  | 3   | 35.2      | 0.46  | 3   | 0.52                | 75.9      | 2.30             | 23       | 3.91              | 14               |
|fluorene                  | 85.6      | 3.57  | 3   | 65.5      | 0.69  | 3   | -0.95               | 85.0      | 7.65             | 39       | 7.24              | 9                |
triphenylene              | 111.0     | 2.65  | 3   | 141.0     | 8.50  | 3   | 0.49                | 108.0     | 2.55             | 11       | 4.46              | 3                |
|1-methylnaphthalene       | 120.7     | 6.51  | 3   | 49.0      | 0.91  | 3   | -0.43               | 127.0     | 7.14             | 30       | 13.89             | 12               |
anthracene                | 181.7     | 6.03  | 3   | 172.3     | 2.52  | 3   | 1.00                | 184.0     | 9.18             | 39       | 10.74             | 9                |
|benzo[j]fluoranthene      | 220.7     | 1.53  | 3   | 281.0     | 8.50  | 3   | -0.01               | 217.0     | 2.55             | 5        | 5.00              | 4                |
|benzo[k]fluoranthene      | 226.3     | 5.03  | 3   | 406.0     | 8.72  | 3   | -0.87               | 225.0     | 9.18             | 24       | 10.38             | 7                |
|2-methylnaphthalene       | 284.0     | 11.79 | 3   | 79.2      | 0.81  | 3   | -0.57               | 276.0     | 27.04            | 32       | 30.26             | 12               |
|chrysene                  | 312.3     | 12.50 | 3   | 547.0     | 3.79  | 3   | -0.97               | 291.0     | 15.82            | 9        | 17.96             | 5                |
|benzo[ghi]perylene        | 327.3     | 5.51  | 3   | 488.0     | 5.00  | 3   | 0.45                | 307.0     | 22.96            | 36       | 23.08             | 11               |
|benzo[e]pyrene            | 312.7     | 6.66  | 3   | 590.0     | 12.01 | 3   | 0.92                | 325.0     | 12.76            | 28       | 21.91             | 14               |
|benz[a]anthracene         | 326.3     | 9.87  | 3   | 378.0     | 4.04  | 3   | 0.96                | 335.0     | 12.76            | 39       | 20.49             | 10               |
|indeno[1,2,3-cd]pyrene    | 359.7     | 8.08  | 3   | 608.0     | 7.94  | 3   | -0.98               | 341.0     | 29.08            | 36       | 32.45             | 11               |
|benzo[a]pyrene            | 361.7     | 5.03  | 3   | 636.0     | 5.69  | 3   | 0.99                | 358.0     | 8.67             | 30       | 19.72             | 12               |
|perylene                  | 412.3     | 9.61  | 3   | 184.0     | 7.00  | 3   | 0.68                | 397.0     | 22.96            | 33       | 35.29             | 11               |
|phenanthrene              | 433.3     | 10.02 | 3   | 539.0     | 7.00  | 3   | -0.76               | 406.0     | 22.45            | 39       | 25.97             | 9                |
|benzo[b]fluoranthene      | 463.0     | 6.93  | 3   | 1089.0    | 19.50 | 3   | -0.61               | 453.0     | 10.71            | 10       | 19.17             | 4                |
|pyrene                    | 565.3     | 10.21 | 3   | 1129.0    | 13.75 | 3   | -0.65               | 581.0     | 19.90            | 38       | 40.48             | 10               |
|fluoranthene              | 638.0     | 33.42 | 3   | 1094.0    | 9.64  | 3   | -0.38               | 651.0     | 25.51            | 38       | 41.96             | 10               |
naphthalene               | 873.7     | 8.08  | 3   | 121.3     | 2.52  | 3   | -0.26               | 848.0     | 48.47            | 38       | 57.54             | 10               |

**Table 3.1.** Example of summary data and reference values.

For each compound, $\bar{Y}$ = mean of the control values, $s_Y$ = standard deviation of the control values, $K$ = sample size of the control sample, $\bar{W}$ = mean of the study values, $s_W$ = standard deviation of the study values, $M$ = sample size of the study sample, $\hat{\rho}_{YW}$ = the estimated correlation coefficient between the control and study samples, $\hat{x}$ = reference value estimate, $\hat{\sigma}_x$ = the uncertainty of the reference value, $\nu_x$ = the degrees of freedom of the reference value uncertainty, $\hat{\sigma}_{\epsilon c}$ = the estimated within lab standard deviation and $\nu_{\epsilon c}$ = degrees of freedom of the within lab uncertainty.
<table>
<thead>
<tr>
<th>Compound</th>
<th>( \hat{\beta}_0 )</th>
<th>( \hat{\sigma}_{\hat{\beta}_0} )</th>
<th>( \nu_{\hat{\beta}_0} )</th>
<th>( lcl_B )</th>
<th>( ucl_B )</th>
<th>( lcl_{exact} )</th>
<th>( ucl_{exact} )</th>
<th>( lcl_T )</th>
<th>( ucl_T )</th>
<th>( lcl_N )</th>
<th>( ucl_N )</th>
</tr>
</thead>
<tbody>
<tr>
<td>acenaphthene</td>
<td>-2.43</td>
<td>2.94</td>
<td>25</td>
<td>-10.10</td>
<td>5.14</td>
<td>-10.00</td>
<td>5.15</td>
<td>-8.48</td>
<td>3.62</td>
<td>-8.20</td>
<td>3.33</td>
</tr>
<tr>
<td>acenaphthylene</td>
<td>3.50</td>
<td>3.48</td>
<td>30</td>
<td>-4.94</td>
<td>11.90</td>
<td>-4.91</td>
<td>11.90</td>
<td>-3.60</td>
<td>10.60</td>
<td>-3.31</td>
<td>10.30</td>
</tr>
<tr>
<td>1-methylphenanthrene</td>
<td>-1.40</td>
<td>3.03</td>
<td>27</td>
<td>-7.88</td>
<td>5.06</td>
<td>-7.86</td>
<td>5.06</td>
<td>-7.63</td>
<td>4.83</td>
<td>-7.35</td>
<td>4.55</td>
</tr>
<tr>
<td>biphenyl</td>
<td>-4.00</td>
<td>4.25</td>
<td>30</td>
<td>-13.80</td>
<td>5.81</td>
<td>-13.80</td>
<td>5.82</td>
<td>-12.70</td>
<td>4.67</td>
<td>-12.30</td>
<td>4.32</td>
</tr>
<tr>
<td>2,6-dimethylnaphthalene</td>
<td>-1.87</td>
<td>2.56</td>
<td>21</td>
<td>-8.63</td>
<td>4.87</td>
<td>-8.61</td>
<td>4.87</td>
<td>-7.20</td>
<td>3.46</td>
<td>-6.89</td>
<td>3.16</td>
</tr>
<tr>
<td>fluorene</td>
<td>0.60</td>
<td>7.93</td>
<td>41</td>
<td>-17.40</td>
<td>18.60</td>
<td>-17.40</td>
<td>18.60</td>
<td>-15.40</td>
<td>16.60</td>
<td>-14.90</td>
<td>16.10</td>
</tr>
<tr>
<td>triphenylene</td>
<td>3.00</td>
<td>2.97</td>
<td>8</td>
<td>-8.48</td>
<td>4.87</td>
<td>-8.48</td>
<td>4.87</td>
<td>-7.20</td>
<td>3.46</td>
<td>-6.89</td>
<td>3.16</td>
</tr>
<tr>
<td>benzo[j]fluoranthene</td>
<td>3.67</td>
<td>2.70</td>
<td>6</td>
<td>-8.48</td>
<td>8.50</td>
<td>-8.48</td>
<td>8.50</td>
<td>-7.20</td>
<td>4.67</td>
<td>-6.89</td>
<td>3.16</td>
</tr>
<tr>
<td>2-methylnaphthalene</td>
<td>8.00</td>
<td>27.90</td>
<td>34</td>
<td>-55.00</td>
<td>70.90</td>
<td>-55.00</td>
<td>70.90</td>
<td>-48.70</td>
<td>64.70</td>
<td>-46.70</td>
<td>62.70</td>
</tr>
<tr>
<td>chrysene</td>
<td>21.30</td>
<td>17.40</td>
<td>11</td>
<td>-26.50</td>
<td>68.90</td>
<td>-26.50</td>
<td>68.90</td>
<td>-16.90</td>
<td>59.60</td>
<td>-12.70</td>
<td>55.40</td>
</tr>
<tr>
<td>benzo[a]anthracene</td>
<td>-8.67</td>
<td>14.00</td>
<td>32</td>
<td>-44.00</td>
<td>26.50</td>
<td>-43.80</td>
<td>26.40</td>
<td>-37.10</td>
<td>19.80</td>
<td>-36.00</td>
<td>18.70</td>
</tr>
<tr>
<td>indeno[1,2,3-cd]pyrene</td>
<td>18.70</td>
<td>29.50</td>
<td>37</td>
<td>-44.50</td>
<td>81.60</td>
<td>-44.30</td>
<td>81.60</td>
<td>-41.00</td>
<td>78.30</td>
<td>-39.10</td>
<td>76.40</td>
</tr>
<tr>
<td>benzo[a]pyrene</td>
<td>3.67</td>
<td>9.15</td>
<td>31</td>
<td>-18.10</td>
<td>25.30</td>
<td>-18.00</td>
<td>25.40</td>
<td>-15.00</td>
<td>22.30</td>
<td>-14.30</td>
<td>21.60</td>
</tr>
<tr>
<td>perylene</td>
<td>15.30</td>
<td>23.60</td>
<td>35</td>
<td>-37.70</td>
<td>68.20</td>
<td>-37.50</td>
<td>68.20</td>
<td>-32.60</td>
<td>63.30</td>
<td>-31.00</td>
<td>61.60</td>
</tr>
<tr>
<td>phenanthrene</td>
<td>27.30</td>
<td>23.20</td>
<td>41</td>
<td>-25.00</td>
<td>79.50</td>
<td>-24.80</td>
<td>79.40</td>
<td>-19.50</td>
<td>74.20</td>
<td>-18.10</td>
<td>72.80</td>
</tr>
<tr>
<td>benzo[b]fluoranthene</td>
<td>10.00</td>
<td>11.40</td>
<td>12</td>
<td>-19.80</td>
<td>39.70</td>
<td>-19.60</td>
<td>39.60</td>
<td>-15.00</td>
<td>35.00</td>
<td>-12.40</td>
<td>32.40</td>
</tr>
<tr>
<td>pyrene</td>
<td>-15.70</td>
<td>20.80</td>
<td>39</td>
<td>-63.60</td>
<td>32.20</td>
<td>-63.40</td>
<td>32.10</td>
<td>-57.60</td>
<td>26.30</td>
<td>-56.30</td>
<td>25.00</td>
</tr>
<tr>
<td>fluoranthene</td>
<td>-13.00</td>
<td>32.00</td>
<td>13</td>
<td>-109.00</td>
<td>82.70</td>
<td>-106.00</td>
<td>80.50</td>
<td>-82.10</td>
<td>56.10</td>
<td>-75.70</td>
<td>49.70</td>
</tr>
<tr>
<td>naphthalene</td>
<td>25.70</td>
<td>48.70</td>
<td>39</td>
<td>-75.50</td>
<td>127.00</td>
<td>-75.20</td>
<td>127.00</td>
<td>-72.90</td>
<td>124.00</td>
<td>-69.80</td>
<td>121.00</td>
</tr>
</tbody>
</table>

Table 3.2. Example of point and interval estimation of a method constant bias.

For each compound, \( \hat{\beta}_0 \) = the estimated bias, \( \hat{\sigma}_{\hat{\beta}_0} \) = standard error of the estimated bias, \( \nu_{\hat{\beta}_0} \) = the degrees of freedom of the standard error, the 95% confidence intervals using: the bootstrap method (\( lcl_B, ucl_B \)), the exact distribution (via numerical integration) (\( lcl_{exact}, ucl_{exact} \)), the \( t \) distribution with Welch-Satterthwaite approximated degrees of freedom (\( lcl_T, ucl_T \)), and the normal distribution (\( lcl_N, ucl_N \)).
<table>
<thead>
<tr>
<th>Compound</th>
<th>$\hat{\sigma}$</th>
<th>$\nu_\epsilon$</th>
<th>lcl($\hat{\sigma}$)</th>
<th>ucl($\hat{\sigma}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2,3,5-trimethylnaphthalene</td>
<td>1.1</td>
<td>2</td>
<td>0.6</td>
<td>6.6</td>
</tr>
<tr>
<td>acenaphthene</td>
<td>2.2</td>
<td>2</td>
<td>1.1</td>
<td>14.0</td>
</tr>
<tr>
<td>dibenz[a,h]anthracene</td>
<td>0.7</td>
<td>2</td>
<td>0.4</td>
<td>4.4</td>
</tr>
<tr>
<td>acenaphthylene</td>
<td>2.1</td>
<td>2</td>
<td>1.1</td>
<td>13.0</td>
</tr>
<tr>
<td>1-methylphenanthrene</td>
<td>0.7</td>
<td>2</td>
<td>0.3</td>
<td>4.1</td>
</tr>
<tr>
<td>biphenyl</td>
<td>2.0</td>
<td>2</td>
<td>1.1</td>
<td>13.0</td>
</tr>
<tr>
<td>2,6-dimethylnaphthalene</td>
<td>2.0</td>
<td>2</td>
<td>1.0</td>
<td>12.0</td>
</tr>
<tr>
<td>fluorene</td>
<td>3.6</td>
<td>2</td>
<td>1.9</td>
<td>22.0</td>
</tr>
<tr>
<td>triphenylene</td>
<td>2.6</td>
<td>2</td>
<td>1.4</td>
<td>17.0</td>
</tr>
<tr>
<td>1-methylnaphthalene</td>
<td>6.5</td>
<td>2</td>
<td>3.4</td>
<td>41.0</td>
</tr>
<tr>
<td>anthracene</td>
<td>6.0</td>
<td>2</td>
<td>3.1</td>
<td>38.0</td>
</tr>
<tr>
<td>benzo[j]fluoranthene</td>
<td>1.5</td>
<td>2</td>
<td>0.8</td>
<td>9.6</td>
</tr>
<tr>
<td>benzo[k]fluoranthene</td>
<td>5.0</td>
<td>2</td>
<td>2.6</td>
<td>32.0</td>
</tr>
<tr>
<td>2-methylnaphthalene</td>
<td>12.0</td>
<td>2</td>
<td>6.1</td>
<td>74.0</td>
</tr>
<tr>
<td>chrysene</td>
<td>13.0</td>
<td>2</td>
<td>6.5</td>
<td>79.0</td>
</tr>
<tr>
<td>benzo[ghi]perylene</td>
<td>5.5</td>
<td>2</td>
<td>2.9</td>
<td>35.0</td>
</tr>
<tr>
<td>benzo[e]pyrene</td>
<td>6.7</td>
<td>2</td>
<td>3.5</td>
<td>42.0</td>
</tr>
<tr>
<td>benzo[a]anthracene</td>
<td>9.9</td>
<td>2</td>
<td>5.1</td>
<td>62.0</td>
</tr>
<tr>
<td>indeno[1,2,3-cd]pyrene</td>
<td>8.1</td>
<td>2</td>
<td>4.2</td>
<td>51.0</td>
</tr>
<tr>
<td>benzo[a]pyrene</td>
<td>5.0</td>
<td>2</td>
<td>2.6</td>
<td>32.0</td>
</tr>
<tr>
<td>perylene</td>
<td>9.6</td>
<td>2</td>
<td>5.0</td>
<td>60.0</td>
</tr>
<tr>
<td>phenanthrene</td>
<td>10.0</td>
<td>2</td>
<td>5.2</td>
<td>63.0</td>
</tr>
<tr>
<td>benzo[b]fluoranthene</td>
<td>6.9</td>
<td>2</td>
<td>3.6</td>
<td>44.0</td>
</tr>
<tr>
<td>pyrene</td>
<td>10.0</td>
<td>2</td>
<td>5.3</td>
<td>64.0</td>
</tr>
<tr>
<td>fluoranthene</td>
<td>33.0</td>
<td>2</td>
<td>17.0</td>
<td>210.0</td>
</tr>
<tr>
<td>naphthalene</td>
<td>8.1</td>
<td>2</td>
<td>4.2</td>
<td>51.0</td>
</tr>
</tbody>
</table>

Table 3.3. Example of point and interval estimates of the method within uncertainty.

For each compound, $\hat{\sigma}$ is the estimated within lab standard deviation, $\nu_\epsilon$ is the degrees of freedom of the within lab standard deviation, lcl($\hat{\sigma}$) is the lower confidence limit for the within lab standard deviation, and ucl($\hat{\sigma}$) is the upper confidence limit for the within lab standard deviation.
<table>
<thead>
<tr>
<th>Compound</th>
<th>$t_{score}$</th>
<th>$t_{sat}$</th>
<th>$\sqrt{f_{score}}$</th>
<th>$\sqrt{f_{queL}}$</th>
<th>$\sqrt{f_{satL}}$</th>
<th>$\sqrt{f_{satU}}$</th>
<th>$\sqrt{f_{queU}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2,3,5-trimethylnaphthalene</td>
<td>-1.05</td>
<td>2.08</td>
<td>0.33</td>
<td>0.03</td>
<td>0.15</td>
<td>2.43</td>
<td>4.05</td>
</tr>
<tr>
<td>acenaphthene</td>
<td>-0.83</td>
<td>2.10</td>
<td>0.68</td>
<td>0.03</td>
<td>0.15</td>
<td>2.26</td>
<td>3.51</td>
</tr>
<tr>
<td>dibenz[a,h]anthracene</td>
<td>0.64</td>
<td>2.27</td>
<td>0.22</td>
<td>0.03</td>
<td>0.15</td>
<td>3.34</td>
<td>7.83</td>
</tr>
<tr>
<td>acenaphthylene</td>
<td>1.01</td>
<td>2.09</td>
<td>0.43</td>
<td>0.03</td>
<td>0.15</td>
<td>2.29</td>
<td>3.60</td>
</tr>
<tr>
<td>biphenyl</td>
<td>-0.94</td>
<td>2.08</td>
<td>0.32</td>
<td>0.03</td>
<td>0.15</td>
<td>2.26</td>
<td>3.51</td>
</tr>
<tr>
<td>1-methylphenanthrene</td>
<td>-0.46</td>
<td>2.09</td>
<td>0.12</td>
<td>0.03</td>
<td>0.15</td>
<td>2.29</td>
<td>3.60</td>
</tr>
<tr>
<td>2,6-dimethylnaphthalene</td>
<td>-0.73</td>
<td>2.12</td>
<td>0.50</td>
<td>0.03</td>
<td>0.15</td>
<td>2.24</td>
<td>3.43</td>
</tr>
<tr>
<td>fluorene</td>
<td>0.08</td>
<td>2.06</td>
<td>0.49</td>
<td>0.03</td>
<td>0.15</td>
<td>2.43</td>
<td>4.05</td>
</tr>
<tr>
<td>triphenylene</td>
<td>1.01</td>
<td>2.23</td>
<td>0.59</td>
<td>0.03</td>
<td>0.15</td>
<td>4.13</td>
<td>12.20</td>
</tr>
<tr>
<td>1-methylnaphthalene</td>
<td>-0.78</td>
<td>2.11</td>
<td>0.47</td>
<td>0.03</td>
<td>0.15</td>
<td>2.29</td>
<td>3.60</td>
</tr>
<tr>
<td>anthracene</td>
<td>-0.24</td>
<td>2.07</td>
<td>0.56</td>
<td>0.03</td>
<td>0.15</td>
<td>2.43</td>
<td>4.05</td>
</tr>
<tr>
<td>benzo[j]fluoranthene</td>
<td>1.36</td>
<td>2.50</td>
<td>0.31</td>
<td>0.03</td>
<td>0.15</td>
<td>3.34</td>
<td>7.83</td>
</tr>
<tr>
<td>benzo[k]fluoranthene</td>
<td>0.14</td>
<td>2.10</td>
<td>0.48</td>
<td>0.03</td>
<td>0.15</td>
<td>2.60</td>
<td>4.66</td>
</tr>
<tr>
<td>2-methylnaphthalene</td>
<td>0.29</td>
<td>2.07</td>
<td>0.39</td>
<td>0.03</td>
<td>0.15</td>
<td>2.29</td>
<td>3.60</td>
</tr>
<tr>
<td>chrysene</td>
<td>1.23</td>
<td>2.25</td>
<td>0.70</td>
<td>0.03</td>
<td>0.15</td>
<td>2.97</td>
<td>6.09</td>
</tr>
<tr>
<td>benzo[e]pyrene</td>
<td>-0.93</td>
<td>2.08</td>
<td>0.30</td>
<td>0.03</td>
<td>0.15</td>
<td>2.24</td>
<td>3.43</td>
</tr>
<tr>
<td>benzo[a]anthracene</td>
<td>-0.62</td>
<td>2.08</td>
<td>0.48</td>
<td>0.03</td>
<td>0.15</td>
<td>2.37</td>
<td>3.86</td>
</tr>
<tr>
<td>benzo[ghi]perylene</td>
<td>0.88</td>
<td>2.06</td>
<td>0.24</td>
<td>0.03</td>
<td>0.15</td>
<td>2.33</td>
<td>3.72</td>
</tr>
<tr>
<td>indeno[1,2,3-cd]pyrene</td>
<td>0.63</td>
<td>2.06</td>
<td>0.25</td>
<td>0.03</td>
<td>0.15</td>
<td>2.33</td>
<td>3.72</td>
</tr>
<tr>
<td>benzo[a]pyrene</td>
<td>0.40</td>
<td>2.08</td>
<td>0.26</td>
<td>0.03</td>
<td>0.15</td>
<td>2.29</td>
<td>3.60</td>
</tr>
<tr>
<td>perylene</td>
<td>0.65</td>
<td>2.07</td>
<td>0.27</td>
<td>0.03</td>
<td>0.15</td>
<td>2.33</td>
<td>3.72</td>
</tr>
<tr>
<td>phenanthrene</td>
<td>1.18</td>
<td>2.06</td>
<td>0.39</td>
<td>0.03</td>
<td>0.15</td>
<td>2.43</td>
<td>4.05</td>
</tr>
<tr>
<td>benzo[b]fluoranthene</td>
<td>0.87</td>
<td>2.23</td>
<td>0.36</td>
<td>0.03</td>
<td>0.15</td>
<td>3.34</td>
<td>7.83</td>
</tr>
<tr>
<td>pyrene</td>
<td>-0.76</td>
<td>2.06</td>
<td>0.25</td>
<td>0.03</td>
<td>0.15</td>
<td>2.37</td>
<td>3.86</td>
</tr>
<tr>
<td>fluoranthene</td>
<td>-0.41</td>
<td>2.21</td>
<td>0.80</td>
<td>0.03</td>
<td>0.15</td>
<td>2.37</td>
<td>3.86</td>
</tr>
<tr>
<td>naphthalene</td>
<td>0.53</td>
<td>2.06</td>
<td>0.14</td>
<td>0.03</td>
<td>0.15</td>
<td>2.37</td>
<td>3.86</td>
</tr>
</tbody>
</table>

**Table 3.4.** Evaluation scores under constant method bias and constant method variance.

For each compound, $t_{score}$ = the estimated t score, $t_{sat}$ = the satisfactory critical value for the t score, $\sqrt{f_{score}}$ = the square root of the f score, $\sqrt{f_{queL}}$ = the square root of the lower questionable limit, $\sqrt{f_{satL}}$ = the square root of the satisfactory lower limit, $\sqrt{f_{satU}}$ = the square root of the satisfactory upper limit, $\sqrt{f_{queU}}$ = the square root of the questionable upper limit.
Figure 3.1. Example of an evaluation scores.

This plot shows the individual scores of the biases and within uncertainties for each compound from the PAHs in the Sediments 2003 exercise for one of the participants. The black dotted semicircles represent the approximate boundaries for satisfactory and unsatisfactory results.
3.1.3 Exporting the Model to the Study Data

The models found while working with the control data can be used to estimate the unknown parameters of the study data (Section 2.2). There are at least three ways the control data can be used to update the study data and a decision must be made regarding on this. These adjustments can be performed by each participant or by the coordinator of a multiple data source exercise event. It is important to specify clearly what the results are in order to use them correctly.

- Using the study data only: The simplest approach is to ignore the control data. The study data is not altered at all and it is reported as is.

- Using the control data to discriminate the study data: This option uses the control data only to detect outliers in the control data and assumes that the corresponding study data measurements are also outliers. The detected outliers are excluded from further calculations. In general this is the current strategy for the IEPOCME program.

- Using the control data for bias adjustment: This option requires us to assume we have enough information to model the bias. First we use the control data to estimate the bias and then we use the estimated bias to adjust the study data. In general a reduced bias or even unbiased estimator of the true value of the study data can be obtained. However the uncertainty of this estimate can be smaller or larger in comparison with the variance of the unadjusted estimates. This technique is detailed in the following section.

3.1.3.1 Parameter Estimation

Recall from Section 2.2 that $W_1, ..., W_M$ are the values obtained on M study samples. $W_m = \beta_0 + d + V_m + \delta_m$ with $V_m, \delta_m$ assumed to be independent error terms, $E[V_m] = 0$, $V[V_m] = \sigma^2_V$, $E[\delta_m] = 0$ and $V[\delta_m] = \sigma^2_\delta$, for $m = 1..M$. 

35
**Independent errors:**

First assume that the error terms are independent for the method measuring the control sample and the study sample; this implies $\text{Corr}[\epsilon_k, \delta_k] = \rho = 0$. Also assume $\sigma^2_\epsilon = \sigma^2_\delta$, i.e., the variance of the method is constant across the control and study samples. This can be due to the constancy of the method’s variance over different levels or to a method’s variance dependent on the concentration level but analyzing samples with similar levels.

Under independent error terms, the number of replicates of the control material and the number of replicates of the study material are not required to be equal.

The sample mean $\bar{W} = \frac{1}{M} \sum_{m=1}^{M} W_m$ is unbiased for $E[W_m] = \beta_0 + d$ and the sample variance $s^2_W = \frac{1}{M-1} \sum_{m=1}^{M} (W_m - \bar{W})^2$ is unbiased for $V[W_m] = \sigma^2 + \sigma^2_\epsilon$, since we assumed $\sigma^2_\delta = \sigma^2_\epsilon$. Let

\[
\hat{d} = \bar{W} - \hat{\beta}_0 \quad \text{and} \quad \hat{\sigma}^2_V = s^2_W - \hat{\sigma}^2_\epsilon
\]

then

\[
E[\hat{d}] = E[\bar{W} - \hat{\beta}_0] = \beta_0 + d - E[\hat{\beta}_0] = d, \quad (3.26)
\]

\[
E[\hat{\sigma}^2_V] = E[s^2_W - \hat{\sigma}^2_\epsilon] = V[W_m] - \sigma^2_\epsilon = \sigma^2_\epsilon \quad (3.27)
\]

and

\[
V[\hat{d}] = V[\bar{W} - \hat{\beta}_0] = V[\bar{W}] + V[\hat{\beta}_0] = E[\hat{\sigma}^2_V] \quad (3.28)
\]

where $\hat{\sigma}^2_d = \frac{s^2_W}{M} + \frac{s^2_\epsilon}{K} + \hat{\sigma}^2_\epsilon$. Then $(\hat{d}, \hat{\sigma}^2_d, \hat{\sigma}^2_V)$ are unbiased estimators for $(d, \sigma^2_d, \sigma^2_V)$ respectively. However, there is potential for both estimators for $d$ and $\sigma^2_V$ to become negative. Note that the assumption of $\sigma^2_\delta = \sigma^2_\epsilon$ is needed for estimating $\sigma^2_V$ but not for estimating $d$ nor for estimating $V[\hat{d}]$. 

36
Some additional distributional assumptions on the error terms are required in order to get confidence intervals for \(d\) and \(\sigma_V^2\). Assuming normality for \(V_k\) and \(\epsilon_k\), the distribution of \(\hat{d}\) is

\[
\frac{\hat{d} - d}{\hat{\sigma}_\hat{d}} \sim T(\nu_{\hat{d}})
\]

(3.29)

and an approximate 100\((1 - \alpha)\)% CI for \(d\) then becomes

\[
\left( \hat{d} + t_{(\alpha/2, \nu_{\hat{d}})} \hat{\sigma}_\hat{d}, \hat{d} + t_{(1-\alpha/2, \nu_{\hat{d}})} \hat{\sigma}_\hat{d} \right)
\]

(3.30)

where \(\nu_{\hat{d}}\) can be estimated with the Welch-Satterthwaite approximation

\[
\nu_{\hat{d}} = \frac{\left( \frac{s_Y^2}{K} + \frac{s_W^2}{M} + \hat{\sigma}_\epsilon^2 \right)^2}{\frac{s_Y^2}{K(K-1)} + \frac{s_W^2}{M(M-1)} + \frac{\hat{\sigma}_\epsilon^2}{\nu_j}}.
\]

(3.31)

When \(K\) and \(M\) are large the distribution is

\[
\frac{\hat{d} - d}{\hat{\sigma}_\hat{d}} \sim Z,
\]

(3.32)

with the approximate 100\((1 - \alpha)\)% CI

\[
\left( \hat{d} + z_{(\alpha/2)} \hat{\sigma}_\hat{d}, \hat{d} + z_{(1-\alpha/2)} \hat{\sigma}_\hat{d} \right).
\]

(3.33)

These confidence intervals can contain zero or even contain only negative values. In such a case the interpretation is that the true value of the compound in the study sample may be below the detection limit of the analytical method, i.e., a censored data point.

Similarly, the approximate distribution of \(\hat{\sigma}_V^2\) can be obtained from the exact distribution of \(s_W^2\)

\[
\frac{(M - 1)s_W^2}{\sigma_V^2 + \sigma_\delta^2} \sim \chi^2_{(M-1)}
\]

and, assuming \(\hat{\sigma}_\epsilon^2 = \sigma_\epsilon^2\), an approximate CI for \(\sigma_V^2\) can be obtained by
Similarly to (3.16), since a CI for $\hat{\sigma}_\epsilon^2$ is available via (3.15) or (3.16), say \( \text{CI}(\hat{\sigma}_\epsilon^2) = (lcl(\hat{\sigma}_\epsilon^2), ucl(\hat{\sigma}_\epsilon^2)) \), then (3.35) can be modified as
\[
\left( \frac{(M - 1)s_{W}^2}{\chi^2_{(1-\alpha/2, M-1)}} - lcl(\hat{\sigma}_\epsilon^2), \frac{(M - 1)s_{W}^2}{\chi^2_{(\alpha/2, M-1)}} - ucl(\hat{\sigma}_\epsilon^2) \right).
\]

(3.36)

There some cases where the probability of the estimate taking negative values is smaller than the probability of the estimate taking positive values. We can use
\[
Pr(\hat{\sigma}_V^2 < 0) = Pr\left(\chi^2_{(M-1)} < \frac{(M - 1)\hat{\sigma}_\epsilon^2}{s_{W}^2}\right)
\]

(3.37)
to estimate that probability.

If $\sigma_{\delta|D}^2 \neq \sigma_{\epsilon|X}^2$, as may be the case if the method’s variance depends on the concentration of the compound and the concentration of the compound in the control material and the study material are very different, then $\sigma_{\delta|D}^2$ and $\sigma_{\epsilon|X}^2$ should be modeled accordingly. In Chapter 5 we briefly review a variance model proportional to the square of the concentration.

Correlated errors:

Now assume that the error terms are correlated for the method while measuring the control sample and the study sample, so $K = M$ with pairing and $\text{Corr}[\epsilon_k, \delta_k] = \rho$. Also assume $\sigma_\epsilon^2 = \sigma_\delta^2$. Define the sample statistics $\bar{Y}, s_Y^2, \bar{W}, s_W^2$ as above, and define the sample covariance as $s_{YW} = \frac{1}{K-1} \left[ \sum_{k=1}^{K} (Y_kW_k) - KY\bar{W} \right]$. By using the method of moments it is straightforward to show that: $E[\bar{W}] = \beta_0 + d$, $E[s_{W}^2] = \sigma_V^2 + \sigma_\epsilon^2$, and $E[s_{YW}] = \rho \sigma_\epsilon \sigma_\delta = \rho \sigma_\epsilon^2$. Hence the point estimators defined as (3.24) and (3.25) are still unbiased. However $V[\hat{d}]$ changes due to the covariance term
\[
V[\hat{d}] = V[\bar{W} + \hat{x} - \bar{Y}] = V[\bar{Q} + \hat{x}] = V[\bar{Q}] + V[\hat{x}]
\]

(3.38)
where $Q_k = W_k - Y_k$ and $s_Q^2 = s_{W}^2 + s_{Y}^2 - 2s_{YW}$ is an unbiased estimator for $\sigma_Q^2 = V[\bar{Q}]$. 38
Some additional distributional assumptions on the error terms are required in order to get confidence intervals for \( d \) and \( \sigma^2_V \). Assuming normality of \( V_k \) and \( \epsilon_k \), the distribution and confidence interval of \( d \) is

\[
\frac{\hat{d} - d}{\sqrt{s_Q^2 + \hat{\sigma}_x^2}} \sim T(\nu_d)
\]

and an approximate 100(1 - \( \alpha \))% CI for \( d \) is

\[
(W - \hat{\beta}_0) \pm t(\alpha/2, \nu_d) \sqrt{s_Q^2 + \hat{\sigma}_x^2}
\]

where \( \nu_d \) is estimated with the Welch-Satterthwaite approximation

\[
\nu_d = \frac{2 \left( s_y^2 + s_Y^2 - 2s_{s_{WY}} \right)^2}{2\left( s_y^2 + s_Y^2 \right)^2} + \frac{1 + \frac{4\operatorname{Var}(s_{WY}) + 4\operatorname{Cov}(s_{Y}, s_{WY}) - 4\operatorname{Cov}(s_{Y}, s_{WY})}{K^2} + 2\hat{\sigma}_x^4}{\nu_y^2}
\]

\[
= \frac{s_Y^2 + s_Y^2 - 2s_{WY}}{\nu_y^2} + \frac{s_y^2 + s_Y^2 - 2s_{WY}}{K^2} + \frac{\hat{\sigma}_x^4}{\nu_y^2}
\]

When \( K \) is large the distribution and confidence interval of \( d \) is

\[
\frac{\hat{d} - d}{\sqrt{s_Q^2 + \hat{\sigma}_x^2}} \sim Z
\]

and an approximate CI for \( d \) becomes

\[
(W - \hat{\beta}_0) \pm z(\alpha/2) \sqrt{s_Q^2 + \hat{\sigma}_x^2}
\]

As with the point estimators, the distribution and approximate confidence interval of \( \sigma^2_V \) remain as stated by (3.34) and (3.35).

3.1.3.2 Examples

Table 3.5 contains the estimates of the adjusted study values and their variances by using the data in Table 3.1 and (3.24), (3.28), (3.30), (3.25), under the assumption
of independent errors and equal within-method variance for each compound across samples.

The adjusted values, the $d'$s, are similar to the observed values. However there are cases where $\hat{d}$ is rather different than $\bar{W}$. In eleven out of 26 cases the confidence interval for $\bar{W}$ does not contain the adjusted value $\hat{d}$. Note that relatively small differences in the control and reference values can lead to large differences in the raw study and adjusted study values when the levels of the control and the study materials are very different. Consider the naphthalene compound, where the control material level (873.7) and the reference value (848) are similar and the estimated bias is relatively small (25.7 about 3% of the control value) but considerably larger than the study material level (121.3), leading to an adjusted study value of (95.7 about 21% of the study value). The adjusted variances are in general larger than the observed variances.

The study material can be considered of homogeneous composition with respect to the majority of the compounds. However, almost 25% of the compounds (seven out of 26) appear to have a within-material variance. A probable explanation may be the assumption $\sigma_{\epsilon|x}^2 = \sigma_{\delta|d}^2$ is not satisfied (i.e., the variance of each compound can be very different across both samples).

Table 3.6 contains the adjusted study values by using the data in Table 3.1 and (3.24), (3.38) and (3.40), under the assumption of correlated errors.

Comparing the results with error assumed correlated (Table 3.6) to the results with error assumed independent (Table 3.5) we observe that the adjusted variances are in general similar to the unadjusted variances despite the large correlation coefficients, this is due to the relatively small covariance compared to the variance of the reference material $\sigma_x^2$. However a few estimated variances get bigger, in those cases the covariance term is significant when compared to the variance of the reference material. The variance of the estimated study value $\hat{d}$ tends to decrease as the
### Table 3.5. Example of study values and variances corrected for bias using control data when errors are independent.

<table>
<thead>
<tr>
<th>Compound</th>
<th>$\bar{W}$</th>
<th>$\bar{d}$</th>
<th>$\hat{\sigma}_d^2$</th>
<th>$\nu_d$</th>
<th>$lcl(\bar{d})$</th>
<th>$ucl(\bar{d})$</th>
<th>$\hat{\sigma}_V^2$</th>
<th>$lcl(\sigma^2_V)$</th>
<th>$ucl(\sigma^2_V)$</th>
<th>$Pr(\hat{\sigma}_V^2 &lt; 0)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2,3,5-trimethynaphthalene</td>
<td>25.7</td>
<td>28.5</td>
<td>7.2</td>
<td>32</td>
<td>23.0</td>
<td>34.0</td>
<td>0</td>
<td>0</td>
<td>11.1</td>
<td>0.972</td>
</tr>
<tr>
<td>acenaphthene</td>
<td>34.7</td>
<td>37.1</td>
<td>8.7</td>
<td>26</td>
<td>31.1</td>
<td>43.2</td>
<td>0</td>
<td>0</td>
<td>3.48</td>
<td>1</td>
</tr>
<tr>
<td>dibenzo[a,h]anthracene</td>
<td>112.0</td>
<td>109.0</td>
<td>32.0</td>
<td>12</td>
<td>96.4</td>
<td>121.0</td>
<td>16.9</td>
<td>4.22</td>
<td>684</td>
<td>0.0273</td>
</tr>
<tr>
<td>acenaphthylene</td>
<td>29.6</td>
<td>26.1</td>
<td>12.7</td>
<td>31</td>
<td>18.9</td>
<td>33.4</td>
<td>0</td>
<td>0</td>
<td>68.5</td>
<td>0.901</td>
</tr>
<tr>
<td>1-methylphenanthrene</td>
<td>69.4</td>
<td>70.8</td>
<td>9.3</td>
<td>27</td>
<td>64.5</td>
<td>77.1</td>
<td>0</td>
<td>0</td>
<td>15</td>
<td>0.668</td>
</tr>
<tr>
<td>biphenyl</td>
<td>29.2</td>
<td>33.2</td>
<td>18.2</td>
<td>30</td>
<td>24.5</td>
<td>41.9</td>
<td>0</td>
<td>0</td>
<td>15.8</td>
<td>1</td>
</tr>
<tr>
<td>2,6-dimethynaphthalene</td>
<td>35.2</td>
<td>37.1</td>
<td>6.6</td>
<td>21</td>
<td>31.7</td>
<td>42.4</td>
<td>0</td>
<td>0</td>
<td>4.39</td>
<td>1</td>
</tr>
<tr>
<td>fluorene</td>
<td>65.5</td>
<td>64.9</td>
<td>63.0</td>
<td>41</td>
<td>48.9</td>
<td>80.9</td>
<td>0</td>
<td>0</td>
<td>6.23</td>
<td>1</td>
</tr>
<tr>
<td>triphenylene</td>
<td>141.0</td>
<td>138.0</td>
<td>33.0</td>
<td>4</td>
<td>121.0</td>
<td>155.0</td>
<td>65.3</td>
<td>12.6</td>
<td>2850</td>
<td>0.0922</td>
</tr>
<tr>
<td>1-methylphenanthalene</td>
<td>49.0</td>
<td>55.3</td>
<td>65.4</td>
<td>23</td>
<td>38.6</td>
<td>72.0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>anthracene</td>
<td>172.0</td>
<td>175.0</td>
<td>98.6</td>
<td>38</td>
<td>155.0</td>
<td>195.0</td>
<td>0</td>
<td>0</td>
<td>214</td>
<td>0.997</td>
</tr>
<tr>
<td>benzo[j]fluoranthene</td>
<td>281.0</td>
<td>277.0</td>
<td>31.4</td>
<td>3</td>
<td>260.0</td>
<td>294.0</td>
<td>70</td>
<td>17.3</td>
<td>2850</td>
<td>0.0317</td>
</tr>
<tr>
<td>benzo[k]fluoranthene</td>
<td>406.0</td>
<td>405.0</td>
<td>118.0</td>
<td>21</td>
<td>382.0</td>
<td>427.0</td>
<td>50.7</td>
<td>0</td>
<td>2980</td>
<td>0.283</td>
</tr>
<tr>
<td>2-methylnaphthalene</td>
<td>79.2</td>
<td>71.2</td>
<td>778.0</td>
<td>34</td>
<td>14.6</td>
<td>128.0</td>
<td>0</td>
<td>0</td>
<td>410</td>
<td>1</td>
</tr>
<tr>
<td>chrysene</td>
<td>547.0</td>
<td>526.0</td>
<td>307.0</td>
<td>11</td>
<td>487.0</td>
<td>564.0</td>
<td>0</td>
<td>0</td>
<td>957</td>
<td>0.703</td>
</tr>
<tr>
<td>benzo[ghi]perylene</td>
<td>488.0</td>
<td>468.0</td>
<td>546.0</td>
<td>38</td>
<td>420.0</td>
<td>515.0</td>
<td>0</td>
<td>0</td>
<td>957</td>
<td>0.703</td>
</tr>
<tr>
<td>benzo[e]pyrene</td>
<td>590.0</td>
<td>602.0</td>
<td>226.0</td>
<td>23</td>
<td>571.0</td>
<td>633.0</td>
<td>100</td>
<td>0</td>
<td>5660</td>
<td>0.264</td>
</tr>
<tr>
<td>benzo[a]anthracene</td>
<td>378.0</td>
<td>387.0</td>
<td>201.0</td>
<td>33</td>
<td>358.0</td>
<td>415.0</td>
<td>0</td>
<td>0</td>
<td>548</td>
<td>0.997</td>
</tr>
<tr>
<td>indeno[1,2,3-cd]pyrene</td>
<td>608.0</td>
<td>589.0</td>
<td>889.0</td>
<td>39</td>
<td>529.0</td>
<td>650.0</td>
<td>0</td>
<td>0</td>
<td>2420</td>
<td>0.645</td>
</tr>
<tr>
<td>benzo[a]pyrene</td>
<td>636.0</td>
<td>632.0</td>
<td>94.5</td>
<td>32</td>
<td>613.0</td>
<td>652.0</td>
<td>7</td>
<td>0</td>
<td>1250</td>
<td>0.543</td>
</tr>
<tr>
<td>perylene</td>
<td>184.0</td>
<td>169.0</td>
<td>574.0</td>
<td>36</td>
<td>120.0</td>
<td>217.0</td>
<td>0</td>
<td>0</td>
<td>1840</td>
<td>0.848</td>
</tr>
<tr>
<td>phenanthrene</td>
<td>539.0</td>
<td>512.0</td>
<td>554.0</td>
<td>43</td>
<td>464.0</td>
<td>559.0</td>
<td>0</td>
<td>0</td>
<td>1840</td>
<td>0.871</td>
</tr>
<tr>
<td>benzo[b]fluoranthene</td>
<td>1090.0</td>
<td>1080.0</td>
<td>258.0</td>
<td>7</td>
<td>1040.0</td>
<td>1120.0</td>
<td>332</td>
<td>55.1</td>
<td>15000</td>
<td>0.119</td>
</tr>
<tr>
<td>pyrene</td>
<td>1130.0</td>
<td>1140.0</td>
<td>494.0</td>
<td>36</td>
<td>1100.0</td>
<td>1190.0</td>
<td>84.7</td>
<td>0</td>
<td>7360</td>
<td>0.424</td>
</tr>
<tr>
<td>fluoranthene</td>
<td>1090.0</td>
<td>1110.0</td>
<td>1050.0</td>
<td>14</td>
<td>1040.0</td>
<td>1180.0</td>
<td>0</td>
<td>0</td>
<td>2560</td>
<td>1</td>
</tr>
<tr>
<td>naphthalene</td>
<td>121.0</td>
<td>95.7</td>
<td>2370.0</td>
<td>39</td>
<td>-2.9</td>
<td>194.0</td>
<td>0</td>
<td>0</td>
<td>185</td>
<td>1</td>
</tr>
</tbody>
</table>

For each compound, $\bar{W}$ = the observed mean study value, $\bar{d}$ = the corrected mean of the study value, $\hat{\sigma}_d^2$ = the variance of the corrected study value, $\nu_d$ = the degrees of freedom of the variance of the corrected study value, $lcl(\bar{d})$ = a lower 95% confidence limit for the corrected study value, $ucl(\bar{d})$ = an upper 95% confidence limit for the corrected study value, $\hat{\sigma}_V^2$ = estimated within study material variance, $lcl(\sigma^2_V)$ = a lower 95% confidence limit for the within study material variance, $ucl(\sigma^2_V)$ = an upper 95% confidence limit for the within study material variance and $Pr(\hat{\sigma}_V^2 < 0)$ = estimated probability of getting a negative estimate of the within study material variance under the assumption of independent errors.
<table>
<thead>
<tr>
<th>Compound</th>
<th>$\hat{\rho}_{YW}$</th>
<th>$\hat{d}$</th>
<th>$\hat{\sigma}_{d}^2$</th>
<th>$\nu_{\hat{d}}$</th>
<th>$lcl(\hat{d})$</th>
<th>$ucl(\hat{d})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2,3,5-trimethylnaphthalene</td>
<td>0.69</td>
<td>28.5</td>
<td>7.0</td>
<td>30</td>
<td>23.1</td>
<td>33.9</td>
</tr>
<tr>
<td>acenaphthene</td>
<td>-0.88</td>
<td>37.1</td>
<td>9.3</td>
<td>20</td>
<td>30.8</td>
<td>43.5</td>
</tr>
<tr>
<td>dibenz[a,h]anthracene</td>
<td>0.28</td>
<td>109.0</td>
<td>31.4</td>
<td>12</td>
<td>96.5</td>
<td>121.0</td>
</tr>
<tr>
<td>acenaphthylene</td>
<td>-0.52</td>
<td>26.1</td>
<td>13.7</td>
<td>22</td>
<td>18.5</td>
<td>33.8</td>
</tr>
<tr>
<td>1-methylphenanthrene</td>
<td>0.92</td>
<td>70.8</td>
<td>9.1</td>
<td>26</td>
<td>64.6</td>
<td>77.0</td>
</tr>
<tr>
<td>biphenyl</td>
<td>-1.00</td>
<td>33.2</td>
<td>19.2</td>
<td>28</td>
<td>24.3</td>
<td>42.2</td>
</tr>
<tr>
<td>2,6-dimethylnaphthalene</td>
<td>0.52</td>
<td>37.1</td>
<td>6.3</td>
<td>23</td>
<td>31.9</td>
<td>42.3</td>
</tr>
<tr>
<td>fluorene</td>
<td>-0.95</td>
<td>64.9</td>
<td>64.5</td>
<td>39</td>
<td>48.7</td>
<td>81.1</td>
</tr>
<tr>
<td>triphenylene</td>
<td>0.49</td>
<td>138.0</td>
<td>25.6</td>
<td>3</td>
<td>123.0</td>
<td>153.0</td>
</tr>
<tr>
<td>1-methylnaphthalene</td>
<td>-0.43</td>
<td>55.3</td>
<td>67.1</td>
<td>21</td>
<td>38.3</td>
<td>72.3</td>
</tr>
<tr>
<td>anthracene</td>
<td>1.00</td>
<td>175.0</td>
<td>88.5</td>
<td>41</td>
<td>156.0</td>
<td>194.0</td>
</tr>
<tr>
<td>benzo[j]fluoranthene</td>
<td>-0.01</td>
<td>277.0</td>
<td>31.5</td>
<td>3</td>
<td>260.0</td>
<td>295.0</td>
</tr>
<tr>
<td>benzo[k]fluoranthene</td>
<td>-0.87</td>
<td>405.0</td>
<td>143.0</td>
<td>10</td>
<td>378.0</td>
<td>431.0</td>
</tr>
<tr>
<td>2-methylnaphthalene</td>
<td>-0.57</td>
<td>71.2</td>
<td>781.0</td>
<td>34</td>
<td>14.4</td>
<td>128.0</td>
</tr>
<tr>
<td>chrysene</td>
<td>-0.97</td>
<td>526.0</td>
<td>338.0</td>
<td>11</td>
<td>485.0</td>
<td>566.0</td>
</tr>
<tr>
<td>benzo[ghi]perylene</td>
<td>0.45</td>
<td>468.0</td>
<td>537.0</td>
<td>37</td>
<td>421.0</td>
<td>515.0</td>
</tr>
<tr>
<td>benzo[e]pyrene</td>
<td>0.92</td>
<td>602.0</td>
<td>176.0</td>
<td>30</td>
<td>575.0</td>
<td>629.0</td>
</tr>
<tr>
<td>benzo[a]anthracene</td>
<td>0.96</td>
<td>387.0</td>
<td>175.0</td>
<td>41</td>
<td>360.0</td>
<td>413.0</td>
</tr>
<tr>
<td>inden[1,2,3-cd]pyrene</td>
<td>-0.98</td>
<td>589.0</td>
<td>931.0</td>
<td>37</td>
<td>528.0</td>
<td>651.0</td>
</tr>
<tr>
<td>benzo[a]pyrene</td>
<td>0.99</td>
<td>632.0</td>
<td>75.6</td>
<td>30</td>
<td>615.0</td>
<td>650.0</td>
</tr>
<tr>
<td>perylene</td>
<td>0.68</td>
<td>169.0</td>
<td>544.0</td>
<td>34</td>
<td>121.0</td>
<td>216.0</td>
</tr>
<tr>
<td>phenanthrene</td>
<td>-0.76</td>
<td>512.0</td>
<td>589.0</td>
<td>34</td>
<td>462.0</td>
<td>561.0</td>
</tr>
<tr>
<td>benzo[b]fluoranthene</td>
<td>-0.61</td>
<td>1080.0</td>
<td>312.0</td>
<td>5</td>
<td>1030.0</td>
<td>1130.0</td>
</tr>
<tr>
<td>pyrene</td>
<td>-0.65</td>
<td>1140.0</td>
<td>555.0</td>
<td>18</td>
<td>1100.0</td>
<td>1190.0</td>
</tr>
<tr>
<td>fluoranthene</td>
<td>-0.38</td>
<td>1110.0</td>
<td>1140.0</td>
<td>10</td>
<td>1030.0</td>
<td>1180.0</td>
</tr>
<tr>
<td>naphthalene</td>
<td>-0.26</td>
<td>95.7</td>
<td>2380.0</td>
<td>39</td>
<td>-3.0</td>
<td>194.0</td>
</tr>
</tbody>
</table>

Table 3.6. Example of study values corrected for bias using control data when errors are correlated.

For each compound, $\hat{\rho}_{YW} =$ correlation coefficient, $\hat{d} =$ the corrected mean of the study value, $\hat{\sigma}_{\hat{d}}^2 =$ estimated variance of the corrected study value, $\nu_{\hat{d}} =$ the degrees of freedom of the variance of the corrected study value, $lcl(\hat{d}) =$ a lower 95% confidence limit for the corrected study value, $ucl(\hat{d}) =$ an upper 95% confidence limit for the corrected study value under the assumption of correlated errors.

correlation coefficient increases, as a consequence the estimated confidence interval tends to get narrower as the correlation coefficient increases. The estimated values $\hat{d}$ are unaffected by the possible covariance.
3.2 Multiple laboratories, one compound at one point in time

Now assume we have a group of $J$ participants measuring the same compound. Using Section 3.1 we can obtain estimates of each participating method’s bias and variance for the specific compound. We want to estimate the true value for each compound in the study material among all the participants, so we are now interested on pooling information across different participants about the same compound.

3.2.1 The Basic Model

The model for the observed concentration of the particular compound in the control material is

$$Y_{jk} = \beta_{0j} + x_j + U_{jk} + \epsilon_{jk}, \ j = 1..J, \ k = 1..K_j$$

(3.44)

where $U_{jk} \sim N(0, \sigma^2_U)$ and $\epsilon_{jk} \sim N(0, \sigma^2_{\epsilon_j})$ are independent error terms. Note that $x_j$ can change with the laboratory; that is, the model allows different CRMs to be allocated to the laboratories.

The model for the observed concentration of the particular compound in the study material is

$$W_{jm} = \beta_{0j} + d + V_{jm} + \delta_{jm}, \ j = 1..J, \ m = 1..M_j$$

(3.45)

where $V_{jm} \sim N(0, \sigma^2_V)$ and $\delta_{jm} \sim N(0, \sigma^2_{\delta_j})$ are independent error terms and the $\beta_{0j}$ are the same unknown constants as in the control material model.

3.2.2 Working with the Study Data Only

Most of the intercomparison literature work is about using the study sample data only. The common assumption is that the reported study values are unbiased with no control material as a reference. No evidence is available to support or to reject this assumption, so automatically the unbiased model must be used for all the methods.

There are several guidelines (ISO 9001, ISO Guide 43-1, ISO/IEC 17025 series) describing the way the multiple source data should be analyzed. The general approach
includes three steps in common: (i) obtaining an assigned value for the unknown material (consensus value), (ii) obtaining the performance statistics, and (iii) evaluating the performance of the participants.

### 3.2.2.1 Parameter Estimation

**Previous Work**

In order to obtain an assigned consensus value several statistical methods can be used. These methods range from using the plain mean and standard deviation to the use of some robust statistics such as the median and mean absolute deviation (mad), trimmed mean or some kind of windsorized mean (Huber (1981) [26]), the weighted mean (Graybill and Deal (1959) [23]), or the more elaborated MLE for the one way random effects model (Rukhin and Vangel (1998) [46], Rukhin (2007) [44] Rukhin (2009) [45]) and some approximations to the MLE (Mandel and Paule (1970) [37], Paule and Mandel (1982) [41], DerSimonian and Laird (1986) [17]).

The problem of summarizing data on measurements made on virtually the same quantity by different individuals is of particular importance to evaluating and comparing test methods, characterizing materials, or evaluating the individuals/laboratories themselves. These cover a wide range on scientific, engineering and international trading aspects. It has been addressed by several researchers.

Cochran (1937) [14] considered this problem for the first time. He investigated Maximum Likelihood Estimates (MLE) for the one-way random effect ANOVA model under balanced design and heteroscedasticity. Meier (1953) [40] studied the approximation to the limiting variance of the weighted mean when the number of individuals/laboratories is small but with no between variance, so this is not a one-way random effect ANOVA. Cochran and Carrol (1953) [13] studied the efficiency of the weighted mean under balanced design and both small sample size and small number of participants.
Cochran (1954) [12], extended to the unbalanced scenario and the heteroscedasticity situation with \( W_{jm} = d + \delta_{jm}, j = 1..J, m = 1..M_j, \delta_{jm} \sim N(0, \sigma_j^2) \). He also stated a first approximation to the random effects model (which he refers to as model with interactions) when random bias terms are involved: \( W_{jm} = d_j + \delta_{jm}, j = 1..J, m = 1..M_j, \delta_{jm} \sim N(0, \sigma_j^2), d_j = d + (d_j - d) = d + b_j \), where \( b_j \) is a random bias.

Graybill and Deal (1959) [23] studied the combination of unbiased estimators of the common mean by using the inverse of the sample variance as weights.

\[
\tilde{x}_{GD} = \frac{\sum x_i s_i^{-2}}{\sum s_i^{-2}}.
\]  

(3.46)

This was revisited by Mandel and Paule (1970) [37]. Rao et al (1981) [42] reported a numerical investigation of a set of estimators, however they did not consider MLE due to computational difficulty. Paule and Mandel (1982) [41] presented a weighting process to estimate iteratively the mean from multiple experiments. Rukhin and Vangel (1998) [46] studied the common mean model and the weighted means model and compared their statistics. DerSimonian and Laird (1986) [17] discussed the random effects model and showed a noniterative procedure to combine evidence from a set of experiments. Zhang (2006) [52], studied the variance associated with the weighted mean of the common mean model: \( W_{jm} = d + \alpha_j + \delta_{jm}, j = 1..J, m = 1..M, \delta_{jm} \sim N(0, \sigma_j^2), \alpha_j \sim N(0, \sigma^2_\alpha) \).

Vangel and Rukhin (1999) [51] reviewed these works and studied the MLE for heteroscedastic one-way random effects (allowing for random bias terms), both theoretical and empirically. They proved that the interative solution always exists (Theorem 1) and also stated the parameter bounds for \( d \) and \( \sigma_\alpha \) to look for the MLE numerically. Multiple real roots of the cubic equation in(3.49) in the interval \([0,1]\) suggests the \( j^{th} \) laboratory is an outlier, and gave a necessary condition for multiple weights.
Dempster et al (1977) [16] reviewed the problem on estimating the mean and variance when data is censored or truncated from a normal distribution using the EM algorithm.

Eberhardt et al (1989) [19] considered a model with bounded random bias; that is, the support for the cdf of the random bias is an interval \([m_i, M_i]\) where \(m_i\) and \(M_i\) are known finite constants. In this case they use the minimax approach. Iyer et al (2004) [32] studied via simulation the interval estimates for the unbounded bias model, the bounded model, and the ISO GUM type models, especially for small number of participants. Rukhin (2007) [44] compared the different approaches to estimate the mean from multiple sources and Rukhin (2009) [45] studied the metrological implications of the different models and studied a meta-model which included explicitly the participant bias as fixed effects or random effects, i.e., \(W_{jm} = d + r_j + \delta_{jm}, j = 1...J, m = 1...M_j, \delta_{jm} \sim N(0, \sigma^2_j), r_j \text{ fixed constants or } r_j \sim N(0, \sigma^2_{rj}).\)

For the purpose of reviewing the use of the maximum likelihood estimators, consider the one-way random effects model with unequal variances applied to the study data. Assuming all the laboratories received the same study material with true value \(d\) and normality of the error terms,

\[
W_{jm} = d + B_{0j} + V_{jm} + \delta_{jm}, j = 1,...,J, m = 1,...,M_j, V_{jm} \sim N(0, \sigma^2_{Vj}), \delta_{jm} \sim N(0, \sigma^2_{\delta j}), B_{0j} \sim N(0, \sigma^2_{B_{0j}}), \text{ where all random terms are independent, } (\sigma^2_{Vj} + \sigma^2_{\delta j}) \text{ are the within-method variance (the components due to heterogeneity and the variance of the error in the equation are confounded) and are allowed to change for each method, and } \sigma^2_{B_{0j}} \text{ is the between method variance. This model treats the methods’ responses as random with a random effect attached to the bias of each one of them.}
\]

Then the log-likelihood function becomes
\[
\ell(\theta|W) = \sum_{j=1}^{J} \log(f(\bar{W}_j, s_{Wj}^2, M_j|d, \sigma_{B0}^2, \sigma_{Vj}^2 + \sigma_{\delta j}^2)) \\
= -\frac{N}{2} \log(2\pi) - \frac{1}{2} \sum (M_j - 1) \log(\sigma_{Vj}^2 + \sigma_{\delta j}^2) - \frac{1}{2} \sum \frac{M_j - 1}{\sigma_{Vj}^2 + \sigma_{\delta j}^2} s_{Wj}^2 \\
- \frac{1}{2} \sum \log(\sigma_{Vj}^2 + \sigma_{\delta j}^2 + \sigma_{B0}^2) - \frac{1}{2} \sum \frac{M_j(\bar{W}_j - d)^2}{\sigma_{Vj}^2 + \sigma_{\delta j}^2 + \sigma_{B0}^2} 
\tag{3.47}
\]

where \( \theta = (d, \sigma_{B0}^2, \sigma_{Vj}^2 + \sigma_{\delta j}^2) \) is the vector of \( J+2 \) parameters and \( N = \sum_j M_j \).

The score equations are

\[
S(\theta, w) = \begin{bmatrix}
\frac{\partial}{\partial d} \ell(\theta) \\
\frac{\partial}{\partial \sigma_{B0}^2} \ell(\theta) \\
\frac{\partial}{\partial (\sigma_{Vj}^2 + \sigma_{\delta j}^2)} \ell(\theta)
\end{bmatrix} = 0 
\tag{3.48}
\]

and the system of \( J+2 \) equations becomes after some algebra and using Vangel and Rukhin (1999) [51] notation

\[
\hat{d} = \sum_\gamma \frac{\gamma_j}{\gamma_j} \bar{W}_j, \\
\hat{\sigma}_{B0}^2 = \frac{\sum \gamma_j^2 (\bar{W}_j - \hat{d})^2}{\sum \gamma_j} \quad \text{and} \quad \gamma_j = \frac{M_j \hat{\sigma}_{B0}^2}{\sigma_{Vj}^2 + \sigma_{\delta j}^2 + \sigma_{B0}^2} \\
0 = \gamma_j^3 - (a_j + 2) \gamma_j^2 + ((M_j + 1) a_j + (M_j - 1) b_j + 1) \gamma_j - M_j a_j, 
\tag{3.49}
\]

where \( \gamma_j = \frac{M_j \sigma_{B0}^2}{\sigma_{Vj}^2 + \sigma_{\delta j}^2 + \sigma_{B0}^2}, \quad a_j = \frac{\sigma_{B0}^2}{(\bar{W}_j - d)^2}, \quad b_j = \frac{s_{Wj}^2/M_j}{(\bar{W}_j - d)^2} \) and \( j = 1, \ldots, J. \)

The equations are non-linear in the parameters, so an iterative approach is required. The Modifying Estimating Equations (MEE) approach is used to fit this non-linear model iteratively. The variance of the estimator of the true value \( \hat{d} \) can be obtained by applying the law of total variance (Casella and Berger (2002) [11])

\[
\sigma_d^2 = V[\hat{d}] = V \left[ \sum_\gamma \frac{\gamma_j \bar{W}_j}{\sum \gamma_j} \right] \\
= E \left[ V \left[ \sum_\gamma \frac{\gamma_j \bar{W}_j}{\sum \gamma_j} \right] \right] + E \left[ \sum_\gamma \frac{\gamma_j \bar{W}_j}{\sum \gamma_j} \right] \\
= E \left[ \sum_\gamma \frac{\gamma_j^2 V[\bar{W}_j]}{\sum \gamma_j^2} \right] = E \left[ \sum_\gamma \frac{\gamma_j^2 \gamma_j^{-1} \sigma_{B0}^2}{\sum \gamma_j^2} \right], 
\tag{3.50}
\]
and an estimator is given by
\[ \hat{\sigma}_d^2 = \frac{\hat{\sigma}_{\hat{d}_0}^2}{\sum \hat{\gamma}_j}. \] (3.51)

### 3.2.3 Working with the Control and Study Data

Using the results in the previous Section 3.1 we can obtain unbiased estimates for \( \beta_{0j} \) and \( \sigma_{\epsilon_j}^2 \). The confidence interval estimates for \( \beta_{0j} \) may give evidence of the presence of fixed effects and the laboratories should be considered as fixed. Consequentially the one way random effects model may not be applicable. If we use the MLE approach and the true underlying model is one of fixed effects then

\[
E[\hat{d}] = E\left[ \sum \frac{\hat{\gamma}_j W_j}{\sum \hat{\gamma}_j} \right] = E\left[ \sum \frac{\hat{\gamma}_j (d + \beta_{0j} + \bar{V}_j + \delta_j)}{\sum \hat{\gamma}_j} \right]
\]

(3.52)

Hence \( \hat{d} \) is unbiased for \( d \) if and only if all the participating methods are unbiased. This imposes a strong assumption that is not always met. The second term in (3.52) is responsible for a potential push-pull effect and if present it is propagated to the users of these estimators.

A more natural approach is to adjust the study values for bias and then obtain the simple average or weighted average of the study values corrected for bias. In this section we obtained \( J \) unbiased estimators for \( d \), let’s say \( \hat{d}_j \) for laboratory \( j \). These are the observed values adjusted for bias.

Let us define the adjusted study values for each laboratory after correcting for bias as \( \hat{d}_j = \bar{W}_j - \hat{\beta}_{0j} \), with

\[
\hat{\sigma}_{\hat{d}_j}^2 = s_{\bar{Q}}^2 + \hat{\sigma}_{\hat{\gamma}_j}^2,
\]

\[
s_{\bar{Q}j}^2 = \begin{cases} 
\frac{s_{Wj}^2}{M_j} + \frac{s_{\gamma_j}^2}{K_j} & \text{if errors in the equation are independent,} \\
\frac{s_{Wj}^2 + s_{\gamma_j}^2 - 2s_{WYj}}{K_j} & \text{if errors in the equation are correlated,}
\end{cases}
\] (3.53)
and \( \nu_{\hat{d}_j} = dof(\hat{\sigma}_{\hat{d}_j}^2) \). It is straightforward to show that each \( \hat{d}_j \) is unbiased for \( d \) as shown in (3.26).

Note that the adjusted study values are correlated if the laboratories are using the same CRM’s for reference value purposes, as is often the case.

Let us define the covariance of the adjusted study values. Let \( R \) different CRM’s be used as control materials. Define a matrix \( A \) coding the information about which laboratory is using what CRM. Let \( A \) be a \( J \times R \) matrix with entries defined as

\[
a_{jr} = \begin{cases} 
1, & \text{if the } j^{th} \text{ laboratory uses the } r^{th} \text{ CRM,} \\
0, & \text{otherwise}
\end{cases}
\]  

with the restriction \( \sum_{r=1}^{R} a_{jr} = 1, \forall j \); that is, one laboratory uses just one CRM. Let \( \Sigma_{\hat{x}} = diag(\sigma_{\hat{x}1}^2, \ldots, \sigma_{\hat{x}R}^2) \) be the diagonal \( R \times R \) matrix of the variances of the different reference values. Then the covariance of \( \hat{d} \) is

\[
\text{Cov}[\hat{d}, \hat{d}] = \Sigma_{\hat{Q}} + A\Sigma_{\hat{x}}A',
\]

\[
\Sigma_{\hat{Q}} = diag(\sigma_{\hat{Q}1}^2, \ldots, \sigma_{\hat{Q}J}^2),
\]

where \( \sigma_{\hat{Q}j}^2 = E[s_{\hat{Q}j}^2], j = 1, \ldots, J. \)

Simple average

Define the estimator of the study value as the simple average of the adjusted study values

\[
\bar{d} = \frac{1}{J} \sum_{j=1}^{J} \hat{d}_j = \frac{1}{J} 1' \hat{d}.
\]  

(3.56)

where \( 1 \) = vector of length \( J \) with all its entries equal 1.

Hence
\[ E[\tilde{d}] = E \left[ \frac{1}{J} \hat{1}' \tilde{d} \right] = \frac{1}{J} \hat{1}' E[\tilde{d}] = d \] and
\[ V[\tilde{d}] = V \left[ \frac{1}{J} \hat{1}' \tilde{d} \right] = \frac{1}{J^2} \hat{1}' V[\hat{d}] \hat{1} = \frac{1}{J^2} \hat{1}' (\Sigma_{\hat{Q}} + A \Sigma_x A') \hat{1} \] (3.57)
\[ = \frac{1}{J^2} \left( \sum_{j=1}^{J} \sigma_{Qj}^2 + \sum_{r=1}^{R} n_r^2 \sigma_{xr}^2 \right) \]

where \([n_1, \ldots, n_R]' = 1' A\) is a vector with the number of laboratories using the same CRM.

The simple mean of the adjusted study values \(\tilde{d}\) is an unbiased estimator for \(d\) and an unbiased estimator for its variance is
\[ \hat{V}[\tilde{d}] = \frac{1}{J^2} \left( \sum_{j=1}^{J} \bar{s}_{Qj}^2 + \sum_{r=1}^{R} n_r^2 \hat{\sigma}_{xr}^2 \right). \] (3.58)

**Weighted average**

Define the estimator of the study value as the weighted average of the adjusted study values. If the true variances are known the weights are \(\omega_j = \sigma_{d_j}^{-2} / \sum_j \sigma_{d_j}^{-2}\). However it is often the case that the true variances are unknown, instead we use estimated weights \(\hat{\omega}_j = \hat{\sigma}_{d_j}^{-2} / \sum_j \hat{\sigma}_{d_j}^{-2}\) and \(\sum_j \hat{\omega}_j = \hat{\omega}' 1 = 1\).

Define the weighted estimator as
\[ \bar{d}_\omega = \sum_{j=1}^{J} \hat{\omega}_j \hat{d}_j = \hat{\omega}' \hat{d}. \] (3.59)

If the weights \(\hat{\omega}_j\) are independent of \(\hat{d}_j\) then it is straightforward to show that \(\bar{d}_\omega\) is unbiased for \(d\), applying the double expectation theorem
\[ E[\bar{d}_\omega] = E \left[ \hat{\omega}' \hat{d} \right] = E \left[ E \left[ \hat{\omega}' \hat{d} | \hat{\omega} \right] \right] = E \left[ \hat{\omega}' E[\hat{d} | \hat{\omega}] \right] = E \left[ \hat{\omega}' E[1d] \right] = E \left[ \hat{\omega}' 1d \right] = E[d] = d. \] (3.60)
The variance of \( \bar{d}_\omega \) can be obtained assuming independence of the weights \( \hat{\omega}_j \) and \( \hat{d}_j \), applying the double expectation theorem, the conditional variance theorem and the expected value of a quadratic form

\[
V[\bar{d}_\omega] = V\left[ \hat{\omega}' \hat{d} \right] = V \left[ E \left[ \hat{\omega}' \hat{d} | \hat{\omega} \right] \right] + E \left[ V \left[ \hat{\omega}' \hat{d} | \hat{\omega} \right] \right]
\]

\[
= V \left[ d \right] + E \left[ \hat{\omega}' V \left[ \hat{d} | \hat{\omega} \right] \right] = E \left[ \hat{\omega}' (\Sigma Q + A \Sigma x A') \hat{\omega} \right] + tr \left( (\Sigma Q + A \Sigma x A') \Sigma_\omega \right). \tag{3.61}
\]

Although (3.61) is the exact variance, we need estimators for \( E[\hat{\omega}] \) and \( \Sigma_\hat{\omega} \), which involve ratios of random variables. Instead it is often assumed that the estimated weights \( \hat{\omega} \) are fixed. Under this assumption \( \Sigma_\hat{\omega} = 0 \) and the variance of the weighted average of the adjusted study values can be estimated as

\[
\hat{V}[\bar{d}_\omega] = \hat{\omega}' (\hat{\Sigma} Q + A \hat{\Sigma} x A') \hat{\omega}
\]

\[
= \sum_{j=1}^{J} \hat{\omega}_j^2 \hat{s}_{Qj}^2 + \sum_{r=1}^{R} \left( \sum_{j=1}^{J} \hat{\omega}_j a_{jr} \right)^2 \hat{\sigma}_{xr}^2. \tag{3.62}
\]

Similarly we can use (3.25) to obtain \( J \) unbiased estimators for the heterogeneity variance of the study material, say \( \hat{\sigma}_{Vj}^2 \). Assuming normality and equal variances of the error in the equation for each laboratory \( (\sigma_{\delta j}^2 = \sigma_{\epsilon j}^2) \), we have

\[
\sum_{j=1}^{J} \left( \nu_{Vj} \hat{\sigma}_{Vj}^2 \right) \sim \chi^2 \left( \sum_{j=1}^{J} \nu_{Vj} \right), \tag{3.63}
\]

where
\[ \nu_{Vj} = \begin{cases} \frac{(s_{Wj}^2-s_{Yj}^2+\hat{\sigma}_{Uj}^2)^2}{s_{Wj}^2+s_{Yj}^2+\hat{\sigma}_{Uj}^2} & \text{if the errors in the equation are independent,} \\
\frac{(s_{Wj}^2-s_{Yj}^2+\hat{\sigma}_{Uj}^2)^2}{s_{Wj}^2+s_{Yj}^2+\hat{\sigma}_{Uj}^2} & \text{if the errors in the equation are correlated.} \end{cases} \]

The \( \sigma^2_v \) can be estimated as

\[ \hat{\sigma}_v^2 = \frac{\sum_{j=1}^{J} \nu_{Vj} \hat{\sigma}_{Vj}^2}{\nu} \quad (3.65) \]

where

\[ \nu_v = \begin{cases} \frac{\left( \sum_{j} \nu_{Vj}(s_{Wj}^2-s_{Yj}^2+\hat{\sigma}_{Uj}^2) \right)^2}{\sum_{j} \nu_{Vj} \left( s_{Wj}^2+s_{Yj}^2+\hat{\sigma}_{Uj}^2 \right) + \sum_r \left( \sum_{j} a_{jr} \nu_{Vj} \right)^2 \frac{\hat{\sigma}_{Uj}^2}{\nu_{Ur}}} & \text{if } s_{YWj} = 0, \\
\frac{\left( \sum_{j} \nu_{Vj}(s_{Wj}^2-s_{Yj}^2+\hat{\sigma}_{Uj}^2) \right)^2}{\sum_{j} \nu_{Vj} \left( s_{Wj}^2+s_{Yj}^2+\hat{\sigma}_{Uj}^2 \right) + \sum_r \left( \sum_{j} a_{jr} \nu_{Vj} \right)^2 \frac{\hat{\sigma}_{Uj}^2}{\nu_{Ur}}} & \text{if } s_{YWj} \neq 0. \end{cases} \quad (3.66) \]

Note that if all the control materials are heterogeneous (\( \hat{\sigma}_{Uj}^2 = 0 \)) or if all the allocated control materials are different CRMs then the \( \hat{\sigma}_{Vj}^2 \) are independent and the degrees of freedom \( \nu_v \) are reduced to

\[ \nu_v = \sum_{j=1}^{J} \nu_{Vj}. \quad (3.67) \]

Another option is to utilize parametric bootstrap for inferences on \( d \).

### 3.2.4 Performance Evaluation

In addition to a direct performance evaluation based on the control data as described in 3.1.2.3, we can get a second performance evaluation by using the study data. The difference between these two evaluations is the previous knowledge about the true value of the concentration in the materials. One more consideration must
be taken into account for the performance evaluation of each participant and this is that the new assigned value \( \bar{d} \) (or \( \bar{d}_\omega \)) for the concentration of the compound from the study material is now a function of all the participant’s results \( \hat{d}_j \) corrected for bias. This functional relation implies the covariance \( \text{Cov}[\hat{d}_j, \bar{d}] \) (or \( \text{Cov}[\hat{d}_j, \bar{d}_\omega] \)) is non-zero.

The performance metric is then
\[
t_{\text{score,} dj} = \frac{\hat{d}_j - \bar{d}}{\sqrt{V[\hat{d}_j - \bar{d}]^2}} \quad \text{and} \quad |t_{\text{score,} dj}| < \text{some critical value} \quad (3.68)
\]
where the estimated variance can be obtained using (3.53), (3.56) and (3.58)

\[
\hat{V}[\hat{d}_j - \bar{d}] = \hat{V}[\hat{d}_j] + \hat{\sigma}_d^2 - 2\text{Cov}[\hat{d}_j, \bar{d}]
\]
\[
= \hat{\sigma}_Q^2 + \sum_r a_{jr} \hat{\sigma}_x^2 \frac{1}{J^2} \left( \sum_j' \hat{\sigma}_Q^2 + \sum_r n_r \hat{\sigma}_x^2 \right) - \frac{2}{J} \left( \hat{\sigma}_Q^2 + \sum_r a_{jr} n_r \hat{\sigma}_x^2 \right) ; \quad (3.69)
\]

The evaluation criteria and approximations used for the performance evaluation while using the control data only are still valid.

The evaluation of the participant’s method variability can be done as described in the previous section using the estimate for the within-method variance from the study data. This is
\[
f_{\text{score,} j} = \frac{\hat{\sigma}_W^2}{\hat{\sigma}_e^2} \sim F(K-1, \nu_{ej})
\]
where \( \hat{\sigma}_e^2 \) is the within method variance of the consensus value obtained from the study data corrected for bias as in (3.5) and with degrees of freedom as
\[
\nu_{ej} = \frac{\left( \hat{\sigma}_x^2 - \hat{\sigma}_e^2 \right)^2}{\hat{\sigma}_x^2 \hat{\nu}_{ej}^2 + \hat{\nu}_{ej} \hat{\sigma}_e^2} . \quad (3.70)
\]
3.2.5 Examples

We use the data of the 16 participants about the PAH family of compounds in sediments from the IEPOCME 2003 intercomparison exercise. The estimated biases are shown in a boxplot for different laboratories for each compound in Figure 3.2.

![Boxplot of estimated biases for different laboratories for each compound](image)

**Figure 3.2.** Estimated biases for the different laboratories for each compound, assuming constant bias across samples.

This boxplot shows the estimated bias $\hat{\beta}_{0j}$ for the J laboratories for each compound, from the PAHs in the Sediments 2003 intercomparison exercise.
We observe that for some compounds the bias terms are very large and for some it is hard to justify the assumption of random bias with zero mean.

Table 3.7 lists the MLE solution to the one way random effects model for the raw study values, \( \hat{d} \) = the consensus value was obtained by using (3.49), \( \hat{\sigma}_d \) = the standard error of the consensus was obtained by using (3.51), \( \nu_{\sigma d} \) was obtained by applying the Welch-Satterthwaite approximation and \((\hat{lcl}(d), \hat{ucl}(d))\) = a Wald’s type CI for \( d \) using \((\hat{d}, \hat{\sigma}_d, \nu_{\sigma d})\). \( J \) is the number of the laboratories that actually reported measurements for that compound. Compounds are in ascending order of the reference value \( \hat{x} \).

The MLE approach is based on the assumption that the true underlying model is the one way random effects model, this is, it assumes the laboratories are random with random bias having mean zero and common between-laboratory variance for each compound. These are strong assumptions, especially because the laboratories are not random but also because the assumption of mean bias = 0 may be questionable for some compounds. Since, we can just as easily proceed with fixed labs and associated biases, we do so.

Table 3.8 lists the point and interval estimates of the simple mean and the weighted mean of the study values corrected for constant bias, assuming the errors \( \epsilon \) and \( \delta \) are independent. \( \bar{d} \) was obtained by using (3.56), \( \hat{\sigma}_{\bar{d}} \) was obtained by the square root of (3.58), \( \nu_{\sigma \bar{d}} \) the degrees of freedom were obtained with the Welch-Satterthwaite approximation and \((\bar{lcl}(d), \bar{ucl}(d))\) = a Wald’s type CI for \( d \) using \((\bar{d}, \hat{\sigma}_{\bar{d}}, \nu_{\sigma \bar{d}})\).

Table 3.9 lists the point and interval estimates of the simple mean and the weighted mean of the study values corrected for constant bias assuming the errors \( \epsilon \) and \( \delta \) are correlated. \( \bar{d}_\omega \) was obtained by using (3.56), \( \hat{\sigma}_{\bar{d}_\omega} \) was obtained by the square root of (3.58), \( \nu_{\sigma \bar{d}_\omega} \) = the degrees of freedom were obtained with the Welch-Satterthwaite approximation and \((\bar{lcl}_\omega(d), \bar{ucl}_\omega(d))\) = a Wald’s type CI for \( d \) using \((\bar{d}_\omega, \hat{\sigma}_{\bar{d}_\omega}, \nu_{\sigma \bar{d}_\omega})\).
Comparing Tables 3.7 and 3.8 we observe that about half of the estimated simple means $\bar{d}$ and weighted means $\bar{d}_\omega$ are larger than the MLE solution $\hat{d}$. Also about half of the estimated variances of the simple means $\hat{\sigma}_d^2$ are larger than the estimated variance of the MLE solution $\hat{\sigma}_d^2$ and about half are smaller than that of the MLE solution. However, the estimated variance of the weighted mean $\hat{\sigma}^2_{\bar{d}_\omega}$ tends to be smaller than that of the MLE solution. The estimated variance of the weighted mean also tends to be smaller than the estimated variance of the simple mean of the study values corrected for bias. Comparing the MLE solution in Table 3.7 and the estimates based on the simple and weighted mean of the study values corrected for constant bias under correlation of the error terms in Table 3.9 we arrive to similar conclusions.

In all the cases the point estimates by definition are within the parameter space. However, in three out of the 26 compounds, the CIs based on the simple mean of the corrected study values contain zero. This problem is not present on the CI estimates based on then weighted mean.

The approach making use of both the control data and study data, either by the simple mean or the weighted mean of the study data corrected for constant bias requires weaker assumptions in contrast with the MLE approach. These approaches allows the bias to change with the laboratory and compound but it assumes the bias for each compound is constant over values of the control and study materials. Hence if the study value is very different from the levels of the CRM’s compounds then this may still be a strong assumption. We address this problem, at least partially, in Chapter 4, by introducing models for bias as a function of concentration.

The CIs based on the simple mean of the corrected study data tend to be wider than the CIs based on the MLE solution and these tend to be wider than the CIs based on the weighted mean of the corrected study data, as shown in Figure 3.3.

Comparing Figures 3.1 and 3.4 we observe the $t_{\text{score}}$ are more disperse and some $f_{\text{score}}$ get larger after adjusting for bias correction. Some results now appear ques-
Table 3.7. Example of MLE under the one way random effects model using the raw study data.

<table>
<thead>
<tr>
<th>Compound</th>
<th>$\hat{d}$</th>
<th>$\hat{\sigma}_d^2$</th>
<th>$\nu_d$</th>
<th>$\hat{lc}l(d)$</th>
<th>$\hat{uc}l(d)$</th>
<th>$J$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2,3,5-trimethynaphthalene</td>
<td>33.6</td>
<td>13.0</td>
<td>11</td>
<td>25.8</td>
<td>41.5</td>
<td>6</td>
</tr>
<tr>
<td>acenaphthene</td>
<td>33.2</td>
<td>8.2</td>
<td>21</td>
<td>27.2</td>
<td>39.1</td>
<td>12</td>
</tr>
<tr>
<td>dibenz[a,h]anthracene</td>
<td>145.9</td>
<td>335.3</td>
<td>27</td>
<td>108.4</td>
<td>183.5</td>
<td>14</td>
</tr>
<tr>
<td>acenaphthylene</td>
<td>47.3</td>
<td>90.7</td>
<td>21</td>
<td>27.6</td>
<td>67.1</td>
<td>11</td>
</tr>
<tr>
<td>1-methylphenanthrene</td>
<td>73.8</td>
<td>22.6</td>
<td>17</td>
<td>63.8</td>
<td>83.8</td>
<td>9</td>
</tr>
<tr>
<td>biphenyl</td>
<td>22.2</td>
<td>6.5</td>
<td>14</td>
<td>16.7</td>
<td>27.6</td>
<td>9</td>
</tr>
<tr>
<td>2,6-dimethynaphthalene</td>
<td>44.4</td>
<td>64.4</td>
<td>19</td>
<td>27.7</td>
<td>61.1</td>
<td>10</td>
</tr>
<tr>
<td>fluorene</td>
<td>56.9</td>
<td>30.8</td>
<td>22</td>
<td>45.4</td>
<td>68.4</td>
<td>12</td>
</tr>
<tr>
<td>triphenylene</td>
<td>103.6</td>
<td>688.1</td>
<td>3</td>
<td>30.8</td>
<td>176.4</td>
<td>2</td>
</tr>
<tr>
<td>1-methylnaphthalene</td>
<td>43.4</td>
<td>19.1</td>
<td>24</td>
<td>34.4</td>
<td>52.4</td>
<td>13</td>
</tr>
<tr>
<td>anthracene</td>
<td>137.5</td>
<td>225.9</td>
<td>31</td>
<td>106.9</td>
<td>168.1</td>
<td>16</td>
</tr>
<tr>
<td>benzo[j]fluoranthene</td>
<td>336.5</td>
<td>8472.2</td>
<td>7</td>
<td>124.2</td>
<td>548.9</td>
<td>4</td>
</tr>
<tr>
<td>benzo[k]fluoranthene</td>
<td>569.8</td>
<td>8474.2</td>
<td>25</td>
<td>380.5</td>
<td>759.0</td>
<td>13</td>
</tr>
<tr>
<td>2-methylnaphthalene</td>
<td>65.4</td>
<td>56.6</td>
<td>26</td>
<td>50.0</td>
<td>80.8</td>
<td>14</td>
</tr>
<tr>
<td>chrysene</td>
<td>713.0</td>
<td>3989.0</td>
<td>25</td>
<td>583.2</td>
<td>842.9</td>
<td>13</td>
</tr>
<tr>
<td>benzo[ghi]perylene</td>
<td>579.9</td>
<td>2610.3</td>
<td>31</td>
<td>475.8</td>
<td>684.0</td>
<td>16</td>
</tr>
<tr>
<td>benzo[e]pyrene</td>
<td>612.7</td>
<td>2512.5</td>
<td>27</td>
<td>510.0</td>
<td>715.4</td>
<td>14</td>
</tr>
<tr>
<td>benzo[a]anthracene</td>
<td>421.4</td>
<td>821.1</td>
<td>31</td>
<td>363.0</td>
<td>479.7</td>
<td>16</td>
</tr>
<tr>
<td>indeno[1,2,3-cd]pyrene</td>
<td>639.2</td>
<td>2538.9</td>
<td>31</td>
<td>536.5</td>
<td>741.8</td>
<td>16</td>
</tr>
<tr>
<td>benzo[a]pyrene</td>
<td>670.3</td>
<td>2539.0</td>
<td>31</td>
<td>567.6</td>
<td>773.1</td>
<td>16</td>
</tr>
<tr>
<td>perylene</td>
<td>187.9</td>
<td>350.3</td>
<td>24</td>
<td>149.3</td>
<td>226.5</td>
<td>13</td>
</tr>
<tr>
<td>phenantherene</td>
<td>472.1</td>
<td>836.4</td>
<td>31</td>
<td>413.2</td>
<td>531.1</td>
<td>16</td>
</tr>
<tr>
<td>benzo[b]fluoranthene</td>
<td>981.8</td>
<td>8565.9</td>
<td>29</td>
<td>792.8</td>
<td>1170.9</td>
<td>15</td>
</tr>
<tr>
<td>pyrene</td>
<td>1043.7</td>
<td>4147.7</td>
<td>31</td>
<td>912.5</td>
<td>1174.9</td>
<td>16</td>
</tr>
<tr>
<td>fluoranthene</td>
<td>1028.3</td>
<td>3566.1</td>
<td>31</td>
<td>906.7</td>
<td>1150.0</td>
<td>16</td>
</tr>
<tr>
<td>naphthalene</td>
<td>90.9</td>
<td>109.6</td>
<td>30</td>
<td>69.6</td>
<td>112.3</td>
<td>16</td>
</tr>
</tbody>
</table>

For each compound, the consensus value using study data only, assuming all the laboratories are random with expected value 0: $\hat{d} = \text{the consensus value}$, $\hat{\sigma}_d^2 = \text{the estimated variance of the consensus}$, $\nu_d = \text{the degrees of freedom of the estimated variance of the consensus}$, $\hat{lc}l(d) = \text{the lower 95\% confidence limit for the consensus}$, $\hat{uc}l(d) = \text{the upper 95\% confidence limit for the consensus}$.
For each compound, the simple mean of the study values corrected for constant bias using the control data, assuming errors are independent. $\bar{d}$ = the simple mean, $\hat{\sigma}_d^2$ = the estimated variance of the simple mean, $\nu_d$ = the degrees of freedom of the estimated variance of the simple mean, $lcl(d)$ = the lower 95% confidence limit for the simple mean, $ucl(d)$ = the upper 95% confidence limit for the simple mean, $\bar{d}_\omega$ = the weighted mean, $\hat{\sigma}_\omega^2$ = the estimated variance of the weighted mean, $\nu_{\omega}$ = the degrees of freedom of the estimated variance of the weighted mean, $lcl(\omega)$ = the lower 95% confidence limit for the weighted mean, $ucl(\omega)$ = the upper 95% confidence limit for the weighted mean.

Table 3.8. Example of simple mean and weighted mean of study values corrected for constant bias, under independent errors.
Table 3.9. Example of simple mean and weighted mean of study values corrected for constant bias, under correlated errors.

For each compound, the simple mean of the study values corrected for constant bias using the control data, assuming errors are correlated. $\bar{d}$ = the simple mean, $\hat{\sigma}_\bar{d}^2 = \text{the estimated variance of the simple mean}$, $\nu_\bar{d} = \text{the degrees of freedom of the estimated variance of the simple mean}$, $lcl(\bar{d}) = \text{the lower 95% confidence limit for the simple mean}$, $ucl(\bar{d}) = \text{the upper 95% confidence limit for the simple mean}$, $\bar{d}_\omega = \text{the weighted mean}$, $\hat{\sigma}_{\bar{d}_\omega}^2 = \text{the estimated variance of the weighted mean}$, $\nu_{\bar{d}_\omega} = \text{the degrees of freedom of the estimated variance of the weighted mean}$, $lcl(\bar{d}_\omega) = \text{the lower 95% confidence limit for the weighted mean}$, $ucl(\bar{d}_\omega) = \text{the upper 95% confidence limit for the weighted mean}$. 

<table>
<thead>
<tr>
<th>Compound</th>
<th>$\bar{d}$</th>
<th>$\hat{\sigma}_\bar{d}^2$</th>
<th>$\nu_\bar{d}$</th>
<th>$lcl(\bar{d})$</th>
<th>$ucl(\bar{d})$</th>
<th>$\bar{d}_\omega$</th>
<th>$\hat{\sigma}<em>{\bar{d}</em>\omega}^2$</th>
<th>$\nu_{\bar{d}_\omega}$</th>
<th>$lcl(\bar{d}_\omega)$</th>
<th>$ucl(\bar{d}_\omega)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2,3,5-trimethynaphthalene</td>
<td>24.5</td>
<td>7.6</td>
<td>33</td>
<td>19.0</td>
<td>30.1</td>
<td>25.3</td>
<td>7.1</td>
<td>31</td>
<td>19.9</td>
<td>30.8</td>
</tr>
<tr>
<td>acenaphthene</td>
<td>26.6</td>
<td>14.3</td>
<td>10</td>
<td>18.3</td>
<td>35.0</td>
<td>34.5</td>
<td>7.3</td>
<td>32</td>
<td>29.0</td>
<td>40.1</td>
</tr>
<tr>
<td>dibenz[a,h]anthracene</td>
<td>60.7</td>
<td>138.5</td>
<td>19</td>
<td>36.1</td>
<td>85.4</td>
<td>111.2</td>
<td>27.0</td>
<td>11</td>
<td>99.7</td>
<td>122.6</td>
</tr>
<tr>
<td>acenaphthylene</td>
<td>5.5</td>
<td>17.7</td>
<td>14</td>
<td>-3.5</td>
<td>14.5</td>
<td>35.3</td>
<td>11.0</td>
<td>29</td>
<td>28.5</td>
<td>42.0</td>
</tr>
<tr>
<td>1-methylphenanthrene</td>
<td>82.6</td>
<td>10.8</td>
<td>49</td>
<td>76.0</td>
<td>89.2</td>
<td>77.7</td>
<td>9.6</td>
<td>30</td>
<td>71.4</td>
<td>84.0</td>
</tr>
<tr>
<td>biphenyl</td>
<td>34.8</td>
<td>17.8</td>
<td>31</td>
<td>26.2</td>
<td>43.4</td>
<td>34.4</td>
<td>17.1</td>
<td>29</td>
<td>25.9</td>
<td>42.8</td>
</tr>
<tr>
<td>2,6-dimethynaphthalene</td>
<td>32.7</td>
<td>6.5</td>
<td>28</td>
<td>27.5</td>
<td>37.9</td>
<td>29.9</td>
<td>5.5</td>
<td>24</td>
<td>25.1</td>
<td>34.7</td>
</tr>
<tr>
<td>fluorene</td>
<td>43.9</td>
<td>60.2</td>
<td>45</td>
<td>28.3</td>
<td>59.6</td>
<td>61.8</td>
<td>59.5</td>
<td>42</td>
<td>46.2</td>
<td>77.4</td>
</tr>
<tr>
<td>triphenylene</td>
<td>108.3</td>
<td>12.2</td>
<td>9</td>
<td>100.4</td>
<td>116.1</td>
<td>95.5</td>
<td>9.9</td>
<td>14</td>
<td>88.7</td>
<td>102.2</td>
</tr>
<tr>
<td>1-methylphenanthalene</td>
<td>85.2</td>
<td>53.3</td>
<td>63</td>
<td>70.6</td>
<td>99.8</td>
<td>66.9</td>
<td>46.3</td>
<td>34</td>
<td>53.1</td>
<td>80.7</td>
</tr>
<tr>
<td>anthracene</td>
<td>228.5</td>
<td>1843.8</td>
<td>31</td>
<td>141.0</td>
<td>316.0</td>
<td>133.0</td>
<td>87.3</td>
<td>41</td>
<td>114.1</td>
<td>151.8</td>
</tr>
<tr>
<td>benzo[j]fluorantheine</td>
<td>212.3</td>
<td>7371.0</td>
<td>5</td>
<td>-2.4</td>
<td>427.1</td>
<td>290.1</td>
<td>7.7</td>
<td>6</td>
<td>283.5</td>
<td>296.7</td>
</tr>
<tr>
<td>benzo[k]fluorantheine</td>
<td>75.9</td>
<td>1304.8</td>
<td>12</td>
<td>-2.5</td>
<td>154.2</td>
<td>345.6</td>
<td>94.8</td>
<td>29</td>
<td>325.7</td>
<td>365.5</td>
</tr>
<tr>
<td>2-methynaphthalene</td>
<td>147.5</td>
<td>441.9</td>
<td>43</td>
<td>105.2</td>
<td>189.9</td>
<td>154.3</td>
<td>443.5</td>
<td>41</td>
<td>111.8</td>
<td>196.8</td>
</tr>
<tr>
<td>chrysene</td>
<td>251.3</td>
<td>3098.9</td>
<td>3</td>
<td>95.5</td>
<td>407.0</td>
<td>416.9</td>
<td>272.7</td>
<td>12</td>
<td>380.9</td>
<td>452.9</td>
</tr>
<tr>
<td>benzo[ghi]pyrene</td>
<td>608.3</td>
<td>627.2</td>
<td>56</td>
<td>558.1</td>
<td>658.5</td>
<td>526.7</td>
<td>495.4</td>
<td>39</td>
<td>481.7</td>
<td>571.7</td>
</tr>
<tr>
<td>benzo[e]pyrene</td>
<td>783.7</td>
<td>822.9</td>
<td>12</td>
<td>721.4</td>
<td>845.9</td>
<td>557.3</td>
<td>163.6</td>
<td>34</td>
<td>531.3</td>
<td>583.3</td>
</tr>
<tr>
<td>benz[a]anthracene</td>
<td>424.6</td>
<td>1254.1</td>
<td>4</td>
<td>332.9</td>
<td>516.3</td>
<td>424.9</td>
<td>154.4</td>
<td>44</td>
<td>399.9</td>
<td>449.9</td>
</tr>
<tr>
<td>indeno[1,2,3-cd]pyrene</td>
<td>666.6</td>
<td>760.5</td>
<td>64</td>
<td>611.6</td>
<td>721.7</td>
<td>614.3</td>
<td>763.0</td>
<td>41</td>
<td>558.6</td>
<td>670.1</td>
</tr>
<tr>
<td>benzo[a]pyrene</td>
<td>695.9</td>
<td>2431.7</td>
<td>4</td>
<td>568.8</td>
<td>823.0</td>
<td>524.6</td>
<td>80.6</td>
<td>35</td>
<td>506.4</td>
<td>542.8</td>
</tr>
<tr>
<td>perylene</td>
<td>326.9</td>
<td>1869.6</td>
<td>52</td>
<td>240.1</td>
<td>413.6</td>
<td>227.8</td>
<td>455.1</td>
<td>39</td>
<td>184.7</td>
<td>271.0</td>
</tr>
<tr>
<td>phenanthrene</td>
<td>413.3</td>
<td>1784.9</td>
<td>12</td>
<td>321.6</td>
<td>504.9</td>
<td>437.6</td>
<td>473.9</td>
<td>43</td>
<td>393.8</td>
<td>481.5</td>
</tr>
<tr>
<td>benzo[b]fluorantheine</td>
<td>840.8</td>
<td>4444.2</td>
<td>22</td>
<td>702.8</td>
<td>978.8</td>
<td>872.2</td>
<td>148.2</td>
<td>15</td>
<td>846.4</td>
<td>898.0</td>
</tr>
<tr>
<td>pyrene</td>
<td>1329.0</td>
<td>8456.3</td>
<td>8</td>
<td>1117.8</td>
<td>1540.1</td>
<td>1020.1</td>
<td>388.5</td>
<td>45</td>
<td>980.5</td>
<td>1659.8</td>
</tr>
<tr>
<td>fluorantheine</td>
<td>857.2</td>
<td>5659.8</td>
<td>6</td>
<td>677.5</td>
<td>1036.9</td>
<td>989.1</td>
<td>636.9</td>
<td>44</td>
<td>938.3</td>
<td>1040.0</td>
</tr>
<tr>
<td>naphthalene</td>
<td>334.0</td>
<td>2803.6</td>
<td>66</td>
<td>228.3</td>
<td>439.7</td>
<td>278.6</td>
<td>2070.9</td>
<td>41</td>
<td>186.7</td>
<td>370.5</td>
</tr>
</tbody>
</table>
Figure 3.3. Estimated CI lengths for each compound, assuming constant bias across samples.

This plot shows the estimated CI length for each compound, on the x-axis the estimate based on the MLE solution, on the y-axis the estimates based on the simple mean (empty dots for independent errors, solid dots for correlated errors) and the weighted mean (empty triangles for independent errors and solid triangles for correlated errors), both of the study values corrected for constant bias.

Some results were questionably and some appear unsatisfactory in contrast with the evaluation with control data where all the results were satisfactory. These results are preliminary since we
have made assumptions that we have not tested, such as the constant bias of the method and constant within method variance for each compound.

Figure 3.4. Example of performance evaluation using the adjusted study data. This target plot shows the $t_{\text{score}}$ and $\sqrt{J_{\text{score}}}$ of the study data corrected for bias, from the PAHs in the Sediments 2003 exercise Program 1. Empty dots represent regular data, filled triangles represent data with absolute value larger than 4 in any of the scores, filled dot represents the mean of the scores.
3.3 One laboratory, multiple compounds at one point in time

When working with multiple compounds measured at one point in time with the same method, we can try to improve the bias and within-method variance estimations by pooling information across compounds. Another approach is to model the bias and the within-method variance as functions of the true value associated with the compound. We will discuss the bias modeling as a function of the true value associated with the compound in Chapter 4.

3.3.1 The Basic Model

The model for the observed value of the \( a \)th compound in the control material is

\[
Y_{ak} = \beta_0a + x_a + U_{ak} + \epsilon_{ak}
\]

(3.71)

where \( a = 1, \ldots, A \), \( k = 1, \ldots, K \), \( E[U_{ak}] = 0 \), \( V[U_{ak}] = \sigma^2_Ua \), \( E[\epsilon_{ak}] = 0 \), \( V[\epsilon_{ak}] = \sigma^2_{\epsilon a} \) and \((\beta_{0a}, x_a, \sigma^2_{\epsilon a}, \sigma^2_Ua)\) are unknown parameters. \( U_{ak} \) and \( \epsilon_{ak} \) are independent for a given \( a,k \) and are independent among \( a,k \). Assume that \( \hat{x}_a, \hat{\sigma}^2_Ua \) and \( \hat{\sigma}^2_{\epsilon a} \) are known unbiased estimators for \( x_a, \sigma^2_Ua, \) and \( V[\hat{x}_a] \); these come from the certification process as discussed in Section 2.3.

The model for the observed value of the compounds in the study material is

\[
W_{ak} = \beta_0a + d_a + V_{ak} + \delta_{ak},
\]

(3.72)

where \( a = 1, \ldots, A \), \( k = 1, \ldots, K \), \( E[V_{ak}] = 0 \), \( V[V_{ak}] = \sigma^2_{V a} \), \( E[\delta_{ak}] = 0 \), \( V[\delta_{ak}] = \sigma^2_{\delta a} \) and \((d_a, \sigma^2_{V a})\) are unknown parameters. The \( \beta_{0a} \) are shared by the same compound on both models, \( V_{ak} \) and \( \delta_{ak} \) are independent for a given \( a,k \) and are independent among \( a,k \). Also assume that \( \delta_{ak}, V_{ak} \) are independent from \( \epsilon_{ak}, U_{ak} \). Then

\[
E[Y_{ak}] = \beta_{0a} + x_a,
\]

(3.73)

\[
V[Y_{ak}] = \sigma^2_{U a} + \sigma^2_{\epsilon a},
\]

\[
E[W_{ak}] = \beta_{0a} + d_a \text{ and }
\]

(3.74)

\[
V[W_{ak}] = \sigma^2_{V a} + \sigma^2_{\delta a}.
\]
3.3.2 Working with the Control Data Only

Here there is no interest on pooling information across different participants. However, there is interest on testing whether a participant’s method is biased. When working with the control data only, the performance of the participants can be assessed directly, provided that the conventional value of the control material and its related variance components are known.

If $\beta_0$ and $\sigma^2_{\epsilon a}$ are different for each compound then unbiased estimators for $\beta_0$ and $\sigma^2_{\epsilon a}$ can be obtained by using the results from Section 3.1.2.1.

### 3.3.2.1 Parameter Estimation

Often the $g$ transformation and the $h$ function are shared among a set of compounds. Often all or many compounds in a material undergo the same processes and transformations in order to be measured and thus any error related to the measurement procedure is shared. In such cases, the assumption of sharing parameters among a whole set of compounds appears to be reasonable and the information across compounds can be used to estimate the parameters.

Under the assumption of a constant bias shared among compounds, $\beta_{0a} = \beta_0$. Consider the point estimator

$$\hat{\beta}_0 = \sum_{a=1}^{A} \hat{\omega}_a \hat{\beta}_{0a}$$

(3.75)

where $\hat{\beta}_{0a} = \bar{Y}_a - \hat{x}_a$, $\hat{\omega}_a = \frac{\hat{\sigma}_{\beta 0a}^{-2}}{\sum \hat{\sigma}_{\beta 0a}^{-2}}$, $\hat{\sigma}_{\beta 0a}^2 = \frac{s^2_{Y a} + K_a \hat{\sigma}_{\epsilon a}^2}{K_a} \text{ and } \sum_{a=1}^{A} \hat{\omega}_a = 1$. Under normality $\hat{\beta}_{0a}$ and $\hat{\omega}_a$ are independent and using double expectation (e.g. Casella & Berger 2002 [11], Theorem 4.4.3.)

$$E[\hat{\beta}_0] = E \left[ E \left[ \sum_{a=1}^{A} \omega_a \hat{\beta}_{0a} \bigg| \hat{\omega}_1, \ldots, \hat{\omega}_A \right] \right] = \beta_0 E \left[ \sum_{a=1}^{A} \hat{\omega}_a \right] = \beta_0.$$

(3.76)

Hence $\hat{\beta}_0$ as defined in (3.75) is unbiased for $\beta_0$ under normality or more generally under the assumption of $\hat{\omega}_a$ random and independent of $\hat{\beta}_{0a}$. If the distribution of
the error terms is not normal then the sample variance and sample mean may not be independent and in general (3.76) becomes an approximation.

Again assuming independence of \( \hat{\beta}_0 a \) and \( \hat{\omega}_a \), and using the conditional variance identity (e.g. Casella & Berger 2002 [11], Theorem 4.4.7.) we have

\[
\sigma^2_{\hat{\beta}_0} = V[\hat{\beta}_0] = \left[ E \left( \sum_{a=1}^{A} \hat{\omega}_a \hat{\beta}_0 a \right) \right] + \left[ E \left( \sum_{a=1}^{A} \hat{\omega}_a^2 \right) \right] = V[\hat{\beta}_0] + \sum_{a=1}^{A} \sigma^2_{\hat{\beta}_0 a} = \sum_{a=1}^{A} V[\hat{\beta}_0] E[\hat{\omega}_a^2].
\]

(3.77)

If we treat \( \hat{\omega}_a \) as fixed, the variance of \( \hat{\beta}_0 \) can be estimated as

\[
\hat{\sigma}^2_{\hat{\beta}_0} \approx \sum_{a} \hat{\omega}_a^2 V[\hat{\beta}_0 a] = \sum_{a} \hat{\sigma}^2_{\hat{\beta}_0 a} \left( \sum_{a'} \sigma^{-2}_{\hat{\beta}_0 a'} \right)^2 = \sum_{a} \hat{\sigma}^{-2}_{\hat{\beta}_0 a}.
\]

(3.78)

By Jensen’s inequality, it is straightforward to show that (3.78) overestimates \( \frac{1}{\sum_a \sigma^2_{\hat{\beta}_0 a}} \). Also it is straightforward to see that \( \hat{\sigma}^2_{\hat{\beta}_0} \) is a consistent estimator for \( \frac{1}{\sum_a \sigma^2_{\hat{\beta}_0 a}} \).

There are more options to be considered for estimating the variance function \( h \). For example, \( h \) can be constant but in practice is often some function of the concentration.

- Under the assumption of normality, homogeneity of the control material (i.e., \( \sigma^2_{U_a} = 0, \forall a \)) and constant \( h \) across compounds, \( E[\hat{\sigma}^2_{e a}] = \sigma^2_{e a} = \sigma^2_{e} \), see section 3.1.2.1, specifically (3.5) and (3.7) we have

\[
\frac{(K_a - 1) \hat{\sigma}^2_{e a}}{\sigma^2_{e}} \sim_{\text{ind}} \chi^2_{(K_a - 1)}
\]

(3.79)

or

\[
\sum_{a=1}^{A} (K_a - 1) \hat{\sigma}^2_{e a} \frac{\sigma^2_{e}}{\sigma^2_{e}} \sim \chi^2_{\sum (K_a - 1)}.
\]

(3.80)
Now define
\[
\hat{\sigma}_\epsilon^2 = \frac{\sum_{a=1}^{A} (K_a - 1) \hat{\sigma}_{\epsilon a}^2}{\sum_{a=1}^{A} (K_a - 1)},
\] (3.81)
then
\[
E[\hat{\sigma}_\epsilon^2] = E \left[ \frac{\sum_{a=1}^{A} (K_a - 1) \hat{\sigma}_{\epsilon a}^2}{\sum_{a=1}^{A} (K_a - 1)} \right] = \frac{\sum_{a=1}^{A} (K_a - 1) E[\hat{\sigma}_{\epsilon a}^2]}{\sum_{a=1}^{A} (K_a - 1)} = \sigma_\epsilon^2. \tag{3.82}
\]
Hence \(\hat{\sigma}_\epsilon^2\) as defined in (3.81) is unbiased for \(\sigma_\epsilon^2\).

We are assuming the estimated variances of the error in the equation for different compounds are independent.

- Under the assumption of normality, and heterogeneity of the control material (i.e., exists \(a\) such that \(\sigma_{UA}^2 \neq 0\)) and constant \(h\) across compounds, \(E[\hat{\sigma}_{\epsilon a}^2] = \sigma_{\epsilon a}^2 = \sigma_\epsilon^2\). Assume we have estimates for \(\nu_{\epsilon a} = \text{degrees of freedom of } \hat{\sigma}_{\epsilon a}^2\), then we have

\[
\frac{\nu_{\epsilon a} \hat{\sigma}_{\epsilon a}^2}{\sigma_\epsilon^2} \sim \chi^2(\nu_{\epsilon a}), \tag{3.83}
\]
or
\[
\frac{\sum_{a=1}^{A} \nu_{\epsilon a} \hat{\sigma}_{\epsilon a}^2}{\sigma_\epsilon^2} \sim \chi^2(\sum \nu_{\epsilon a}). \tag{3.84}
\]

Now define
\[
\hat{\sigma}_\epsilon^2 = \frac{\sum_{a=1}^{A} \nu_{\epsilon a} \hat{\sigma}_{\epsilon a}^2}{\sum_{a=1}^{A} \nu_{\epsilon a}} \tag{3.85}
\]
then
\[
E[\hat{\sigma}_\epsilon^2] = E \left[ \frac{\sum_{a=1}^{A} \nu_{\epsilon a} \hat{\sigma}_{\epsilon a}^2}{\sum_{a=1}^{A} \nu_{\epsilon a}} \right] = \frac{\sum_{a=1}^{A} \nu_{\epsilon a} E[\hat{\sigma}_{\epsilon a}^2]}{\sum_{a=1}^{A} \nu_{\epsilon a}} = \sigma_\epsilon^2. \tag{3.86}
\]
Hence \(\hat{\sigma}_\epsilon^2\) as defined in (3.85) is approximately unbiased for \(\sigma_\epsilon^2\). The estimated degrees of freedom of \(\hat{\sigma}_{UA}^2\) must be known from the CRM in order to obtain estimated degrees of freedom of \(\hat{\sigma}_{\epsilon a}^2\). The approximation is due to the assumption in the individual distributions (3.83).
Assuming the CRM provides both $\hat{\sigma}^2_{Ua}$ and $\nu_{Ua}$ for each compound and assuming normality, then applying Welch-Satterthwaite approximation

$$\nu_{\epsilon a} \approx \frac{(s^2_{Ya} - \hat{\sigma}^2_{Ua})^2}{K_a - 1 + \frac{\hat{\sigma}^4_{\nu_{Ua}}}{\nu_{Ua}}}.$$ (3.87)

Some additional distributional assumptions on the error terms are required in order to get confidence intervals for $\beta_0$ and $\sigma^2_\epsilon$. Assuming normality of $U_{ak}$ and $\epsilon_{ak}$, the distribution of $\beta_0$ is

$$\frac{\hat{\beta}_0 - \beta_0}{\hat{\sigma}_{\beta_0}} \sim T(\nu_{\beta_0})$$ (3.88)

where $\nu_{\beta_0}$ can be approximated by the Welch-Satterthwaite method, using (3.78) and the delta method, we obtain

$$\nu_{\beta_0} \approx \frac{2 (\hat{\sigma}^2_{\beta_0})}{V[\hat{\sigma}^2_{\beta_0}]} = \frac{(\hat{\sigma}^2_{\beta_0})}{\sum \frac{\hat{\sigma}^{-4}_{\beta_0 \epsilon_{\beta_0}}}{\nu_{\beta_0 \epsilon_{\beta_0}}}} = \frac{\hat{\sigma}^{-4}_{\beta_0}}{\sum \frac{\hat{\sigma}^{-4}_{\beta_0 \epsilon_{\beta_0}}}{\nu_{\beta_0 \epsilon_{\beta_0}}}},$$ (3.89)

and $\nu_{\beta_0} \epsilon_{\beta_0}$ is given in (3.11).

An approximate $100(1 - \alpha)\%$ confidence interval for $\beta_0$ is then

$$\hat{\beta}_0 \pm t_{(1 - \alpha/2; \nu_{\beta_0})} \hat{\sigma}_{\beta_0}.$$ (3.90)

When all $K_a$ are large the distribution of $\beta_0$ is

$$\frac{\hat{\beta}_0 - \beta_0}{\hat{\sigma}_{\beta_0}} \sim Z,$$ (3.91)

and the approximate $100(1 - \alpha)\%$ confidence interval becomes

$$\hat{\beta}_0 \pm z_{(1 - \alpha/2)} \hat{\sigma}_{\beta_0}.$$ (3.92)
Similarly, using all observations together and assuming a constant \( h \) model shared among all the compounds, using (3.80) we can obtain a 100(1\( - \alpha \))% confidence interval for \( \sigma^2 \) under the assumption of homogeneous control material (i.e., \( \sigma^2_{Ua} = 0, \forall a \)) by

\[
\left( \frac{\hat{\sigma}^2 \sum (K_a - 1)}{\chi^2_{(1 - \alpha/2; \sum (K_a - 1))}}, \frac{\hat{\sigma}^2 \sum (K_a - 1)}{\chi^2_{(\alpha/2; \sum (K_a - 1))}} \right)
\]

and under the assumption of heterogeneous control material (i.e., exists \( a \) such that \( \sigma^2_{Ua} \neq 0 \)) we have approximately

\[
\left( \frac{\hat{\sigma}^2 \sum \nu_{ea}}{\chi^2_{(1 - \alpha/2; \sum \nu_{ea})}}, \frac{\hat{\sigma}^2 \sum \nu_{ea}}{\chi^2_{(\alpha/2; \sum \nu_{ea})}} \right).
\]

The approximation comes from the use of an approximated \( \chi^2 \) distribution with approximated degrees of freedom \( \nu_{ea} \).

We have assumed that both the bias and the variance are constant, we now introduce a way to test these assumptions.

### 3.3.2.2 Inferences about the Parameters

**Testing for zero bias**

Under the hypothesis of an unbiased \( g \) model shared among all the compounds \((H_0 : \beta_0 = 0)\), (3.88) reduces to

\[
\frac{\hat{\beta}_0}{\hat{\sigma}_0} \overset{\sim}{\sim} T(\nu_{\beta_0}),
\]

or

\[
\frac{f^*}{\hat{\sigma}_0^2} \overset{\sim}{\sim} F(1, \nu_{\beta_0}).
\]

\( H_0 \) can be rejected at a significance level of approximately \( \alpha \) if

\[
f^* > f_{(1 - \alpha; 1, \nu_{\beta_0})}.
\]

The approximation comes from using an approximate distribution with approximate degrees of freedom. The exact distribution of (3.88) is not a standard t distribution.
Testing for constancy of the bias

In order to test that all of the biases are equal (i.e., \( H_0 : \beta_{0a} = \beta_0, \forall a \)), a general linear hypotheses test (Ravishanker and Dey 2002 [43]) can be conducted. Consider \( H_0 : C\hat{\beta}_0 = 0 \), where \( \hat{\beta}_0 = [\beta_{01}, \ldots, \beta_{0A}]^T \), and \( C \) is a \((A-1) \times A\) matrix of independent contrasts. Then, under normality and known covariance matrix \( \Sigma_{\beta} \) we have:

\[
(C\hat{\beta}_0)^T(C\Sigma_{\beta}C^T)^{-1}(C\hat{\beta}_0) \sim \chi^2_A.
\]

In practice \( \Sigma_{\beta} \) is unknown, instead we use the estimated covariance matrix \( \hat{\Sigma}_{\beta} \) and we define:

\[
\chi^2_* = (C\hat{\beta}_0)^T(C\hat{\Sigma}_{\beta}C^T)^{-1}(C\hat{\beta}_0)
\]

and \( \chi^2_* \) is approximately a \( \chi^2_A \) random variable.

\( H_0 \) can be rejected at a significance level of approximately \( \alpha \) if

\[
\chi^2_* > \chi^2_{(1-\alpha; A)}.
\]

The approximation comes from using an estimated covariance matrix for \( \beta \).

Testing for constant within-laboratory variances

Similarly the assumption of constant within variances can be tested by using the well known Levene’s Test (Brown and Forsythe (1974) [5]). This test allows changing biases over compounds.

- For homogeneous materials (i.e., \( \sigma^2_{Ua} = 0, \forall a \)). Consider \( H_0 : \hat{\sigma}^2_{a} = \hat{\sigma}^2_{a'}, \forall a, a' \), then under normality the Levene’s test statistic:

\[
f^* = \frac{A}{A-1} \left[ \sum_{a=1}^{A} K_a (\bar{D}_a - \bar{D})^2 \right]^{-1} \left[ \sum_{a=1}^{A} \sum_{k=1}^{K_a} (D_{ak} - \bar{D}_a)^2 \right],
\]

\[
f^* \sim F_{(A-1, \sum (K_a-1))}
\]

where \( D_{ak} = |\epsilon_{ak} - \bar{\epsilon}_a| = |Y_{ak} - \bar{Y}_a|; H_0 \) can be rejected at a significance level of approximately \( \alpha \) if

\[
f^* > f_{(1-\alpha; A-1, \sum (K_a-1))}.
\]
• For heterogeneous materials (i.e., \( \exists a \) such that \( \sigma^2_{Ua} \neq 0 \)) we can use the version of Levene’s Test based on squared residuals using (3.100) where \( D_{ak} = (Y_{ak} - \bar{Y}_a)^2 - \hat{\sigma}^2_{Ua} \). \( H_0 \) can be rejected at a significance level of approximately \( \alpha \) as above. Details are found in Appendix B.

### 3.3.2.3 Performance Evaluation

We can proceed as in the previous section assuming a constant bias over compounds, by using an upgraded \( t_{score} \) from (3.88) and upgraded critical values by using the corresponding degrees of freedom.

If the within-method variance is constant, then the evaluation of the participant’s method variability can be done by comparing the pooled variance estimate against the variability of the methods used during the certification. Let \( \sigma^2_{\epsilon c} \) be the pooled variance across compounds of the CRM and \( \nu_{\epsilon c} \) be its degrees of freedom. Under normality and independence of the error terms we have: \( \hat{\sigma}^2_{\epsilon} \sim \frac{\sigma^2_{\epsilon c}}{\nu_{\epsilon}} \chi^2_{\nu_{\epsilon}} \) and \( \hat{\sigma}_{\epsilon} \sim \frac{\sigma_{\epsilon} \nu_{\epsilon}}{\nu_{\epsilon}} \chi^2_{\nu_{\epsilon}} \), hence using (3.22) and the same argument

\[
\text{f}_{score} = \frac{\hat{\sigma}^2_{\epsilon}}{\sigma^2_{\epsilon c}} \sim F_{(\nu_{\epsilon}, \nu_{\epsilon c})}.
\]

If the confidence interval of \( f_{score} \) contains one then the observed within-method variance is considered satisfactory.

### 3.3.2.4 Examples

Here we use the data from Tables 3.1 and 3.2 to illustrate the methods just described. Since our main assumption is that the bias and the within-method variance are constant across compounds, we will test for these assumptions first.

In order to test for constant bias across compounds we use (3.98) and we find the value \( \chi^2 = 16.15 \) and \( \chi^2_{0.95; A=26} = 38.89 \), since \( \chi^2 < \chi^2_{0.95, 26} \) there is no evidence of unequal biases. Note that the estimated biases differ dramatically but the CI’s for bias are very wide so that each of them contains zero. Furthermore, we can use (3.96)
to test for zero bias and we find the value $f^* = 0.00024$ and $f_{0.95;1,137} = 3.91$, since $f^* < f_{0.95;1,137}$ there is no evidence of bias.

In order to test for constant within-method variance we use the detailed data from the participant and (3.100), we obtain: $f^* = 4.56$ and $f_{0.95;25,52} = 1.72$, since $f^* > f_{0.95;25,52}$ hence there is enough evidence to reject $H_0$ and to conclude the within-method variance is not constant across compounds, our assumption of constant within-method variance is inadequate and in general $\sigma_{ea}^2 \neq \sigma_{ea'}^2$, $\forall a \neq a'$.

We show in Table 3.10 the estimates for the constant bias when the within-method variance changes across compounds and the estimates for constant bias when the within-method variance is constant, using (3.75), (3.82) and (3.81). The estimates for the constant bias are to be taken with caution, since the test of constant bias is approximate. Although we decide to accept the hypothesis of equal biases, from Table 3.2 we observe the estimated individual biases vary considerably and this can lead us to some power reduction of the test. Hence the conclusion of constant bias has some type II error risk involved. There is also strong evidence to consider the within-method variance is not constant across compounds. The estimate for both are shown for illustration purpose but proceeding assuming constant within-method variance has some risk. The estimated degrees of freedom of $\beta_0$ are reduced when $\sigma_\epsilon^2$ is assumed constant since the estimate $\hat{\sigma}_\epsilon^2$ is reused for each $\hat{\sigma}_\beta_0^2$ and additional covariance terms are involved in the derivation of the degrees of freedom.

<table>
<thead>
<tr>
<th>parameter</th>
<th>estimate</th>
<th>variance</th>
<th>dof</th>
<th>lcl</th>
<th>ucl</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_0</td>
<td>\sigma_{ea}^2 \neq \sigma_{ea'}^2$</td>
<td>0.0152</td>
<td>0.961</td>
<td>146</td>
<td>-1.92</td>
</tr>
<tr>
<td>$\beta_0</td>
<td>\sigma_{ea}^2 = \sigma_{ea'}^2$</td>
<td>0.3244</td>
<td>3.127</td>
<td>129</td>
<td>-3.17</td>
</tr>
<tr>
<td>$\sigma_\epsilon^2$</td>
<td>85.8</td>
<td>283</td>
<td>52</td>
<td>60.4</td>
<td>131</td>
</tr>
</tbody>
</table>

Table 3.10. Parameter pooled estimates assuming constant bias and constant within-method variance.

$dof$ = the degrees of freedom of the estimated variance, $lcl$ = the lower confidence limit, $ucl$ = the upper confidence limit.
Figure 3.5 shows the constant bias estimate and its 95% confidence interval along with the individual estimates of the bias of each compound, see Table 3.2. Figure 3.6 shows the constant within-method variance estimate with its 95% confidence interval along with the individual estimates of the within-method variance of each compound. We can observe that four out of the 26 individual estimates appear to be different from the estimated constant within-method variance.

Using the value found in Table 3.10 we get a global $t_{\text{score}} = 0.0155$ and $t_{\text{sat}} = t_{(0.977,137)} = 2.01$. Under this criteria we conclude the performance about biasedness of the participating method is satisfactory; this is, the laboratory’s method is unbiased. Following the same argument we used for the constant bias this assessment has some risk. As we tested previously, the within variance is not constant so we do not offer an example for a global $f_{\text{score}}$ to evaluate the performance.
Figure 3.5. Point and interval estimation of the method constant bias.

This plot shows the individual estimates of the biases ($\hat{\beta}_0$) for each compound and the overall estimate of the bias $\hat{\beta}_0$ from the PAHs in the Sediments 2003 exercise for one of the participants. The solid line is the estimated constant bias of the method, the dotted lines are the 95% confidence interval for the estimated constant bias.
Figure 3.6. Point and interval estimation of the method constant variance

This plot shows the individual estimates of the within variances ($\sigma^2_{\text{comp}}$) for each compound and the overall estimate of the within variance $\sigma^2_{\text{c}}$ from the PAHs in the Sediments 2003 exercise for one of the participants. The solid line is the estimated constant within variance of the method, the dotted lines are the 95% confidence interval for the estimated constant within variance.
3.3.3 Exporting the Model to the Study Data

The models found while working with the control data can be used to estimate the unknown parameters of the study data.

3.3.3.1 Parameter Estimation

The focus is on estimating the true values of the study sample \( d_a \) under the constant bias model across compounds. The only change is the use of the pooled bias estimate \( \hat{\beta}_0 \). Otherwise we can proceed one compound at a time as described in the section 3.1.3.

**Independent errors:**

As before, we first assume that the error terms are independent for the method measuring the control sample and the study sample; this implies \( \text{Corr} \{ \epsilon_{ak}, \delta_{ak} \} = \rho_a = 0 \). Under the assumption of normality and constant within-method variance for each compound \( \sigma_{\delta a}^2 = \sigma_{\epsilon a}^2 \). Let

\[
\hat{d}_a = \bar{W}_a - \hat{\beta}_0
\]  

and

\[
\hat{\sigma}_{Va}^2 = s_{Wa}^2 - \hat{\sigma}_{\epsilon a}^2.
\]

Then

\[
E[\hat{d}_a] = E[\bar{W}_a - \hat{\beta}_0] = d_a + \beta_0 - E[\hat{\beta}_0] = d_a
\]  

and

\[
V[\hat{d}_a] = V[\bar{W}_a] + V[\hat{\beta}_0] - 2\text{Cov}[\bar{W}_a, \hat{\beta}_0] = V[\bar{W}_a] + V[\hat{\beta}_0]
\]

which can be estimated by \( \hat{\sigma}_{da}^2 = \frac{s_{Wa}^2}{M_a} + \hat{\sigma}_{\beta_0}^2 \). Also

\[
E[\hat{\sigma}_{Va}^2] = E[s_{Wa}^2 - \hat{\sigma}_{\epsilon a}^2].
\]

The distribution of \( \hat{d}_a \) is

\[
\frac{\hat{d}_a - d_a}{\hat{\sigma}_{da}} \sim T(\nu_{da})
\]
and the approximate confidence interval for \( d_a \) then becomes

\[
\hat{d}_a \pm t(\alpha/2, \nu_{da}) \hat{\sigma}_{da}
\]  \hspace{1cm} (3.108)

where \( \nu_{da} \) can be estimated with the Welch-Satterthwaite approximation, as shown in (3.31) adding the corresponding subindex.

**Correlated errors:**

Now assume that the error terms are correlated for the method measuring the control sample and the study sample, this is \( K_a = M_a \) with pairing and \( Corr[\epsilon_{ak}, \delta_{ak}] = \rho_a \neq 0 \). Define the sample statistics as in Section 3.1.3.1. Then it is straightforward to show that: \( E[\hat{W}_a] = \beta_0 + d_a, \) \( E[\hat{s}_a^2] = \sigma^2_{V_a} + \sigma^2_{\delta a}, \) and \( E[s_{YW_a}] = \rho_a \sigma_{\epsilon a} \sigma_{\delta a}. \) Hence the point estimators defined as (3.102) and (3.103) are still unbiased. However \( V[\hat{d}_a] \) changes due to the covariance term

\[
V[\hat{d}_a] = V[\hat{W}_a] + V[\hat{\beta}_0] - 2Cov[\hat{W}_a, \hat{\beta}_0].
\]  \hspace{1cm} (3.109)

This is estimated by

\[
\hat{\sigma}^2_{da} = \frac{s^2_{Wa}}{K_a} + \hat{\sigma}^2_{\beta_0} - 2\frac{s_{YW_a}\hat{\omega}_a}{K_a}.
\]  \hspace{1cm} (3.110)

Assuming normality of the error terms the distribution of \( \hat{d}_a \) is

\[
\frac{\hat{d}_a - d_a}{\hat{\sigma}_{da}} \sim T(\nu_{da}),
\]  \hspace{1cm} (3.111)

where \( \nu_{da} \) can be estimated with the Welch-Satterthwaite approximation, as shown in (3.41) adding the corresponding subindex.

An approximate CI for \( d_a \) is

\[
\hat{d}_a \pm t(\alpha/2, \nu_{da}) \hat{\sigma}_{da},
\]  \hspace{1cm} (3.112)

In the case of the confidence interval containing zero, the interpretation is that the true value of the compound in the study sample is below the detection limit of the analytical method.
The point estimators, the distribution and approximate confidence interval of $\sigma^2_{Va}$ remain as described in Section 3.1.3.1.

### 3.3.3.2 Examples

Table 3.11 applies the results above to the data in Table 3.2. The first seven columns show the interval estimates for the observed values and the point and interval estimates for the concentration of each compound in the study sample, assuming the errors are independent. The last four columns show the same estimates under correlated errors but using separate correlations for each compound. For this purpose we use the data in Table 3.1, the estimates in Table 3.10 and (3.102), (3.109) and (3.112). These results differ from those in Tables 3.5 and 3.6 in assuming a constant bias across compounds.

We observe that the variances of the study values and their degrees of freedom are increased in such a way that most of the confidence intervals are narrower than those of the observed data (16 out of 26).

The average correlation coefficient $\hat{\rho}$ becomes $-0.074$. This correlation value is small enough that it makes no significant difference in the final estimates of $V[\hat{d}_a]$ and the interval estimation of $\hat{d}_a$ for the specific participating method. Using this constant correlation coefficient we show the estimates in Table 3.12.

Since the test for constant within-method variance among compounds was rejected we still assume it is not constant compound-wise and the estimates of the within-material variance for the study material remain the same shown in Table 3.5.
Table 3.11. Example of study values corrected for bias using control data assuming independent errors and correlated errors, assuming a constant bias across compounds.

For each compound, $lcl(W_a) = \text{the lower confidence limit for the mean study value}$, $ucl(W_a) = \text{the upper confidence limit for the mean study value}$, $\hat{d}_a = \text{the study data corrected for bias using control data, under independent errors}$: $\hat{\sigma}_{da}^2 = \text{the variance of the study data corrected for bias}$, $\nu_{\hat{d}_a} = \text{the degrees of freedom associated to the variance of the study data corrected for bias}$, $lcl(\hat{d}_a) = \text{the lower confidence limit for the mean of the study data corrected for bias}$, $ucl(\hat{d}_a) = \text{the upper confidence limit for the mean of the study data corrected for bias}$. The last four columns list the corresponding estimates under correlated errors.
Table 3.12. Example of adjusted study values when errors are correlated, constant bias and constant correlation across compounds.

<table>
<thead>
<tr>
<th>Compounds</th>
<th>$\hat{d}_a$</th>
<th>$\hat{\sigma}^2_{da}$</th>
<th>$\nu_{\hat{\sigma}_{da}}$</th>
<th>$lcl_{\hat{d}_a}$</th>
<th>$ucl_{\hat{d}_a}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2,3,5-trimethylnaphthalene</td>
<td>25.7</td>
<td>1.07</td>
<td>177.1</td>
<td>23.6</td>
<td>27.7</td>
</tr>
<tr>
<td>acenaphthene</td>
<td>34.7</td>
<td>1.04</td>
<td>253.7</td>
<td>32.7</td>
<td>36.7</td>
</tr>
<tr>
<td>dibenz[a,h]anthracene</td>
<td>112.0</td>
<td>6.74</td>
<td>2.7</td>
<td>103.0</td>
<td>121.0</td>
</tr>
<tr>
<td>acenaphthylene</td>
<td>29.6</td>
<td>1.59</td>
<td>12.5</td>
<td>26.9</td>
<td>32.3</td>
</tr>
<tr>
<td>biphenyl</td>
<td>29.2</td>
<td>1.13</td>
<td>83.4</td>
<td>27.1</td>
<td>31.3</td>
</tr>
<tr>
<td>1-methylphenanthrene</td>
<td>69.4</td>
<td>1.09</td>
<td>133.7</td>
<td>67.3</td>
<td>71.5</td>
</tr>
<tr>
<td>2,6-dimethylnaphthalene</td>
<td>35.2</td>
<td>1.04</td>
<td>221.1</td>
<td>33.2</td>
<td>37.2</td>
</tr>
<tr>
<td>fluorene</td>
<td>65.5</td>
<td>1.12</td>
<td>95.0</td>
<td>63.4</td>
<td>67.6</td>
</tr>
<tr>
<td>triphenylene</td>
<td>141.0</td>
<td>25.19</td>
<td>2.2</td>
<td>121.0</td>
<td>161.0</td>
</tr>
<tr>
<td>1-methylnaphthalene</td>
<td>49.0</td>
<td>1.24</td>
<td>38.7</td>
<td>46.7</td>
<td>51.2</td>
</tr>
<tr>
<td>anthracene</td>
<td>172.3</td>
<td>3.08</td>
<td>4.2</td>
<td>168.0</td>
<td>177.0</td>
</tr>
<tr>
<td>benzo[j]fluoranthene</td>
<td>281.0</td>
<td>25.16</td>
<td>2.2</td>
<td>261.0</td>
<td>301.0</td>
</tr>
<tr>
<td>benzo[k]fluoranthene</td>
<td>406.0</td>
<td>26.32</td>
<td>2.2</td>
<td>385.0</td>
<td>427.0</td>
</tr>
<tr>
<td>2-methylnaphthalene</td>
<td>79.2</td>
<td>1.18</td>
<td>58.3</td>
<td>77.0</td>
<td>81.4</td>
</tr>
<tr>
<td>chrysene</td>
<td>547.0</td>
<td>5.75</td>
<td>2.9</td>
<td>539.0</td>
<td>555.0</td>
</tr>
<tr>
<td>benzo[e]pyrene</td>
<td>590.0</td>
<td>49.09</td>
<td>2.1</td>
<td>561.0</td>
<td>619.0</td>
</tr>
<tr>
<td>benzo[a]anthracene</td>
<td>378.0</td>
<td>6.41</td>
<td>2.8</td>
<td>370.0</td>
<td>386.0</td>
</tr>
<tr>
<td>benzo[ghi]perylene</td>
<td>488.0</td>
<td>9.30</td>
<td>2.5</td>
<td>477.0</td>
<td>499.0</td>
</tr>
<tr>
<td>indeno[1,2,3-cd]pyrene</td>
<td>608.0</td>
<td>21.96</td>
<td>2.2</td>
<td>589.0</td>
<td>627.0</td>
</tr>
<tr>
<td>benzo[a]pyrene</td>
<td>636.0</td>
<td>11.75</td>
<td>2.4</td>
<td>623.0</td>
<td>649.0</td>
</tr>
<tr>
<td>perylene</td>
<td>184.0</td>
<td>17.30</td>
<td>2.2</td>
<td>168.0</td>
<td>200.0</td>
</tr>
<tr>
<td>phenanthrene</td>
<td>539.0</td>
<td>17.30</td>
<td>2.2</td>
<td>523.0</td>
<td>555.0</td>
</tr>
<tr>
<td>benzo[b]fluoranthene</td>
<td>1089.0</td>
<td>127.79</td>
<td>2.0</td>
<td>1040.0</td>
<td>1140.0</td>
</tr>
<tr>
<td>pyrene</td>
<td>1129.0</td>
<td>63.98</td>
<td>2.1</td>
<td>1100.0</td>
<td>1160.0</td>
</tr>
<tr>
<td>fluoranthene</td>
<td>1094.0</td>
<td>31.98</td>
<td>2.1</td>
<td>1070.0</td>
<td>1120.0</td>
</tr>
<tr>
<td>naphthalene</td>
<td>121.3</td>
<td>3.07</td>
<td>4.2</td>
<td>117.0</td>
<td>126.0</td>
</tr>
</tbody>
</table>

For each compound, $\hat{d}_a =$ the study data corrected for bias using control data, $\hat{\sigma}^2_{da} =$ the variance of the study data corrected for bias, $\nu_{\hat{\sigma}_{da}} =$ the degrees of freedom associated to the variance of the study data corrected for bias, $lcl_{\hat{d}_a} =$ the lower confidence limit for the mean of the study data corrected for bias and $ucl_{\hat{d}_a} =$ the upper confidence limit for the mean of the study data corrected for bias, under correlated errors with constant correlation across compounds.
3.4 One laboratory, one compound at several points in time

This section examines the bias and within method variance of a laboratory over time on a single compound. Explicit modeling of the bias as a function of time and or concentration will be discussed in Chapter 4.

Here the primary goal is a general test of constant bias over time and to show a way to estimate a common bias if it is constant over time. This section is very similar to the preceding section looking at different compounds at one point in time except that:

- the independent variable is time, that replaces the former variable the compound,
- that time has a natural order of sequence while the compounds do not,
- the possible reuse of the same reference material over time.

Working with one laboratory with just one compound at several points in time, the notation can be simplified by just dropping the related indexes to the corresponding lab and compound but retaining the time and replicate indexes.

3.4.1 The Basic Model

The model for the observed value of the compound in the control material over time is:

\[ Y_{tk} = \beta_{0t} + x_t + U_{tk} + \epsilon_{tk} \]  

where \( t = 1, \ldots, T \), \( k = 1, \ldots, K_t \), \( E[U_{tk}] = 0 \), \( V[U_{tk}] = \sigma^2_{Ut} \), \( E[\epsilon_{tk}] = 0 \), \( V[\epsilon_{tk}] = \sigma^2_{\epsilon t} \) and \((\beta_{0t}, x_t, \sigma^2_{Ut}, \sigma^2_{\epsilon t})\) are unknown parameters. The \( U_{tk} \) and \( \epsilon_{tk} \) errors are assumed independent for a given \( t, k \) and are independent among \( t, k \). Assume that \( \hat{x}_t \), \( \hat{\sigma}^2_{Ut} \) and \( \hat{\sigma}^2_{\epsilon t} \) are known unbiased estimators for \( x_t \), \( \sigma^2_{Ut} \), and \( V[\hat{x}_t] \) respectively; see Section 3.2.

The model for the observed value of the compound in the study material over time is
\[ W_{tk} = \beta_{0t} + d_t + V_{tk} + \delta_{tk} \quad (3.114) \]

where \( t = 1, \ldots, T, k = 1, \ldots, K_t \), \( E[V_{tk}] = 0, V[V_{tk}] = \sigma^2_{Vt}, E[\delta_{tk}] = 0, V[\delta_{tk}] = \sigma^2_{\epsilon t} \) and \((d_t, \sigma^2_{Vt})\) are unknown parameters. The \( \beta_{0t} \) are shared by both models and note that they depend on time. The \( V_{tk} \) and \( \delta_{tk} \) terms are independent for a given \( t, k \) and are independent among \( t, k \). Also assume that \( \delta_{tk}, V_{tk} \) are independent from \( \epsilon_{tk}, U_{tk} \).

Then

\[ E[Y_{tk}] = \beta_{0t} + x_t, V[Y_{tk}] = \sigma^2_{Ut} + \sigma^2_{\epsilon t} \quad (3.115) \]

and

\[ E[W_{tk}] = \beta_{0t} + d_t, V[W_{tk}] = \sigma^2_{Vt} + \sigma^2_{\epsilon t}. \quad (3.116) \]

We have assumed that the observed values are independent, but now the same CRM may be repeatedly used over time and in consequence some of the true reference values \( x_t \) may stand for the same underlying true reference value \( x_r \). Let \( R \leq T \) different CRMs be used over time, with conventional reference values \((\hat{x}_r, \hat{\sigma}^2_{x_r}, \nu_{\hat{x}_r})\), \( r = 1..R \). Let \( a_{tr} \) be defined as:

\[
a_{tr} = \begin{cases} 
1, & \text{if the } r^{th} \text{ CRM is used at time } t, \\
0, & \text{otherwise},
\end{cases} \quad (3.117)
\]

then the conventional reference value used at time \( t \) is

\[ \hat{x}_t = \sum_{r=1}^{R} a_{tr} \hat{x}_r. \quad (3.118) \]

Define the covariance matrices as

\[ V[\hat{Y}] = diag(V[\hat{Y}_1], ..., V[\hat{Y}_T]) \text{ and } V[\hat{x}] = diag(V[\hat{x}_1], ..., V[\hat{x}_R]). \quad (3.119) \]
Then the covariance of the conventional reference values used over time is

\[
Cov[\hat{x}_t, \hat{x}_{t'}] = Cov \left[ \sum_{r=1}^{R} a_{tr} \hat{x}_r, \sum_{r'=1}^{R} a_{t'r'} \hat{x}_{t'} \right] = \sum_{r=1}^{R} a_{tr} a_{t'r'} \sigma_{\hat{x}_r}^2. \tag{3.120}
\]

Let \( \hat{x} = (\hat{x}_1, ..., \hat{x}_R)' \) be the vector of different reference values, \( \bar{Y} = (\bar{Y}_1, ..., \bar{Y}_T)' \) be the vector of observed values, \( A \) be the \([T \times R]\) matrix with entries defined by \( a_{tr} \) in (3.117) and \( \hat{\beta}_0 = (\hat{\beta}_{01}, ..., \hat{\beta}_{0T}) \) be the vector of separate estimates of bias over time.

Then

\[
\hat{\beta}_0 = \bar{Y} - A \hat{x},
\]

\[
A = \begin{bmatrix}
a_{11} & \cdots & a_{1R} \\
a_{21} & \cdots & a_{2R} \\
\vdots & \ddots & \vdots \\
a_{T1} & \cdots & a_{TR}
\end{bmatrix} \tag{3.121}
\]

and

\[
\Sigma_{\beta_0} = Cov[\hat{\beta}_0] = V[\bar{Y}] + AV[\hat{x}]A'.
\tag{3.122}
\]

Using (3.122) we can estimate the covariance matrix as

\[
\hat{\Sigma}_{\beta_0} = \hat{V}[\bar{Y}] + \hat{A}V[\hat{x}]A'
\tag{3.123}
\]

with \((s^2_{Y_t}/K_t, \hat{\sigma}_{\hat{x}_r}^2)\) the estimators for \((V[\bar{Y}_t], V[\hat{x}_r])\) respectively.

### 3.4.2 Working with the Control Data Only

Here there is no interest on pooling information across different participants. However, there is interest on testing whether a participant’s method is consistently biased over time. When working with the control data only the performance of the participants can be assessed directly, provided that the conventional value of the control material and its related variance components are known.
If $\beta_0$ and $\sigma^2_{\epsilon}$ are different for each point in time then unbiased estimates for $\beta_0$ and $\sigma^2_{\epsilon}$ can be obtained by using the results from Section 3.1.2.1, specifically using (3.4) and (3.5).

The results in this section match those from the case with multiple compounds at one point in time where we assumed constancy across compounds for both the bias and the within variance of the method. For this reason we will mainly remark only on the differences below.

### 3.4.2.1 Estimation and inferences

As we reviewed in the case of multiple compounds at one point in time, the estimation of a global parameter makes sense only if the assumptions hold. Thus we will test for the assumptions followed by the parameters estimation.

- **Testing for constancy of the bias**

  We can reuse the results we found for the case of one laboratory measuring multiple compounds at one point in time described in Section 3.3.2.2.

  In order to test that all of the biases are equal over time (i.e., $H_0 : \beta_0 = \beta_0, \forall t$) we use $\chi^2 = (C\hat{\beta}_0)'^{(C\hat{\Sigma}_{\hat{\beta}_0}C')^{-1}(C\hat{\beta}_0)}$ where $C$ is a $(T - 1) \times T$ matrix of independent contrasts and $\hat{\Sigma}_{\hat{\beta}_0}$ as in (3.123), then under normality $\chi^2$ is approximately a $\chi^2_T$ random variable.

  $H_0$ can be rejected at a significance level of approximately $\alpha$ if $\chi^2 > \chi^2_{(1-\alpha;T)}$.

  The approximation comes from using an estimated covariance matrix for $\hat{\beta}_0$.

- **Testing for constant within laboratory variances**

  Similarly the assumption of constant within method variances can be tested by using the Levene’s Test as described in (3.100) for homogeneous and heterogeneous control material and using the test statistic in (3.101) for both homogeneous and heterogeneous control materials.
• *Estimation of a constant bias over time*

Let us assume that the method used by the participant is stable enough such that the parameters that are related to the method can be considered as fixed constants over time \( E[\hat{\beta}_{0t}] = \beta_0 \). That is, the test for constancy of the bias failed to reject the hypothesis.

Under the assumption of constant bias shared over time, \( \beta_{0t} = \beta_0 \). Consider the point estimators:

*Simple average*

Define the simple average of the biases over time

\[
\bar{\beta}_0 = \frac{1}{T} \sum_{t=1}^{T} \hat{\beta}_{0t} = \frac{1}{T} \hat{\beta}_0^\prime. 
\]  

(3.124)

Then

\[
E[\bar{\beta}_0] = E \left[ \frac{1}{T} \sum_{t=1}^{T} \hat{\beta}_{0t} \right] = \frac{1}{T} \hat{\beta}_0^\prime E[\hat{\beta}_0] = \beta_0, \]

\[
V[\bar{\beta}_0] = \frac{1}{T^2} \hat{\Sigma}_{\hat{\beta}_0}^1 \]

(3.125)

where \( \hat{\Sigma}_{\hat{\beta}_0} \) is given in (3.122). \( V[\bar{\beta}_0] \) can be estimated as

\[
\hat{\sigma}^2_{\bar{\beta}_0} = \frac{1}{T^2} \hat{\Sigma}_{\hat{\beta}_0}^1. \]

(3.126)

Hence \( \bar{\beta}_0 \) as defined in (3.124) is unbiased for \( \beta_0 \) under normality. If the distribution of the error terms is not normal then the sample variance and sample mean may not be independent and in general (3.124) becomes an approximation.
**Weighted average**

Define the weighted average of the biases over time

$$\bar{\beta}_0 \omega = \sum_{t=1}^{T} \hat{\omega}_t \hat{\beta}_{0t} = \hat{\omega}' \hat{\beta}_0$$

(3.127)

where $\hat{\beta}_{0t} = \bar{Y}_t - \hat{x}_t$, $\hat{\omega}_t = \hat{\sigma}_{\beta_0t}^{-2} / \sum_t \hat{\sigma}_{\beta_0t}^{-2}$, $\hat{\omega} = [\hat{\omega}_1, \ldots, \hat{\omega}_T]'$ and $\hat{\omega}'1 = 1$.

Under normality these are independent and using double expectation (e.g. Casella & Berger 2002 [11, Theorem 4.4.3.])

$$E[\bar{\beta}_0 \omega] = E[\hat{\omega}' \hat{\beta}_0] = \beta_0 E[\hat{\omega}'] = \beta_0.$$  

(3.128)

Hence $\bar{\beta}_0 \omega$ as defined in (3.127) is unbiased for $\beta_0$ under normality. If the distribution of the error terms is not normal then the sample variance and sample mean may not be independent and in general (3.128) becomes an approximation.

Note that

$$\bar{\beta}_0 \omega = \hat{\omega}' \hat{\beta}_0 = \hat{\omega}' (\bar{Y} - A\hat{x}).$$  

(3.129)

Now assuming normality and using the conditional variance identity (e.g., Casella & Berger 2002, Theorem 4.4.7, [11]) we have

$$V[\bar{\beta}_0 \omega] = V[\hat{\omega}' \hat{\beta}_0] = V \left[ E \left[ \hat{\omega}' \hat{\beta}_0 | \hat{\omega} \right] \right] + E \left[ V \left[ \hat{\omega}' \hat{\beta}_0 | \hat{\omega} \right] \right]$$

$$= V \left[ \hat{\omega}' E \left[ \hat{\beta}_0 | \hat{\omega} \right] \right] + E \left[ \hat{\omega}' V \left[ \hat{\beta}_0 | \hat{\omega} \right] \hat{\omega} \right]$$

$$= V \left[ \hat{\omega}' \beta_0 1 \right] + E \left[ \hat{\omega}' \Sigma_{\beta_0} \hat{\omega} \right]$$

$$= V \left[ \beta_0 \right] + E \left[ \hat{\omega}' \Sigma_{\beta_0} E \left[ \hat{\omega} \right] \right] = E \left[ \hat{\omega}' \Sigma_{\beta_0} E \left[ \hat{\omega} \right] \right]$$

(3.130)

where $V[\bar{Y}]$ and $V[\hat{x}]$ are given in (3.119).
For the case where all the reference materials are different then $A = I$ the identity matrix and (3.130) takes the form of (3.78) replacing the compound suffix $a$ with the time suffix $t$.

We can obtain an approximation for $V[\tilde{\beta}_0]$ or $V[\tilde{\beta}_{0\omega}]$ by plugging $(s^2_t/K_t, \hat{\sigma}^2_{\tilde{x}_r})$, the estimators for $(V[\tilde{Y}_t], V[\tilde{x}_r])$ respectively, into (3.125) or (3.130) respectively and dropping the expectation operator, obtaining

$$\hat{\sigma}^2_{\tilde{\beta}_{0\omega}} = \hat{\omega}' \hat{\Sigma}_{\tilde{\beta}_0} \hat{\omega}. \tag{3.131}$$

This is the same as treating $\hat{\omega}_t$ as fixed.

Some additional distributional assumptions on the error terms are required in order to get confidence intervals for $\beta_0$. Assuming normality of $U_{tk}$ and $\epsilon_{tk}$, the distribution of $\beta_0$ is

$$\frac{\hat{\beta}_0 - \beta_0}{\hat{\sigma}_{\tilde{\beta}_0}} \sim T(\nu_{\tilde{\beta}_0}), \tag{3.132}$$

where $\hat{\beta}_0$ and $\hat{\sigma}_{\tilde{\beta}_0}$ are given in (3.124) and (3.126) for simple average of the constant biases or (3.129) and (3.131) for weighted average of the constant biases. The respective degrees of freedom $\nu_{\tilde{\beta}_0}$ can be estimated by the Welch-Satterthwaite approximation as

$$\nu_{\tilde{\beta}_0} = \frac{2(\hat{\sigma}^2_{\tilde{\beta}_0})^2}{\hat{V}[\hat{\sigma}^2_{\tilde{\beta}_0}]} \tag{3.133}$$

where $\hat{V}[\hat{\sigma}^2_{\tilde{\beta}_0}]$ takes into account the covariance due to the CRM’s reuse over time.

An approximate $100(1 - \alpha)$% confidence interval for $\beta_0$ under normality is then

$$\hat{\beta}_0 \pm t(\alpha/2, \nu_{\tilde{\beta}_0}) \hat{\sigma}_{\tilde{\beta}_0}. \tag{3.134}$$
• **Testing for zero bias**

We can apply exactly the same results we found for the case of one laboratory measuring multiple compounds at one point in time described in Section 3.3.2.2 considering the updated variance and degrees of freedom using (3.131) or simply by testing if the CI in (3.134) contains zero. This test assumes there is no different bias over time, otherwise this test is invalid.

• **Performance evaluation for unbiasedness**

We can proceed as in the previous section, by using the \( t\) score to assess the performance of the method for unbiasedness.

• **Estimation of a constant within method variance over time**

Assuming normality and independence of the error terms, and constant within method variance over time, \( \sigma_{et|x}^2 = \sigma_{et|x}^2 = h(x, \theta) \), with \( h \) unspecified,

\[ \diamond \] Under the assumption that the control materials are homogeneous (i.e., \( \sigma_{Ut}^2 = 0, \forall t \)), we have independent \( \hat{\sigma}_{et}^2 \) and we can refer to the results in Section 3.3.2.1, specifically using (3.80), (3.81) and (3.82), replacing the compound index \( a \) with the time index \( t \).

Using all observations together we can obtain a \( 100(1 - \alpha)\% \) confidence interval for \( \sigma_{e}^2 \) by

\[
\left( \frac{\hat{\sigma}_{e}^2 \sum (K_t - 1)}{\chi^2_{1-\alpha/2; \sum (K_t - 1)}} , \frac{\hat{\sigma}_{e}^2 \sum (K_t - 1)}{\chi^2_{\alpha/2; \sum (K_t - 1)}} \right). \tag{3.135}
\]

\[ \diamond \] Under the assumption that the control materials are heterogeneous (i.e., \( \exists t \) such that \( \sigma_{Ut}^2 \neq 0 \)). Assume we have estimates for \( \nu_{et} = \) degrees of freedom of \( \hat{\sigma}_{et}^2 \), then we have dependent \( \hat{\sigma}_{et}^2 \) distributed

\[
\frac{\nu_{et}\hat{\sigma}_{ut}^2}{\sigma_{e}^2} \sim \chi^2_{(\nu_{et})} \tag{3.136}
\]

and
\[
\sum_{t=1}^{T} \frac{\nu_t \hat{\sigma}_{et}^2}{\sigma_t^2} \sim \chi^2_{(\nu_t)} \tag{3.137}
\]

for some \( \nu_t \leq \sum \nu_{et} \), estimated with the Welch-Satterthwaite approximation. Define

\[
\hat{\sigma}_t^2 = \frac{\sum_{t=1}^{T} \nu_{et} \hat{\sigma}_{et}^2}{\nu_t} \tag{3.138}
\]

then

\[
E[\hat{\sigma}_t^2] = E\left[\frac{\sum_{t=1}^{T} \nu_{et} \sigma_{et}^2}{\nu_t}\right] = \frac{E[\sum_{t=1}^{T} \nu_{et} \hat{\sigma}_{et}^2]}{\nu_t} = \sigma_t^2. \tag{3.139}
\]

Hence \( \hat{\sigma}_t^2 \) as defined in (3.138) is unbiased for \( \sigma_t^2 \) under the corresponding assumptions.

Using all observations together we can obtain a 100(1 − \( \alpha \))% confidence interval for \( \sigma_t^2 \) by

\[
\left( \frac{\hat{\sigma}_t^2 \sum \nu_{et}}{\chi^2_{(1-\alpha/2,\nu_t)}}, \frac{\hat{\sigma}_t^2 \sum \nu_{et}}{\chi^2_{(\alpha/2,\nu_t)}} \right) \tag{3.140}
\]

provided that the estimates of \( \nu_{et} \) are known.

### 3.4.2.2 Examples

Table 3.13 lists the summary data for the benzo[ghi]perylene compound measured by one laboratory in sediment samples over time. In this case, three different CRMs are used as control material, the groups using the same CRM are separated by a horizontal line. The separate estimates and CIs for each point in time are listed in Table 3.14. Figure 3.7 shows the separate estimates of the bias at each point in time.

We have no evidence to reject the null hypothesis of an unbiased method at each point in time. Note each 95% CI contains zero. This suggests the method is conditionally unbiased at each time. However the CI’s are very wide indicating poor power for testing for unbiasedness.

In order to test for constant bias over time we find the value \( \chi^2_t = 36.03 \) and \( \chi^2_{(0.95;T=10)} = 23.21 \), since \( \chi^2_t > \chi^2_{(0.95,10)} \) then there is evidence of unequal biases.
over time. Note that the \( \hat{\beta}_{0t} \) are not independent since they may be sharing the same reference material as described by (3.123). When contrasting a pair of them a covariance term appears and modifies the final result. The assumption of constant method bias is inadequate.

<table>
<thead>
<tr>
<th>Year</th>
<th>( t )</th>
<th>( \bar{Y}_t )</th>
<th>( s_{Y_t} )</th>
<th>( K_t )</th>
<th>( W_t )</th>
<th>( s_{W_t} )</th>
<th>( M_t )</th>
<th>( \hat{\rho}_{YW_t} )</th>
<th>( \hat{x}_t )</th>
<th>( \hat{\sigma}_{\hat{x}_t} )</th>
<th>( \nu_{\hat{x}_t} )</th>
<th>( \hat{\sigma}_{\text{wt}} )</th>
<th>( \nu_{\text{wt}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1994</td>
<td>533</td>
<td>4.73</td>
<td>3</td>
<td>552</td>
<td>10.69</td>
<td>3</td>
<td>-0.99</td>
<td>525</td>
<td>34.18</td>
<td>33</td>
<td>32.01</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>1995</td>
<td>517</td>
<td>18.19</td>
<td>3</td>
<td>2750</td>
<td>56.52</td>
<td>3</td>
<td>0.05</td>
<td>525</td>
<td>34.18</td>
<td>33</td>
<td>32.01</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>1996</td>
<td>531</td>
<td>23.63</td>
<td>3</td>
<td>829</td>
<td>28.10</td>
<td>3</td>
<td>-0.92</td>
<td>525</td>
<td>34.18</td>
<td>33</td>
<td>32.01</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>1997</td>
<td>526</td>
<td>2.08</td>
<td>3</td>
<td>734</td>
<td>19.22</td>
<td>3</td>
<td>-0.09</td>
<td>525</td>
<td>34.18</td>
<td>33</td>
<td>32.01</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>1998</td>
<td>484</td>
<td>20.55</td>
<td>3</td>
<td>903</td>
<td>13.05</td>
<td>3</td>
<td>-0.94</td>
<td>525</td>
<td>34.18</td>
<td>33</td>
<td>32.01</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>1999</td>
<td>519</td>
<td>10.41</td>
<td>3</td>
<td>186</td>
<td>4.16</td>
<td>3</td>
<td>0.88</td>
<td>525</td>
<td>34.18</td>
<td>33</td>
<td>32.01</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>2846</td>
<td>17.06</td>
<td>3</td>
<td>930</td>
<td>17.21</td>
<td>3</td>
<td>-0.85</td>
<td>2840</td>
<td>51.02</td>
<td>45</td>
<td>70.77</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td>2794</td>
<td>48.69</td>
<td>3</td>
<td>653</td>
<td>7.09</td>
<td>3</td>
<td>-1.00</td>
<td>2840</td>
<td>51.02</td>
<td>45</td>
<td>70.77</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>2003</td>
<td>327</td>
<td>5.51</td>
<td>3</td>
<td>488</td>
<td>5.00</td>
<td>3</td>
<td>0.45</td>
<td>307</td>
<td>22.96</td>
<td>36</td>
<td>23.08</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>2005</td>
<td>312</td>
<td>5.51</td>
<td>3</td>
<td>316</td>
<td>5.03</td>
<td>3</td>
<td>-0.44</td>
<td>307</td>
<td>22.96</td>
<td>36</td>
<td>23.08</td>
<td>11</td>
<td></td>
</tr>
</tbody>
</table>

**Table 3.13.** Example of summary data reported over time.

For each time (year), \( \bar{Y}_t \) = mean of control values, \( s_{Y_t} \) = standard deviation of control value, \( K_t \) = sample size of the control value, \( W_t \) = mean of study values, \( s_{W_t} \) = standard deviation of study value, \( M_t \) = sample size of the study value, \( \hat{\rho}_{YW_t} \) = correlation coefficient between the control and the study data, \( \hat{x}_t \) = the reference value, \( \hat{\sigma}_{\hat{x}_t} \) = uncertainty of the reference value, \( \nu_{\hat{x}_t} \) = the degrees of freedom associated to the uncertainty of the reference value, \( \hat{\sigma}_{\text{wt}} \) = within-method uncertainty of the reference value and \( \nu_{\text{wt}} \) = the degrees of freedom associated to the within-method uncertainty of the reference value.

The separate estimates and CI's for each within method variance \( \sigma^2_{ct} \) are also listed in Table 3.14 and displayed in Figure 3.8. We observe that the estimates of the within method variance for 1997 and 2002 differ; this suggests the within method variance is not constant over time.

Similarly for testing the assumption of constant within laboratory variances we use the Levene Test as described in Section 3.3.2.2. Using the detailed data from the participant and the data from the previous section we obtain: \( f^* = 3.2206 \), \( f_{(0.95;9,20)} = 2.3928 \). There is enough evidence to reject \( H_0 \) and to conclude the within
Table 3.14. Method bias and within variance estimated separately over time using the control data.

<table>
<thead>
<tr>
<th>Year</th>
<th>$\hat{\beta}_0t$</th>
<th>$\hat{\sigma}_{\hat{\beta}0t}$</th>
<th>$\nu_{\hat{\beta}0t}$</th>
<th>$lcl_{\hat{\beta}0t}$</th>
<th>$ucl_{\hat{\beta}0t}$</th>
<th>$\hat{\sigma}_{\epsilon t}^2$</th>
<th>$K_t$</th>
<th>$lcl_{\sigma_{\epsilon t}^2}$</th>
<th>$ucl_{\sigma_{\epsilon t}^2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1994</td>
<td>7.67</td>
<td>34.3</td>
<td>33</td>
<td>-62.1</td>
<td>77.4</td>
<td>22.3</td>
<td>3</td>
<td>0.56</td>
<td>82.38</td>
</tr>
<tr>
<td>1995</td>
<td>-8</td>
<td>35.8</td>
<td>34</td>
<td>-80.6</td>
<td>64.6</td>
<td>331</td>
<td>3</td>
<td>8.38</td>
<td>1221.02</td>
</tr>
<tr>
<td>1996</td>
<td>5.67</td>
<td>36.8</td>
<td>31</td>
<td>-69.4</td>
<td>80.7</td>
<td>558</td>
<td>3</td>
<td>14.14</td>
<td>2059.62</td>
</tr>
<tr>
<td>1997</td>
<td>1.33</td>
<td>34.2</td>
<td>33</td>
<td>-68.3</td>
<td>70.9</td>
<td>4.33</td>
<td>3</td>
<td>0.11</td>
<td>15.99</td>
</tr>
<tr>
<td>1998</td>
<td>-41.3</td>
<td>36.2</td>
<td>33</td>
<td>-115</td>
<td>32.2</td>
<td>422</td>
<td>3</td>
<td>10.69</td>
<td>1557.94</td>
</tr>
<tr>
<td>1999</td>
<td>-6.33</td>
<td>34.7</td>
<td>35</td>
<td>-76.8</td>
<td>64.2</td>
<td>108</td>
<td>3</td>
<td>2.74</td>
<td>399.63</td>
</tr>
<tr>
<td>2000</td>
<td>6</td>
<td>52</td>
<td>29</td>
<td>-100</td>
<td>112</td>
<td>291</td>
<td>3</td>
<td>7.37</td>
<td>1073.46</td>
</tr>
<tr>
<td>2002</td>
<td>-46</td>
<td>58.3</td>
<td>20</td>
<td>-167</td>
<td>75.3</td>
<td>2370</td>
<td>3</td>
<td>60.03</td>
<td>8746.33</td>
</tr>
<tr>
<td>2003</td>
<td>20.3</td>
<td>23.2</td>
<td>37</td>
<td>-26.6</td>
<td>67.3</td>
<td>30.3</td>
<td>3</td>
<td>0.77</td>
<td>111.90</td>
</tr>
<tr>
<td>2005</td>
<td>4.67</td>
<td>23.2</td>
<td>37</td>
<td>-42.3</td>
<td>51.6</td>
<td>30.3</td>
<td>3</td>
<td>0.77</td>
<td>111.90</td>
</tr>
</tbody>
</table>

For each point in time (year), $\hat{\beta}_0t$ = method bias, $\hat{\sigma}_{\hat{\beta}0t}$ = standard deviation of the bias, $\nu_{\hat{\beta}0t}$ = degrees of freedom of the standard deviation of the bias, $lcl_{\hat{\beta}0t}$ = lower confidence limit for the bias, $ucl_{\hat{\beta}0t}$ = upper confidence limit for the bias, $\hat{\sigma}_{\epsilon t}^2$ = within-method variance, $K_t$ = sample size of the control data, $lcl_{\sigma_{\epsilon t}^2}$ = lower confidence limit for the within-method variance, $ucl_{\sigma_{\epsilon t}^2}$ = upper confidence limit for the within-method variance.

Variance is not constant over time, our assumption of constant within method variance is inadequate and in general $\exists t \neq t' : \sigma_{\epsilon t}^2 \neq \sigma_{\epsilon t'}^2$.

Since both hypotheses (constant bias over time and constant within method variance) are rejected, pooled estimates are no longer pursued.
Figure 3.7. Point and interval estimates of the method bias over time.

This plot shows the separate estimates of the biases ($\hat{\beta}_0$) for one compound over time and the overall estimate of the bias $\hat{\beta}_0$ from the benzo[ghi]perylene in the Sediments exercise for one of the participants. No constant bias over time is shown since the hypothesis test for constant bias was rejected. The variance of the estimated bias does not show any apparent pattern.
Figure 3.8. Point and interval estimates of the within-method variance over time.

This plot shows the separate estimates of the within-method variance ($\hat{\sigma}_{\epsilon}$) for one compound over time and the constant estimate of the within variance $\hat{\sigma}_{\epsilon}$ from the benzo[ghi]perylene in the Sediments exercise for one of the participants. No confidence interval is shown for a constant within-method variance since the hypothesis test for an equal within-method variance was rejected.
3.4.3 Exporting the Model to the Study Data

The models found while working with the control data can be used to estimate the unknown parameters of the study data.

3.4.3.1 Parameter Estimation

The focus is on estimating the true values of the study sample \( d_t \) under the constant bias and constant within laboratory variance over time. The only change is the use of the pooled bias estimate \( \hat{\beta}_0 \) and the pooled within laboratory variance estimate \( \hat{\sigma}_t^2 \) if the tests for constant bias of the method and constant within laboratory variance are not rejected. Otherwise we can proceed one point in time separately as described in the Section 3.1.3 for one compound at a time. The approach is very much like that in Section 3.3.3.

Independent errors

Assume first that the error terms are independent for the method measuring the control sample and the study sample. This implies \( \text{Corr}[\epsilon_{tk}, \delta_{tk}] = \rho_t = 0 \). Let

\[
\hat{d}_t = W_t - \hat{\beta}_0
\]  

(3.141)

and

\[
\hat{\sigma}_t^2 = s_{Wt}^2 - \hat{\sigma}_{\epsilon t}^2.
\]  

(3.142)

Then

\[
E[\hat{d}_t] = E[W_t - \hat{\beta}_0] = d_t + \beta_0 - E[\hat{\beta}_0] = d_t
\]  

(3.143)

and

\[
V[\hat{d}_t] = V[W_t] + V[\hat{\beta}_0] - 2\text{Cov}[W_t, \hat{\beta}_0] = V[W_t] + V[\hat{\beta}_0]
\]  

(3.144)

which can be estimated by \( \hat{\sigma}_{dt}^2 = \frac{s_{Wt}^2}{M_t} + \hat{\sigma}_{\beta_0}^2 \). Also

\[
E[\hat{\sigma}_t^2] = E[s_{Wt}^2 - \hat{\sigma}_{\epsilon t}^2].
\]  

(3.145)

The distribution of \( \hat{d}_t \) is
\[
\frac{\hat{d}_t - d_t}{\hat{\sigma}_{dt}} \sim T(\nu_{\hat{d}t})
\] (3.146)

and the approximate confidence interval for \(d_t\) then becomes

\[
\hat{d}_t \pm t(\alpha/2;\nu_{\hat{d}t})\hat{\sigma}_{dt},
\] (3.147)

where \(\nu_{\hat{d}t}\) can be estimated with the Welch-Satterthwaite approximation. For estimating \(\sigma^2_V\) we can refer to the results in Section 3.3.3.1 and consider replacing compound index \(a\) with time index \(t\).

**Correlated errors**

Now assume that the error terms are correlated for the method measuring the control sample and the study sample, this is \(K_t = M_t\) with pairing and \(\text{Corr}[\epsilon_{tk}, \delta_{tk}] = \rho_t \neq 0\). Define the sample statistics as in (3.1.3.1). Then it is straightforward to show that \(E[\hat{W}_t] = \beta_0 + d_t\), \(E[\hat{s}^2_{\hat{W}_t}] = \sigma^2_{\hat{W}_t} + \sigma^2_{\hat{\delta}_t}\), and \(E[s_{YW\hat{t}}] = \rho_t \sigma_{\epsilon_t} \sigma_{\delta_t}\). Hence the point estimators defined as (3.141) and (3.142) are still unbiased. However \(V[\hat{d}_t]\) changes due to the covariance term

\[
V[\hat{d}_t] = V[\hat{W}_t] + V[\hat{\beta}_0] - 2\text{Cov}[\hat{W}_t, \hat{\beta}_0] \quad (3.148)
\]

where \(V[\hat{d}_t]\) can be estimated by

\[
\hat{\sigma}_{dt}^2 = \frac{s^2_{\hat{W}_t}}{K_t} + \hat{\sigma}_{\hat{\beta}_0}^2 - 2\frac{s_{Y\hat{W}_t}}{K_t} \quad (3.149)
\]

Assuming normality of the error terms the distribution of \(\hat{d}_t\) is

\[
\frac{\hat{d}_t - d_t}{\hat{\sigma}_{dt}} \sim T(\nu_t),
\] (3.150)

and an approximate CI for \(d_t\) is

\[
\hat{d}_t \pm t(\alpha/2;\nu_t)\hat{\sigma}_{dt}.
\] (3.151)

In the case of the confidence interval containing zero, the interpretation is that the true value of the compound in the study sample is below the detection limit of the analytical method.
The point estimators, the distribution and approximate confidence interval of $\sigma_{Vt}^2$ remain as stated by (3.34) and (3.35).

If the correlation coefficient changes with each point in time, the results shown in Section 3.1.3.1 are applicable. Consider the case where the correlation is fixed and shared over time as a characteristic of the compound (or a characteristic of the relationship between the compound and the measurement procedure) while the independence between the errors at different points in time is preserved.

Assuming constant correlation over time we can estimate it as

$$\hat{\rho} = \frac{1}{T} \sum_t \frac{s^2_{Y_t W_t}}{s_{Yt} s_{Wt}}$$

with $\sum_t (K_t - 1)$ degrees of freedom.

The bootstrap technique is a preferable way to estimate CI and to test for non-zero correlation.

### 3.4.3.2 Examples

Table 3.15 shows results applying the methods to the data in Table 3.13. The columns shows the interval estimates for the observed values and the point and interval estimates for the concentration of each compound in the study sample, assuming the errors are independent.

It can be observed that the variances of the study values and their degrees of freedom after correcting for bias are increased in such a way that some of the confidence intervals are narrower than those of the observed data (4 out of 10).

Using the estimated constant non-zero correlation coefficient over time ($\rho = -0.385$) for illustration purpose only the new estimates are shown in Table 3.16. The estimated confidence intervals are slightly wider than those obtained with separate estimates of the correlation coefficient for each year.
Since the test for constant within-method variance among compounds was rejected we still assume it is constant compound-wise and the estimates of the within-material variance for the study material remain the same as showed in Table 3.5.

<table>
<thead>
<tr>
<th></th>
<th>(lcl_{\bar{W}_t})</th>
<th>(ucl_{\bar{W}_t})</th>
<th>(\hat{d}_t)</th>
<th>(\hat{\sigma}_t^2)</th>
<th>(\nu_{\sigma dt})</th>
<th>(lcl_{\bar{d}t})</th>
<th>(ucl_{\bar{d}t})</th>
</tr>
</thead>
<tbody>
<tr>
<td>1994</td>
<td>525.1</td>
<td>578.2</td>
<td>544.0</td>
<td>1214.0</td>
<td>35</td>
<td>473.3</td>
<td>614.7</td>
</tr>
<tr>
<td>1995</td>
<td>2609.0</td>
<td>2890.0</td>
<td>2758.0</td>
<td>2344.0</td>
<td>9</td>
<td>2648.0</td>
<td>2867.0</td>
</tr>
<tr>
<td>1996</td>
<td>759.5</td>
<td>899.1</td>
<td>823.7</td>
<td>1618.0</td>
<td>28</td>
<td>741.3</td>
<td>906.1</td>
</tr>
<tr>
<td>1997</td>
<td>686.6</td>
<td>782.1</td>
<td>733.0</td>
<td>1293.0</td>
<td>34</td>
<td>659.9</td>
<td>806.1</td>
</tr>
<tr>
<td>1998</td>
<td>870.9</td>
<td>935.8</td>
<td>944.7</td>
<td>1366.0</td>
<td>35</td>
<td>869.7</td>
<td>1020.0</td>
</tr>
<tr>
<td>1999</td>
<td>175.3</td>
<td>196.0</td>
<td>192.0</td>
<td>1210.0</td>
<td>35</td>
<td>121.4</td>
<td>262.6</td>
</tr>
<tr>
<td>2000</td>
<td>886.9</td>
<td>972.4</td>
<td>923.7</td>
<td>2799.0</td>
<td>30</td>
<td>815.6</td>
<td>1032.0</td>
</tr>
<tr>
<td>2002</td>
<td>635.0</td>
<td>670.3</td>
<td>698.7</td>
<td>3410.0</td>
<td>21</td>
<td>577.1</td>
<td>820.2</td>
</tr>
<tr>
<td>2003</td>
<td>475.6</td>
<td>500.4</td>
<td>467.7</td>
<td>545.6</td>
<td>38</td>
<td>420.4</td>
<td>514.9</td>
</tr>
<tr>
<td>2005</td>
<td>303.8</td>
<td>328.8</td>
<td>311.7</td>
<td>545.7</td>
<td>38</td>
<td>264.4</td>
<td>359.0</td>
</tr>
</tbody>
</table>

**Table 3.15.** Example of adjusted study values when errors are independent and constant bias at each point in time.

For each compound, \(lcl_{\bar{W}_t}\) = the lower confidence limit for the mean study value, \(ucl_{\bar{W}_t}\) = the upper confidence limit for the mean study value, \(\hat{d}_t\) = the study data corrected for bias using control data, under independent errors: \(\hat{\sigma}_t^2\) = the variance of the study data corrected for bias, \(\nu_{\sigma dt}\) = the degrees of freedom associated to the variance of the study data corrected for bias, \(lcl_{\bar{d}t}\) = the lower confidence limit for the mean of the study data corrected for bias, \(ucl_{\bar{d}t}\) = the upper confidence limit for the mean of the study data corrected for bias.
Table 3.16. Example of study values corrected for bias with correlated errors, assuming constant correlation over time and constant bias at each point in time.

For each compound, \( \hat{d}_t \) = the study data corrected for bias using control data, \( \hat{\sigma}^2_{\hat{d}_t} \) = the variance of the study data corrected for bias, \( \nu_{\hat{\sigma}_{\hat{d}_t}} \) = the degrees of freedom associated to the variance of the study data corrected for bias, \( lcl_{\hat{d}_t} \) = the lower confidence limit for the mean of the study data corrected for bias and \( ucl_{\hat{d}_t} \) = the upper confidence limit for the mean of the study data corrected for bias, under correlated errors.

<table>
<thead>
<tr>
<th>t</th>
<th>( \hat{d}_t )</th>
<th>( \hat{\sigma}^2_{\hat{d}_t} )</th>
<th>( \nu_{\hat{\sigma}_{\hat{d}_t}} )</th>
<th>lcl_{\hat{d}_t}</th>
<th>ucl_{\hat{d}_t}</th>
</tr>
</thead>
<tbody>
<tr>
<td>1994</td>
<td>544.0</td>
<td>1215.32</td>
<td>29</td>
<td>473.0</td>
<td>615.0</td>
</tr>
<tr>
<td>1995</td>
<td>2757.7</td>
<td>2366.08</td>
<td>1</td>
<td>2340.0</td>
<td>3170.0</td>
</tr>
<tr>
<td>1996</td>
<td>823.7</td>
<td>1631.45</td>
<td>2</td>
<td>655.0</td>
<td>992.0</td>
</tr>
<tr>
<td>1997</td>
<td>733.0</td>
<td>1294.08</td>
<td>27</td>
<td>659.0</td>
<td>807.0</td>
</tr>
<tr>
<td>1998</td>
<td>944.7</td>
<td>1371.83</td>
<td>7</td>
<td>858.0</td>
<td>1030.0</td>
</tr>
<tr>
<td>1999</td>
<td>192.0</td>
<td>1211.46</td>
<td>30</td>
<td>121.0</td>
<td>263.0</td>
</tr>
<tr>
<td>2000</td>
<td>923.7</td>
<td>2801.99</td>
<td>16</td>
<td>811.0</td>
<td>1040.0</td>
</tr>
<tr>
<td>2002</td>
<td>698.7</td>
<td>3413.13</td>
<td>9</td>
<td>566.0</td>
<td>832.0</td>
</tr>
<tr>
<td>2003</td>
<td>467.7</td>
<td>547.02</td>
<td>30</td>
<td>420.0</td>
<td>515.0</td>
</tr>
<tr>
<td>2005</td>
<td>311.7</td>
<td>547.14</td>
<td>30</td>
<td>264.0</td>
<td>359.0</td>
</tr>
</tbody>
</table>
CHAPTER 4

THE PROPORTIONAL AND LINEAR BIAS MODELS

The first section of this chapter explores the linear bias model for one laboratory, one compound, and one point in time, using only the control data. Under these constraints the linear bias model has infinite solutions thus it is restricted to become a proportional bias model with a single solution. Section 4.2 expands the model to multiple compounds at one point in time. This allows to explore the case when the linear bias model is shared across compounds.

The model for the observed concentration of the particular compound in the control material measured by one laboratory at one point in time is

\[ Y_k = \beta_0 + \beta_1(x + U_k) + \epsilon_k \]  \hspace{1cm} (4.1)

where \( k = 1, \ldots, K \), \( E[U_k] = 0 \), \( V[U_k] = \sigma_U^2 \), \( E[\epsilon_k] = 0 \), \( V[\epsilon_k] = \sigma_\epsilon^2 \), and \( (\beta_0, \beta_1, x, \sigma_U^2 \) and \( \sigma_\epsilon^2 \) \) are unknown parameters. We assume that \( U_k \) and \( \epsilon_k \) are independent for a given \( k \) and are independent among \( k \) and that \( \hat{x}, \hat{\sigma}_U^2 \) and \( \hat{\sigma}_\epsilon^2 \) are known unbiased estimates for \( x, \sigma_U^2 \) and \( V[\hat{x}] \). Recall that \( X_k = x + U_k \) represents the true allocated value of the \( k^{th} \) sample.

The model for the observed concentration of the particular compound in the study material is

\[ W_m = \beta_0 + \beta_1(d + V_m) + \delta_m \]  \hspace{1cm} (4.2)

where \( m = 1, \ldots, M \), \( E[\delta_m] = 0 \), \( V[\delta_m] = \sigma_\epsilon^2 \), \( E[V_m] = 0 \) and \( V[V_m] = \sigma_V^2 \). The parameters \( \beta_0, \beta_1 \) and \( \sigma_\epsilon^2 \) are shared by both models, \( V_m \) and \( \delta_m \) are independent for a given \( m \) and are independent among \( m \). In addition, assuming that \( \delta_m, V_m \) are independent from \( \epsilon_k, U_k \), then
\[ E[Y_k] = \beta_0 + \beta_1 x, \]
\[ V[Y_k] = \beta_1^2 \sigma_U^2 + \sigma_e^2, \]
\[ E[W_m] = \beta_0 + \beta_1 d \text{ and} \]
\[ V[W_m] = \beta_1^2 \sigma_V^2 + \sigma_e^2. \]  

(4.3)

These models will be expanded to allow changes over compounds, laboratories and time. The notation will differ only by adding adequate indexes.

4.1 One laboratory, one compound at one point in time

4.1.1 The Basic Model

Considering one compound at one point in time, we have two equations with three unknowns from the control data only. In order to obtain estimates some constraint is required. Only the proportional bias model will be addressed, assuming \( \beta_0 = 0 \) leads us to

\[ E[Y_k] = \beta_1 x, \]
\[ V[Y_k] = \beta_1^2 \sigma_U^2 + \sigma_e^2, \]
\[ E[W_m] = \beta_1 d \text{ and} \]
\[ V[W_m] = \beta_1^2 \sigma_V^2 + \sigma_e^2. \]  

(4.4)

Hence \( \beta_1 = \frac{E[Y]}{E[X]} \) and \( d = \frac{E[W]}{\beta_1} \). Both can be treated as ratios.

The assumption \( E[Y|X] = \beta_1 X \) may appear somewhat questionable. For parameter estimation purposes this model is related to the results found in Section 3.1 since \( E[Y_k]/x = 1 + \beta_0/x \) where \( \beta_0 = E[Y_k] - x \) and so \( \beta_0 = (\beta_1 - 1)x \). However, the consequences are different when exporting the model to the study data as will be shown.
4.1.2 Working with the Control Data Only

The main goal is to obtain estimates of each individual LMEM, hence there is no interest on pooling information across different participants. However there is interest in testing whether a participant’s method is biased.

When working with the control data only, the performance of the participants can be assessed directly provided that an estimate of \( \sigma^2 \) and of \( \sigma^2_U \) are given.

4.1.2.1 Parameter Estimation

In practice \( E[Y_k] \) and \( V[Y_k] \) are unknown and some estimates must be used instead. Define the sample mean \( \bar{Y} = \frac{1}{K} \sum_{k=1}^{K} Y_k \) which is unbiased for \( E[Y_k] \) and the sample variance \( s^2_Y = \frac{1}{K-1} \sum_{k=1}^{K} (Y_k - \bar{Y})^2 \) which is unbiased for \( V[Y_k] \). Let

\[
\hat{\beta}_1 = \frac{\bar{Y}}{\hat{x}},
\]

(4.5)

\[
\hat{\sigma}^2_\epsilon = s^2_Y - \hat{\beta}_1^2 \hat{\sigma}^2_U.
\]

(4.6)

Although, \( \hat{\beta}_1 \) is the maximum likelihood estimator of \( \beta_1 \), its statistical properties are complicated, depending on the distributional properties of \((\hat{x}, \bar{Y})\). It is well known that the ratio of normal random variables (which is a usual assumption with calibration models) is closely related to the general Cauchy distribution which has all moments undefined [38], [24]. We can however determine asymptotic properties of \( \hat{\beta}_1 \) and obtain confidence intervals for \( \beta_1 \).

We can apply the delta method (theorem 5.5.24 of [11]) in order to obtain approximations for the expected value and variance of \( \hat{\beta}_1 \):

\[
E[\hat{\beta}_1] \approx \beta_1 \left( 1 + \frac{\sigma^2_\epsilon}{\hat{x}^2} \right),
\]

(4.7)

\[
V[\hat{\beta}_1] \approx \frac{\sigma^2_Y}{\hat{x}^2} + \frac{\beta_1^2 \sigma^2_\epsilon}{\hat{x}^2}.
\]

(4.8)

(4.7) provides a simple expression for estimating the bias of \( \hat{\beta}_1 \) analytically,
\[
\text{bias}(\hat{\beta}_1) \approx \hat{\beta}_1 \hat{\sigma}_x^2. \tag{4.9}
\]

In the case of having a homogeneous material (i.e., \( \sigma_r^2 = 0 \)) then \( \hat{\sigma}^2 = s_Y^2 \) is unbiased for \( \sigma_r^2 \).

There are several well known alternatives to obtain interval estimators. The model can be stated as

\[
\begin{bmatrix}
\hat{x} \\
\hat{Y}
\end{bmatrix}
\sim (\mu, \Sigma),
\mu =
\begin{bmatrix}
x \\
\beta_1 x
\end{bmatrix},
\Sigma =
\begin{bmatrix}
\sigma_x^2 & 0 \\
0 & \sigma_Y^2
\end{bmatrix}.
\tag{4.10}
\]

One option to obtain a confidence interval for \( \beta_1 \) is by using the delta method and assuming \( \hat{\beta}_1 \) is approximately normal distributed. Assuming the bias in \( \hat{\beta}_1 \) is negligible (4.8) leads to an approximate 100(1 – \( \alpha \))%CI for \( \beta_1 \) of

\[
\hat{\beta}_1 \pm t_{(1-\alpha/2, \nu)} \sqrt{\frac{\hat{\sigma}_Y^2}{\hat{x}^2} + \frac{\hat{\beta}_1^2 \hat{\sigma}_x^2}{\hat{x}^2}}.
\tag{4.11}
\]

where \( \nu \) is a Welch-Satterthwaite approximated degrees of freedom

\[
\nu = \frac{(\hat{\sigma}_Y^2 + \hat{\beta}_1^2 \hat{\sigma}_x^2)^2}{\frac{\hat{\sigma}_Y^2}{K-1} + \hat{\beta}_1^2 \hat{\sigma}_x^2},
\tag{4.12}
\]

or \( \nu = \infty \) in which case \( t_{(1-\alpha/2, \nu)} = z_{(1-\alpha/2)} \).

Fieller’s method

Another option is the Fieller’s pivotal approach [7, ]. Applying the Fieller’s pivotal approach to (4.10) we obtain the approximate distribution

\[
\frac{(\hat{Y} - \beta_1 \hat{x})}{\sqrt{\hat{\sigma}_Y^2 + \hat{\beta}_1^2 \hat{\sigma}_x^2}} \sim T_\nu,
\tag{4.13}
\]

where \( \hat{\sigma}_x^2 \) is unbiased estimator for \( \sigma_x^2 \). Just assuming \( \nu = \infty \) corresponds to \( T_\nu \) being the standard normal. More revealingly for a specific significance level \( \alpha \):
\( P (q(\beta_1) \leq 0) = 1 - \alpha, \quad (4.14) \)

where \( q(\beta_1) = f_2\beta_1^2 - 2f_1\beta_1 + f_0 \) is a quadratic function of \( \beta_1 \) with \( f_0 = (\bar{Y}^2 - t_{(1-\alpha/2,\nu)}^2 \hat{\sigma}_Y^2), f_1 = \hat{x}\bar{Y} \) and \( f_2 = (\hat{x}^2 - t_{(1-\alpha/2,\nu)}^2 \hat{\sigma}_x^2) \).

Hence an approximate confidence set for \( \beta_1 \) is \( \{ \beta_1 : q(\beta_1) \leq 0 \} \). Defining \( D = f_1^2 - f_0f_2, r_1 = (f_1 - D^{1/2})/f_2, \) and \( r_2 = (f_1 + D^{1/2})/f_2, \) the approximate confidence set can be expressed as:

\[
\begin{cases}
[r_1, r_2] & \text{if } f_2 > 0, \\
(-\infty, r_2] \cup [r_1, \infty) & \text{if } f_2 \leq 0 \text{ and } D > 0, \\
(-\infty, \infty) & \text{if } f_2 \leq 0 \text{ and } D \leq 0.
\end{cases}
\quad (4.15)
\]

The approximation is due to the estimated degrees of freedom and the fact that the joint normality in (4.10) may only be estimated.

Reordering the terms in (4.13) it can be shown that the denominator should be proportional to the square root of an estimate of \( V[\hat{\beta}_1] \) as:

\[
\frac{(\bar{Y} - \beta_1 \hat{x})}{\sqrt{\hat{\sigma}_Y^2 + \beta_1^2 \hat{\sigma}_x^2}} = \frac{\bar{Y} - \beta_1}{\sqrt{\hat{\sigma}_Y^2 + \beta_1^2 \hat{\sigma}_x^2}} = \frac{\hat{\beta}_1 - \beta_1}{\sqrt{\hat{\sigma}_Y^2 + \beta_1^2 \hat{\sigma}_x^2}}
\quad (4.16)
\]

If we replace \( \beta_1 \) in the denominator by \( \hat{\beta}_1 \) this leads to an approximate CI equivalent to the interval estimator obtained by the delta method.

Another alternative is the bootstrap technique [20]. The bootstrap inferences can be obtained by using the so called parametric bootstrap. The nonparametric bootstrap strategy based on resampling the replicates or residuals with replacement is not useful for this case since only few replicates are available. The parameters of the population \( (\mu_x, \mu_Y, \sigma_x^2, \sigma_Y^2) \) are estimated from the sampled data \( (\hat{x}, \bar{Y}, \hat{\sigma}_x^2, s_Y^2) \). Then new re-sampled data \( (\hat{x}_b, \bar{Y}_b), b = 1, \ldots, B \) (large) are obtained from the parameterized distribution, with
\[ \hat{x}_b \sim N(\hat{x}, \hat{\sigma}_x^2), \]
\[ \bar{Y}_b \sim N(\bar{Y}, s_Y^2/K), \]

(4.17)

and \( \hat{x}_b \) and \( \bar{Y}_b \) are independent. In order to show that (4.17) mimics the original model recall that \( \bar{Y}_b = \hat{\beta}_1(\hat{x} + \bar{U}_b) + \bar{\epsilon}_b = \bar{Y} + Z_b\sqrt{\hat{\beta}_1^2 \hat{\sigma}_U^2 + \hat{\sigma}_b^2} = \bar{Y} + Z_b\sqrt{\frac{\hat{\sigma}_b^2}{K}}, \) with \( Z_b \) distributed standard normal.

Finally a set of new estimates of \( \beta_1 \) are obtained based on the re-sampled data \( B = \{ \beta_{1b} = \bar{Y}_b/\hat{x}_b, b = 1 \text{ to } B \} \) and the percentile confidence interval is obtained by computing \((Q_{(\alpha/2)}(B), Q_{(1-\alpha/2)}(B))\) where \( Q_{(\alpha)}(B) \) is the \( \alpha^{th} \) quantile of \( B \).

The bootstrap estimate of bias is

\[ \text{bias}(\hat{\beta}_1)^* = \hat{\beta}_1^* - \hat{\beta}_1, \]

(4.18)

where \( \hat{\beta}_1^* = \frac{1}{B} \sum_b \hat{\beta}_{1b}. \)

**Inferences for \( \sigma_t^2 \)**

Under normality,

\[ \frac{(K - 1)s_Y^2}{\beta_1^2 \sigma_U^2 + \sigma_t^2} \sim \chi^2_{K-1} \]

(4.19)

and an approximate CI for \( \sigma_t^2 \) with confidence level \( 1 - \alpha \), treating \( \hat{\beta}_1 \) and \( \hat{\sigma}_U^2 \) as \( \beta_1 \) and \( \sigma_U^2 \), respectively is

\[ \left( \frac{(K - 1)s_Y^2}{\chi^2_{(1 - \alpha/2, K-1)}} - \hat{\beta}_1^2 \hat{\sigma}_U^2, \frac{(K - 1)s_Y^2}{\chi^2_{(\alpha/2, K-1)}} - \hat{\beta}_1^2 \hat{\sigma}_U^2 \right). \]

(4.20)

There is potential for obtaining estimates outside of the parameter space for \( \sigma_t^2 \). The interval would be truncated so that it contains only non-negative values and if the upper bound of (4.20) is less that zero then the interpretation would be that \( \sigma_t^2 \) is zero.

A better option is to use additional information about \( U_k \) if it is available, suppose \( \nu_U \hat{\sigma}_U^2/\hat{\sigma}_U^2 \sim \chi^2(\nu_U) \) then a 100(1 - \( \alpha \)% CI for \( \sigma_U^2 \), say \([lcl(\sigma_U^2), ucl(\sigma_U^2)]\) can be
constructed. Again, treating $\hat{\beta}_1$ as $\beta_1$ and using Bonferroni’s method a CI for $\sigma^2_\epsilon$ with confidence level $1 - 2\alpha$ is given by

$$\left( \frac{(K-1)s_Y^2}{\chi^2_{(1-\alpha/2,K-1)}} - \hat{\beta}_1^2\text{ucl}(\sigma^2_\epsilon), \frac{(K-1)s_Y^2}{\chi^2_{(\alpha/2,K-1)}} - \hat{\beta}_1^2\text{lcl}(\sigma^2_\epsilon) \right). \quad (4.21)$$

Note that if $\sigma^2_\epsilon = 0$ then (4.20) and (3.15) lead to the same estimates. Both (4.20) and (4.21) require an estimate of $\beta^2_1$ so an additional adjustment is taking into account the uncertainty attached to $\hat{\beta}_1$. We can use $(\hat{\beta}_1^2 - \hat{\sigma}^2_\epsilon)$ instead of $\hat{\beta}_1^2$.

Rather than continue to extend these equations to account for uncertainty in $\hat{\beta}_1$ a better strategy is to use the bootstrap adding in generating $\hat{\sigma}^2_{Ub}$ with

$$\hat{\sigma}^2_{Ub} \sim \frac{\sigma^2_U}{\nu_U} \chi^2_{(\nu_U)},$$

$$s^2_{Yb} \sim \frac{s^2_2}{K-1} \chi^2_{(K-1)},$$

$$\hat{\sigma}^2_{eb} = s^2_{Yb} - \hat{\beta}_1 b \hat{\sigma}^2_{Ub}. \quad (4.22)$$

4.1.2.2 Inferences about the Parameters

**Testing for Zero Bias**

The hypothesis of unbiasedness ($H_0 : \beta_1 = 1$) can be tested approximately using (4.11). $H_0$ can be rejected, and $\beta_1 \neq 1$ concluded, with an approximate significance level of $\alpha$, if

$$|t_{\text{score}}| > t_{(1-\alpha/2,\nu)}$$

where

$$t_{\text{score}} = \frac{\hat{\beta}_1 - 1}{\sqrt{\frac{s^2_Y}{2} + \frac{\hat{\beta}_1^2}{2}}} \quad (4.23)$$

Another way to test for unbiasedness is by using the Fieller’s CI as defined in (4.15). We can reject $H_0$ and conclude $\beta_1 \neq 1$ if $1 \notin (4.15)$.

Note also that under the unbiasedness hypothesis (4.11) and (3.17) are equivalent and both lead to the same conclusion with the same significance. However (4.15) does not reduce to (4.11) unless $\sigma^2_\epsilon = 0$. 

103
While enabling testing if bias is zero, the model does not allow testing for bias constancy. Two or more points are required across compounds or time in order to test for constancy of the bias. This is addressed in the Section 4.2.

4.1.2.3 Performance Evaluation

Using (4.23) we can evaluate the performance of the participant. Under the hypothesis of unbiasedness the statistic (4.23) and (3.17) are the same and the $t_{score}$ can be used to evaluate the performance of the participant.

The evaluation of the participant’s method variability can be done by comparing it against the variability of the methods used during the certification as stated in Section 3.1.2.3.

4.1.2.4 Examples

Here we use the data from Table 3.1 to illustrate the methods just described. Table 3.1 lists the conventional values for the control material SRM 1941b and the observed values as reported by one of the participants about the PAH family of compounds in Sediments from the IEPOCME 2003 inter-comparison exercise.

Table 4.1 shows estimates and inferences for $\beta_1$ using $\sigma_U^2 = 0$, for each compound:

- $\hat{\beta}_1$ = the estimated slope using (4.5),
- $\hat{\sigma}_{\beta_1} = \sqrt{s^2 + \hat{\beta}_1^2 \hat{\sigma}_x^2}$ = standard error of the estimated slope using (4.8) with estimated quantities,
- $\nu_{\beta_1}$ = the degrees of freedom of the standard error using (4.12),
- $(lcl_B, ucl_B) = a 95\%$ confidence interval using the bootstrap technique,
- $\hat{bias}_B$ the estimated bias of the $\hat{\beta}_1$ using the bootstrap technique using (4.18),
- $(lcl_{DM}, ucl_{DM}) = a 95\%$ confidence interval using the delta method with Welch-Satterthwaite approximated degrees of freedom using (4.11),
• \( \text{bias}_{DM} \) = the estimated bias of the \( \hat{\beta}_1 \) using the delta method using (4.9) and

• \((lcl_F, ucl_F)\) = a 95% confidence interval using the Fieller approach using (4.15).

Applying (4.6) and (4.20) using \( \sigma^2_U = 0 \) we obtain the estimates for \( \sigma^2_\epsilon \), its degrees of freedom and confidence intervals. These are the same results shown in Table 3.3.

All the confidence intervals contain 1 hence we fail to reject \( H_0 \) and we conclude there is no evidence that the method is biased, although some of the lower bounds are down near 0.7 and some upper bounds only slightly over 1.

In general the bootstrap CIs and the Fieller’s CIs agree among them while the delta method CIs tend to be shifted towards zero relative to the bootstrap and Fieller’s CIs. This is consistent with the estimated bias, whereas the bias estimated with the delta method tends to be larger than the bias estimated with the bootstrap technique.

### 4.1.3 Exporting the Model to the Study Data

The models found while working with the control data can be used to estimate the unknown parameters of the study data.

#### 4.1.3.1 Parameter Estimation

**Under Independent Errors**

First assume that the error terms are independent for the method measuring the control sample and the study sample; i.e., \( \text{Corr}[\epsilon_k, \delta_k] = \rho = 0 \). Also assume \( \sigma^2_\epsilon = \sigma^2_\delta \); i.e., the variance of the method is constant across the control and study samples. This can be due to the constancy of the method’s variance over different levels or to a method’s variance dependent on the concentration level but analyzing compounds with similar levels in both control and study samples.

Under independent error terms, the number of replicates of the control material and the number of replicates of the study material are not required to be equal.
Table 4.1. Example of point and interval estimates using model with proportional bias.

For each compound, $\hat{\beta}_1$ = the estimated slope, $\hat{\sigma}_{\beta_1}$ = standard error of the estimated slope, $\nu_{\beta_1}$ = the degrees of freedom of the standard error, $(lcl_B, ucl_B)$ = a 95% confidence interval using the bootstrap technique, $\hat{\text{bias}}_B$ the estimated bias of the $\hat{\beta}_1$ using the bootstrap technique (4.18), $(lcl_{DM}, ucl_{DM})$ = a 95% confidence interval using the delta method with Welch-Satterthwaite approximated degrees of freedom (4.11), $\hat{\text{bias}}_{DM}$ = the estimated bias of the $\hat{\beta}_1$ using the delta method (4.9), and $(lcl_F, ucl_F)$ = a 95% confidence interval using the Fieller approach (4.15).
The primary goal is to estimate \( d \) (the true value associated with the total collection of study material) and \( V[W_m] = \sigma^2_V \) (the among sample variance). Define the sample mean \( \bar{W} = \frac{1}{M} \sum_{m=1}^{M} W_m \) which is unbiased for \( E[W_m] = \beta_1 d \) and the sample variance \( s^2_W = \frac{1}{M-1} \sum_{m=1}^{M} (W_m - \bar{W})^2 \) which is unbiased for \( V[W_m] \). This leads to

\[
\hat{d} = \frac{\bar{W}}{\beta_1}, \quad (4.24)
\]

and

\[
\hat{\sigma}^2_V = \frac{s^2_W - \hat{\sigma}^2_\epsilon}{\beta_1^2}. \quad (4.25)
\]

Applying Jensen’s inequality,

\[
E[\hat{d}] = E[\bar{W} \hat{x}/\bar{Y}] \geq \frac{E[\bar{W}]E[\hat{x}]}{E[\bar{Y}]} = \frac{\beta_1 dx}{\beta_1 x} = d \quad (4.26)
\]

and

\[
E[\hat{\sigma}^2_V] = E \left[ \frac{s^2_W - \hat{\sigma}^2_\epsilon}{\beta_1^2} \right] \geq \frac{\sigma^2_W - \sigma^2_\epsilon}{E[\beta_1^2]} = \frac{\beta_1^2}{\beta_1^2 + \sigma^2_\epsilon}. \quad (4.27)
\]

An approximation of the variance of \( \hat{d} \) can be obtained from (4.24), by using the delta method

\[
V[\hat{d}] \approx \frac{V[\bar{W}] + d^2 V[\hat{\beta}_1]}{\beta_1^2}. \quad (4.28)
\]

Plugging estimates in and defining \( s^2_{\bar{W}} = s^2_W/M \), an estimate of (4.28) is

\[
\hat{V}[\hat{d}] = \frac{s^2_{\bar{W}} + \hat{d}^2 \hat{\sigma}^2_{\bar{Y}}}{\hat{\beta}_1^2}. \quad (4.29)
\]

From (4.26) is it clear that \( \hat{d} \) is biased, this is \( E[\hat{d}] \neq d \). The bias of \( \hat{d} \) can also be approximated by the delta method. Defining \( s^2_Y = s^2_{\bar{Y}}/K \), an estimate of the bias is

\[
\widehat{\text{bias}}(\hat{d}) \approx d \frac{s^2_Y}{\bar{Y}^2}. \quad (4.30)
\]

Some additional distributional assumptions on the error terms are required in order to get confidence intervals for \( d \) and \( \sigma^2_V \). Assuming
\[
\frac{\bar{W} - \hat{\beta}_1 d}{\sqrt{s_w^2 + d^2 \hat{\sigma}_{\beta_1}^2}} \sim T_{\nu_d},
\]
(4.31)

approximately, where

\[
\nu_d \approx \left(\frac{s_w^2 + \hat{d}^2 \hat{\sigma}_{\beta_1}^2}{\frac{s_w^4}{\nu_w} + \frac{d^4 \hat{\sigma}_{\beta_1}^4}{\nu_{\beta_1}}}\right)^2,
\]
(4.32)

an approximate 100(1 - \alpha)% delta method confidence interval for \(d\) then becomes

\[
\left(\hat{d} + t_{(\alpha/2;\nu_d)}\sqrt{\hat{V}[\hat{d}]}, \hat{d} + t_{(1-\alpha/2;\nu_d)}\sqrt{\hat{V}[\hat{d}]}\right).
\]
(4.33)

Applying Fieller’s method to (4.31)

\[
P(q(d) \leq 0) = 1 - \alpha,
q(d) = f_2 d^2 - 2 f_1 d + f_0,
\]
\[
f_0 = \bar{W}^2 - t_{(1-\alpha/2;\nu_d)}^2 \hat{\sigma}_{\beta_1}^2,
\]
\[
f_1 = \hat{\beta}_1 \bar{W},
\]
\[
f_2 = \hat{\beta}_1^2 - t_{(1-\alpha/2;\nu_d)}^2 \hat{\sigma}_{\beta_1}^2.
\]
(4.34)

Hence an approximate confidence interval for \(\hat{d}\) is the set \(\{d : q(d) \leq 0\}\). Defining \(D = f_1^2 - f_0 f_2\), \(r_1 = (f_1 + D^{1/2})/f_2\), and \(r_2 = (f_1 + D^{1/2})/f_2\), the approximate confidence set can be expressed as

\[
\begin{cases}
[r_1, r_2] & \text{if } f_2 > 0, \\
(-\infty, r_2] \cup [r_1, \infty) & \text{if } f_2 \leq 0 \text{ and } D > 0, \\
(-\infty, \infty) & \text{if } f_2 \leq 0 \text{ and } D \leq 0.
\end{cases}
\]
(4.35)

This confidence interval can contain zero or even contain only negative values. In such a case the interpretation is that the true value of the compound in the study
sample is below the detection limit of the analytical method; i.e., a censored data point.

Applying the parametric bootstrap method, for \( b = 1, \ldots, B \) (large) generate
\[
\hat{x}_b \sim N(\hat{x}, \hat{\sigma}_x^2),
\]
\[
\hat{Y}_b \sim N\left(\hat{\beta}_1 \hat{x}, \frac{\hat{\beta}_1^2 \hat{\sigma}_x^2 + \hat{\sigma}_\epsilon^2}{K}\right) \equiv N(\hat{Y}, s_{\hat{Y}}^2),
\]
\[
\hat{W}_b \sim N\left(\hat{\beta}_1 \hat{d}, \frac{\hat{\beta}_1^2 \hat{\sigma}_\epsilon^2 + \hat{\sigma}_\epsilon^2}{M}\right) \equiv N(\hat{W}, s_{\hat{W}}^2),
\]
then compute the sets
\[
\mathcal{B} = \left\{ \hat{\beta}_{1b} = \hat{Y}_b / \hat{x}_b, b = 1 \text{ to } B \right\}, \quad \mathcal{D} = \left\{ \hat{d}_b = \hat{W}_b / \hat{\beta}_{1b}, b = 1 \text{ to } B \right\}.
\]

A \((1 - \alpha)100\%\) percentile CI for \( \hat{d} \) is \((Q(\alpha/2)(\mathcal{D}), Q(1-\alpha/2)(\mathcal{D}))\), where \( Q(\alpha)(\mathcal{D}) \) is the \( \alpha^{th} \) quantile of \( \mathcal{D} \).

The bias of the adjusted study value \( \hat{d} \) can be estimated as
\[
\widehat{\text{bias}}(\hat{d})^* = \frac{1}{B} \sum_{b=1}^{B} \hat{d}_b - \hat{d} = \hat{d}^* - \hat{d}.
\]

Under normality
\[
\frac{(M - 1)s_{\hat{W}}^2}{\hat{\beta}_1^2 \hat{\sigma}_\epsilon^2 + \hat{\sigma}_\epsilon^2} \sim \chi^2_{(M-1)}.
\]

Treating \( \hat{\beta}_1 \) as \( \beta_1 \) and \( \hat{\sigma}_\epsilon^2 \) as \( \sigma_\epsilon^2 \) leads to a \((1 - \alpha)100\%\) CI for \( \sigma^2_V \) of
\[
\left( \frac{(M - 1)s_{\hat{W}}^2}{\hat{\beta}_1^2 \chi^2_{(1-\alpha/2;M-1)}} - \hat{\sigma}_\epsilon^2, \frac{(M - 1)s_{\hat{W}}^2}{\hat{\beta}_1^2 \chi^2_{(\alpha/2;M-1)}} - \hat{\sigma}_\epsilon^2 \right).
\]
The approximation comes from use of the estimates \( \hat{\beta}_1 \) and \( \hat{\sigma}_\epsilon^2 \) as \( \beta_1 \) and \( \sigma_\epsilon^2 \). There is potential for obtaining values in the interval that are outside of the parameter space for \( \sigma^2_V \), such as 0 and negative values. The interval should be truncated to include only non-negative values. As with the control data a better approach here is to bootstrap adding in
\[
s_{\hat{W}_b}^2 \sim \frac{s_{\hat{W}}^2}{M - 1} \chi^2_{(M-1)},
\]
\[
\hat{\sigma}^2_{V_b} = \frac{s_{\hat{W}_b}^2 - \hat{\sigma}^2_{\epsilon_b}}{\hat{\beta}^2_{1b}}.
\]
Under Correlated Errors

Now assume the data is paired and that the error terms are correlated for the method while measuring the control sample and the study sample; i.e., $\text{Corr}[\epsilon_k, \delta_k] = \rho \neq 0$. Define the sample statistics $\bar{Y}, s_Y^2, \bar{W}, s_W^2$ as above, define the sample covariance as $s_{Y,W} = \frac{1}{K-1} \sum_{k=1}^{K} (Y_k W_k - \bar{Y} \bar{W})$, and let $s_Y^2 = s_Y^2 / M$, $s_W^2 = s_W^2 / M$ and $s_{\bar{Y},\bar{W}} = s_{Y,W} / M$ be the sample variances and covariance, respectively, of the means.

It is straightforward to show that

$$E[\bar{W}] = \beta_1 \bar{d}, \ E[s_W^2] = \beta_1^2 \sigma^2 + \sigma^2$$ and $E[s_{Y,W}] = \rho \sigma \sigma = \rho \sigma^2$.

The last equality is due to the assumption of equal variances, $\sigma^2 = \sigma^2$. Hence the point estimators defined as (4.24) and (4.25) are still adequate.

However $V[\hat{d}]$ changes due to a covariance term. Using the delta method, the conditional variance identity and assuming $\text{bias}[\hat{\beta}_1]$ to be negligible this can be approximated by

$$V[\hat{d}] \approx \frac{1}{\hat{\beta}_1^2} \left( V[\bar{W}] + \hat{d}^2 V[\hat{\beta}_1] - 2\hat{d} \text{Cov}[\bar{W}, \hat{\beta}_1] \right),$$

where

$$\text{Cov}[\bar{W}, \hat{\beta}_1] = \text{Cov} \left[ \bar{W}, \frac{\bar{Y}}{\bar{x}} \right]$$

$$= E \left[ \text{Cov} \left[ \bar{W}, \frac{\bar{Y}}{\bar{x}} | \hat{x} \right] \right] + \text{Cov} \left[ E[\bar{W} | \hat{x}], E \left[ \frac{\bar{Y}}{\bar{x}} | \hat{x} \right] \right]$$

$$= E \left[ \frac{1}{\hat{x}} \text{Cov}[\bar{W}, \bar{Y}] \right] + E[\bar{W}] \text{Cov} \left[ 1, \frac{\bar{Y}}{\bar{x}} \right]$$

$$= 0$$

Now, $V[\hat{d}]$ can be estimated by

$$\hat{V}[\hat{d}] = \frac{1}{\hat{\beta}_1^2} \left( s_W^2 + \hat{d}^2 \sigma_\beta^2 - 2 \frac{\hat{d}}{\hat{x}} s_{Y,W} \right)$$

$$= \hat{d}^2 \left( \frac{\sigma^2}{\hat{x}^2} + \frac{s_W^2}{W^2} + \frac{s_Y^2}{Y^2} - 2 \frac{s_{Y,W}}{Y \bar{W}} \right).$$

110
Note that if $\hat{\sigma}^2/\hat{x}^2$ is much larger than $s_Y^2/Y^2$ and $s_W^2/W^2$ then any correlation has little effect in $\hat{V}[\hat{d}]$.

The bias of $\hat{d}$ under correlated errors can also be approximated by the delta method, with

$$\text{bias}(\hat{d}) \approx \hat{d} \left( \frac{s_Y^2}{Y^2} - \frac{s_Y s_W}{YW} \right).$$

Assuming normality of $V_k$ and $\epsilon_k$ and $K$ small the distribution and confidence interval of $d$ can be approximated as

$$\frac{\hat{d} - d}{\sqrt{\hat{V}[\hat{d}]}} \sim T(\nu_d),$$

where

$$\nu_d \approx \frac{2 (\hat{V}[\hat{d}])^2}{\hat{V} \left[ \hat{V}[\hat{d}] \right]}.$$  (4.45)

The procedure to obtain $\hat{V}[\hat{V}[\hat{d}]]$ is lengthy and tedious but it is based on well known theory [2], assuming normality of $U_k, \epsilon_k, V_k$ and $\delta_k$,

$$\hat{V} \left[ \hat{V}[\hat{d}] \right] = \frac{1}{\hat{\beta}_1^4} \left\{ \frac{2s_W^4}{(K-1)} + \frac{2d^4\hat{\beta}_1^4}{\nu_{\beta_1}} + \frac{4d^2 S^2 W S^2 \bar{Y}}{\bar{x}^2(K-1)} + \frac{8d^2 S^2 \bar{Y} S W Y}{\bar{x}^2(K-1)} - \frac{8d S^2 W S W Y}{\bar{x}(K-1)} \right\}.$$  (4.46)

An approximate $(1 - \alpha)100\%$ confidence interval for $d$ based on the delta method is

$$\left( \hat{d} + T(\alpha/2\nu_d) \sqrt{\hat{V}[\hat{d}]}, \hat{d} + T(1-\alpha/2\nu_d) \sqrt{\hat{V}[\hat{d}]}, \right).$$  (4.47)

Applying the Fieller’s method, we can use (4.35) but modifying (4.34) to include the correlation adjustments. This leads to
\[ P(q(d) \leq 0) = 1 - \alpha, \]
\[ q(d) = f_2 d^2 - 2f_1 d + f_0, \]
\[ f_0 = \hat{W}^2 - t_{(1-\alpha/2, \nu_d)}^2 \hat{\sigma}_W^2, \]
\[ f_1 = \hat{\beta}_1 \hat{W} - t_{(1-\alpha/2, \nu_d)}^2 \hat{\text{Cov}}[\hat{W}, \hat{\beta}_1], \]
\[ f_2 = \hat{\beta}_1^2 - t_{(1-\alpha/2, \nu_d)}^2 \hat{\sigma}_{\beta_1}^2, \]

where \( \hat{\text{Cov}}[\hat{W}, \hat{\beta}_1] = \frac{1}{\hat{x}} s_{\hat{W}, \hat{Y}} \) and the approximate degrees of freedom are obtained by using (4.45) and (4.46).

Applying the parametric bootstrap method, for \( b = 1, \ldots, B \) (large) generate
\[ \begin{bmatrix} \bar{Y}_b \\ \bar{W}_b \end{bmatrix} \sim N \left( \begin{bmatrix} \hat{\beta}_1 \hat{\beta}_d \\ \hat{\beta}_1 \hat{\beta}_d \end{bmatrix}, \begin{bmatrix} \frac{\hat{\beta}_1^2 \hat{\sigma}_W^2 + \hat{\sigma}_\epsilon^2}{K} & \hat{\sigma}_{\epsilon, \delta} \\ \hat{\sigma}_{\epsilon, \delta} & \frac{\hat{\beta}_1^2 \hat{\sigma}_W^2 + \hat{\sigma}_\epsilon^2}{K} \end{bmatrix} \right) \]
\[ \equiv N \left( \begin{bmatrix} \bar{Y} \\ \bar{W} \end{bmatrix}, \begin{bmatrix} s_{\bar{Y}}^2 & s_{\bar{Y}, \bar{W}} \\ s_{\bar{Y}, \bar{W}} & s_{\bar{W}}^2 \end{bmatrix} \right). \]

The bootstrap inference then proceeds as in the uncorrelated case. Also note that the estimator for the correlation coefficient \( \hat{\rho} \) does not alter the point estimators, so the bootstrap only needs to generate the means.

As with the point estimators, the distribution and approximate confidence interval of \( \sigma_V^2 \) remain as stated by (4.38) and (4.39).

If \( \sigma_{\delta k|D_k}^2 \neq \sigma_{\epsilon k|X_k}^2 \), as may be the case if the method’s variance depends on the concentration of the compound and the concentration of the compound in the control material and the study material are very different, then \( \sigma_{\delta D}^2 \) and \( \sigma_{\epsilon X}^2 \) should be modeled accordingly. In Chapter 5 we briefly review a variance model proportional to the square of the concentration.
4.1.3.2 Example

Table 4.2 contains the estimates of the adjusted study values and their variances for the data in Table 3.1, under the assumption of independent errors and equal within method variance for each compound across samples. Table 4.3 contains the estimates of the adjusted study values and their variances and confidence intervals under the assumption of correlated errors and equal within method variance for each compound across samples.

In Table 4.2 it can be observed that the corrected values \( \hat{d} \) change more or less evenly around the observed values \( \bar{W} \) and their magnitudes are very similar. Recalling the results from Section 3.1.3.2, the proportional bias model produces confidence interval estimates within reasonable values in contrast to the constant bias model that can produce confidence interval estimates outside the parameter space (negative concentration levels). In the best case, a censored value is produced, although the observed value has a high likelihood of being positively present.

However, the corrected variance estimate \( \hat{\sigma}_d^2 \) is larger than the observed variance \( s^2_{\bar{W}} \) in all the cases. This is due to the additional variance terms as expressed in (4.28) and (4.42). Only two corrected confidence intervals appear to be narrower than their corresponding observed confidence intervals (triphenylene and benzo[j]fluoranthene compounds, both are assigned the lowest corrected degrees of freedom). In general, we can expect to have wider confidence intervals for \( d \) when allowing a proportional bias model since there is need to account for the uncertainty in \( \hat{\beta}_1 \) compared to the confidence intervals obtained from the study data \( W_m \) and assuming unbiasedness. This is clear from (4.29) and the unbiasedness assumption that \( \hat{V} [\hat{d}] \approx s^2_{\bar{W}} + \hat{d}^2 \hat{\sigma}_d^2 > s^2_{\bar{W}} \).

Under the assumption of unbiasedness no adjustment is required and the naive estimators \( \hat{d}_{naive} = \bar{W} \) and \( \hat{\sigma}^2_{d, naive} = s^2_{\bar{W}} \) would lead to a CI for \( \hat{d}_{naive} \) of
\[(l_{clW}, ucl_W) = (\hat{d}_{\text{naive}} + t_{(\alpha/2;K-1)}\hat{\sigma}_{d,\text{naive}}, \hat{d}_{\text{naive}} + t_{(1-\alpha/2;K-1)}\hat{\sigma}_{d,\text{naive}}), \quad (4.50)\]

this is the CI obtained by just using the study data \(W_m\). The coverage of the naive confidence interval for \(\hat{d}\) is less than 47\%, this is \(\hat{Pr}(\hat{d} \in (l_{clW}, ucl_W)) < 0.47\). This is true under the assumption of independent or correlated error terms for the specific data shown. In contrast, the adjusted confidence interval for the study data contains most likely the naive estimator \(\bar{W}\), this is \(\hat{Pr}(\bar{W} \in (l_{cl\hat{d}}, ucl_{\hat{d}})) \approx 1\).

When comparing the delta method estimates against the Fieller’s method and bootstrap estimates, a remarkable distinction is the asymmetry of the Fieller’s method and the bootstrap intervals against the the delta method intervals, which are symmetric by construction. One consequence is that the delta method intervals often overlap with the bootstrap and Fieller’s intervals on the shortest tail side of the re-sampled distribution, suggesting that the delta method intervals are biased. A second distinction is that in general the delta method intervals are shorter than the bootstrap and Fieller’s intervals, and this is also related to the skewness of the re-sampled values. The combination of these two effects (biased and too optimistic intervals) leads to a lowered confidence coefficient when compared to the Fieller’s and the bootstrap confidence coefficient. The fact that Fieller’s method outperforms the delta method intervals for estimating ratios is well documented (under normality see Cox(1990) [15] and also for ratio of binary variables Sitter & Wu(1993) [49]).

Table 4.4 contains only the Fieller’s estimates from Tables 4.2 and 4.3 and the correlation coefficient from Table 3.1 for easy of comparison. It can be observed that in general the estimated confidence intervals under correlation of the error terms are narrower when the correlation is positive. CIs are larger when the correlation is negative in comparison with the CIS under independence of the error terms. This is true for the bootstrap, delta method and Fieller’s method estimates. One compound (triphenylene) shows a larger confidence interval under correlated errors while having
positive correlation. We can confirm that the estimates of \( d \) and the CI for \( d \) differ slightly when the fitting is done using the proportional bias model than when fitting the constant bias model seen in Section 3.1; this can be done by comparing Table 3.5 and Table 4.2 or Table 3.6 and Table 4.3, although the models are related as mentioned in Section 4.1.1. It may be surprising that the differences in the CIs are small considering some of the estimated correlation coefficients are relatively large. We can explain this behavior by analyzing the variance components in (4.42) and using the data in Table 3.1. In most cases \( \hat{\sigma}_x^2 / \hat{x}^2 \) is larger than both \( s_Y^2 / \bar{Y}^2 \) and \( s_W^2 / \bar{W}^2 \) so we can expect no significant changes in \( \hat{V}[\hat{d}] \) due to correlation.

Table 4.5 contains the estimated bias of the adjusted study value \( \hat{d} \) from the analytical approach and the bootstrap estimate for both under independent errors and correlated errors. In general, the estimated bias using the delta method tends to be smaller than the bootstrap technique, which is not surprising since the analytical estimate is a second order approximation and the discarded terms are positive. When considering the correlation the estimates tend to be more similar independently of the sign of the correlation coefficient.

Table 4.6 contains the estimates of the within material variance for the study data by using (4.25) and (4.39), while assuming the measurement error variance for each compound is the same across samples. The estimates of the within material variance are those on Table 3.5 scaled by the factor \( 1 / \hat{\beta}_1^2 \). The estimated probability of obtaining negative estimates is the same. This shows the difficulties of estimating \( \sigma_V^2 \).
For each compound, $W$ = the observed mean study value, $(lcl_W, ucl_W)$ = a 95% confidence interval for $W$. $\hat{d}$ = the corrected mean of the study value under proportional bias; $\hat{\sigma}_{dDM}$ = the estimated standard error of $\hat{d}$; $\nu_{\hat{d}}$ = the degrees of freedom of the variance of $\hat{d}$; $(lcl_{DM}, ucl_{DM})$ = 95% confidence interval for $\hat{d}$ based on the delta method; $(lcl_F, ucl_F)$ = 95% confidence interval for $\hat{d}$ using Fieller’s method; $\hat{\sigma}_{dB}$ = bootstrap estimate of standard error of $\hat{d}$; $(lcl_B, ucl_B)$ = 95% bootstrap confidence interval for $\hat{d}$.

### Table 4.2: Example of study values and variances corrected for proportional bias using control data when errors are independent.

<table>
<thead>
<tr>
<th>Compound</th>
<th>$W$</th>
<th>$lcl_W$</th>
<th>$ucl_W$</th>
<th>$\hat{d}$</th>
<th>$\nu_{\hat{d}}$</th>
<th>$lcl_{DM}$</th>
<th>$ucl_{DM}$</th>
<th>$lcl_F$</th>
<th>$ucl_F$</th>
<th>$\hat{\sigma}_{dB}$</th>
<th>$lcl_B$</th>
<th>$ucl_B$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2,3,5-trimethynaphthalene</td>
<td>25.7</td>
<td>24.32</td>
<td>27.08</td>
<td>28.87</td>
<td>30.07</td>
<td>32</td>
<td>22.62</td>
<td>35.12</td>
<td>3.053</td>
<td>22.97</td>
<td>35</td>
<td></td>
</tr>
<tr>
<td>acenaphthene</td>
<td>34.7</td>
<td>33.56</td>
<td>35.84</td>
<td>37.05</td>
<td>2.887</td>
<td>24</td>
<td>31.09</td>
<td>43</td>
<td>31.91</td>
<td>44.13</td>
<td>2.865</td>
<td>31.62</td>
</tr>
<tr>
<td>dibenz[a,h]anthracene</td>
<td>112</td>
<td>101.7</td>
<td>122.3</td>
<td>105.4</td>
<td>10.43</td>
<td>11</td>
<td>82.49</td>
<td>128.4</td>
<td>86.41</td>
<td>134.4</td>
<td>10.39</td>
<td>84.95</td>
</tr>
<tr>
<td>acenaphthylene</td>
<td>29.63</td>
<td>26.26</td>
<td>33.01</td>
<td>27.81</td>
<td>1.945</td>
<td>28</td>
<td>23.82</td>
<td>31.79</td>
<td>24.26</td>
<td>32.35</td>
<td>1.929</td>
<td>24.13</td>
</tr>
<tr>
<td>biphenyl</td>
<td>29.23</td>
<td>27.47</td>
<td>31</td>
<td>30.9</td>
<td>1.833</td>
<td>32</td>
<td>27.17</td>
<td>34.64</td>
<td>27.55</td>
<td>35.12</td>
<td>1.82</td>
<td>27.41</td>
</tr>
<tr>
<td>1-methylphenanthrene</td>
<td>69.4</td>
<td>67.85</td>
<td>70.95</td>
<td>70.75</td>
<td>2.956</td>
<td>28</td>
<td>64.69</td>
<td>76.81</td>
<td>65.16</td>
<td>77.37</td>
<td>2.946</td>
<td>64.93</td>
</tr>
<tr>
<td>2,6-dimethynaphthalene</td>
<td>35.2</td>
<td>34.06</td>
<td>36.34</td>
<td>36.09</td>
<td>1.255</td>
<td>22</td>
<td>33.49</td>
<td>38.69</td>
<td>33.65</td>
<td>38.88</td>
<td>1.241</td>
<td>33.69</td>
</tr>
<tr>
<td>fluorene</td>
<td>65.5</td>
<td>63.78</td>
<td>67.22</td>
<td>65.04</td>
<td>6.075</td>
<td>41</td>
<td>52.7</td>
<td>77.31</td>
<td>54.71</td>
<td>80.15</td>
<td>6.047</td>
<td>53.27</td>
</tr>
<tr>
<td>triphenylene</td>
<td>141</td>
<td>119.9</td>
<td>162.1</td>
<td>137.2</td>
<td>6.074</td>
<td>5</td>
<td>121.5</td>
<td>152.9</td>
<td>122.1</td>
<td>153.6</td>
<td>6.022</td>
<td>125.7</td>
</tr>
<tr>
<td>1-methylphenanthlene</td>
<td>48.97</td>
<td>46.71</td>
<td>51.22</td>
<td>51.54</td>
<td>3.359</td>
<td>22</td>
<td>44.58</td>
<td>58.5</td>
<td>45.38</td>
<td>59.55</td>
<td>3.327</td>
<td>51.17</td>
</tr>
<tr>
<td>anthracene</td>
<td>172.3</td>
<td>166.1</td>
<td>178.6</td>
<td>174.5</td>
<td>9.447</td>
<td>37</td>
<td>155.4</td>
<td>193.7</td>
<td>157.3</td>
<td>196</td>
<td>9.372</td>
<td>156.6</td>
</tr>
<tr>
<td>benzo[j]fluoranthene</td>
<td>281</td>
<td>259.9</td>
<td>302.1</td>
<td>276.3</td>
<td>5.924</td>
<td>4</td>
<td>260.2</td>
<td>292.5</td>
<td>260.5</td>
<td>292.8</td>
<td>5.883</td>
<td>264.9</td>
</tr>
<tr>
<td>benzo[k]fluoranthene</td>
<td>406</td>
<td>384.3</td>
<td>427.7</td>
<td>403.6</td>
<td>17.98</td>
<td>28</td>
<td>366.8</td>
<td>440.4</td>
<td>369.6</td>
<td>443.8</td>
<td>17.84</td>
<td>369.2</td>
</tr>
<tr>
<td>2-methylnaphthalene</td>
<td>79.23</td>
<td>77.23</td>
<td>81.24</td>
<td>77</td>
<td>7.78</td>
<td>34</td>
<td>61.19</td>
<td>92.81</td>
<td>63.88</td>
<td>96.88</td>
<td>7.748</td>
<td>61.85</td>
</tr>
<tr>
<td>chrysene</td>
<td>547</td>
<td>537.6</td>
<td>556.4</td>
<td>509.6</td>
<td>30.17</td>
<td>11</td>
<td>443.3</td>
<td>576</td>
<td>450.9</td>
<td>585.9</td>
<td>29.94</td>
<td>451.9</td>
</tr>
<tr>
<td>benzo[e]pyrene</td>
<td>590</td>
<td>560.2</td>
<td>619.8</td>
<td>613.3</td>
<td>26.23</td>
<td>32</td>
<td>559.8</td>
<td>666.7</td>
<td>563.8</td>
<td>671.4</td>
<td>26.02</td>
<td>563.1</td>
</tr>
<tr>
<td>benz[a]anthracene</td>
<td>378</td>
<td>368</td>
<td>388</td>
<td>388</td>
<td>16.43</td>
<td>32</td>
<td>354.6</td>
<td>421.5</td>
<td>357.2</td>
<td>424.6</td>
<td>16.28</td>
<td>356.8</td>
</tr>
<tr>
<td>benzo[ghi]perylene</td>
<td>488</td>
<td>475.6</td>
<td>500.4</td>
<td>457.7</td>
<td>34.62</td>
<td>37</td>
<td>387.6</td>
<td>527.8</td>
<td>396.8</td>
<td>540.4</td>
<td>34.52</td>
<td>389.3</td>
</tr>
<tr>
<td>indeno[1,2,3-cd]pyrene</td>
<td>608</td>
<td>588.3</td>
<td>627.7</td>
<td>576.4</td>
<td>49.92</td>
<td>38</td>
<td>475.4</td>
<td>677.5</td>
<td>490.4</td>
<td>698.8</td>
<td>49.75</td>
<td>478.1</td>
</tr>
<tr>
<td>benzo[a]pyrene</td>
<td>636</td>
<td>621.9</td>
<td>650.1</td>
<td>629.6</td>
<td>16.39</td>
<td>33</td>
<td>596.2</td>
<td>662.9</td>
<td>597.8</td>
<td>664.7</td>
<td>16.26</td>
<td>598.3</td>
</tr>
<tr>
<td>perylene</td>
<td>184</td>
<td>166.6</td>
<td>201.4</td>
<td>177.2</td>
<td>11.22</td>
<td>34</td>
<td>154.4</td>
<td>199.9</td>
<td>156.7</td>
<td>202.9</td>
<td>11.14</td>
<td>155.6</td>
</tr>
<tr>
<td>phenanthrene</td>
<td>539</td>
<td>521.6</td>
<td>556.4</td>
<td>505</td>
<td>28.97</td>
<td>42</td>
<td>446.5</td>
<td>563.5</td>
<td>452.5</td>
<td>571</td>
<td>28.82</td>
<td>449</td>
</tr>
<tr>
<td>benzo[b]fluoranthene</td>
<td>1089</td>
<td>1041</td>
<td>1137</td>
<td>1065</td>
<td>29</td>
<td>14</td>
<td>1003</td>
<td>1128</td>
<td>1006</td>
<td>1131</td>
<td>28.72</td>
<td>1010</td>
</tr>
<tr>
<td>pyrene</td>
<td>1129</td>
<td>1095</td>
<td>1163</td>
<td>1160</td>
<td>42.33</td>
<td>41</td>
<td>1075</td>
<td>1246</td>
<td>1080</td>
<td>1252</td>
<td>42.02</td>
<td>1080</td>
</tr>
<tr>
<td>fluoranthene</td>
<td>1094</td>
<td>1070</td>
<td>1118</td>
<td>1116</td>
<td>55.55</td>
<td>13</td>
<td>996.1</td>
<td>1237</td>
<td>1008</td>
<td>1251</td>
<td>54.91</td>
<td>1012</td>
</tr>
<tr>
<td>naphthalene</td>
<td>121.3</td>
<td>115.1</td>
<td>127.6</td>
<td>117.8</td>
<td>6.906</td>
<td>41</td>
<td>103.8</td>
<td>131.7</td>
<td>105.2</td>
<td>133.5</td>
<td>6.881</td>
<td>104.2</td>
</tr>
</tbody>
</table>

*Table 4.2: Example of study values and variances corrected for proportional bias using control data when errors are independent.*
<table>
<thead>
<tr>
<th>Compound</th>
<th>$\hat{\sigma}_{dDM}$</th>
<th>$\nu_d$</th>
<th>$lcl_{DM}$</th>
<th>$ucl_{DM}$</th>
<th>$lcl_F$</th>
<th>$ucl_F$</th>
<th>$\hat{\sigma}_{dB}$</th>
<th>$lcl_B$</th>
<th>$ucl_B$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2,3,5-trimethylnaphthalene</td>
<td>3.004</td>
<td>30</td>
<td>22.73</td>
<td>35.01</td>
<td>23.79</td>
<td>36.72</td>
<td>2.994</td>
<td>22.97</td>
<td>34.77</td>
</tr>
<tr>
<td>acenaphthene</td>
<td>3.22</td>
<td>30.83</td>
<td>43.27</td>
<td>32.90</td>
<td>23.98</td>
<td>32.75</td>
<td>2.034</td>
<td>23.99</td>
<td>31.97</td>
</tr>
<tr>
<td>dibenz[a,h]anthracene</td>
<td>1.952</td>
<td>26</td>
<td>26.89</td>
<td>34.92</td>
<td>27.33</td>
<td>35.41</td>
<td>1.934</td>
<td>27.23</td>
<td>34.79</td>
</tr>
<tr>
<td>1-methylnaphthalene</td>
<td>2.914</td>
<td>26</td>
<td>64.76</td>
<td>76.74</td>
<td>65.23</td>
<td>77.3</td>
<td>2.911</td>
<td>64.96</td>
<td>76.46</td>
</tr>
<tr>
<td>2,6-dimethylnaphthalene</td>
<td>1.19</td>
<td>20</td>
<td>33.6</td>
<td>38.57</td>
<td>33.76</td>
<td>38.77</td>
<td>1.18</td>
<td>33.81</td>
<td>38.45</td>
</tr>
<tr>
<td>fluorene</td>
<td>6.173</td>
<td>41</td>
<td>52.57</td>
<td>77.51</td>
<td>54.55</td>
<td>80.39</td>
<td>6.14</td>
<td>53.14</td>
<td>77.26</td>
</tr>
<tr>
<td>triphenylene</td>
<td>5.299</td>
<td>3</td>
<td>121.4</td>
<td>152.9</td>
<td>122.1</td>
<td>153.5</td>
<td>5.282</td>
<td>126.9</td>
<td>147.7</td>
</tr>
<tr>
<td>1-methylanthracene</td>
<td>3.471</td>
<td>20</td>
<td>44.29</td>
<td>58.79</td>
<td>45.16</td>
<td>59.87</td>
<td>3.434</td>
<td>45.03</td>
<td>58.55</td>
</tr>
<tr>
<td>anthracene</td>
<td>8.911</td>
<td>32</td>
<td>156.4</td>
<td>192.7</td>
<td>158.1</td>
<td>194.9</td>
<td>8.879</td>
<td>157.1</td>
<td>191.9</td>
</tr>
<tr>
<td>benzo[j]fluoranthene</td>
<td>5.935</td>
<td>4</td>
<td>259.8</td>
<td>292.8</td>
<td>260.2</td>
<td>293.2</td>
<td>5.893</td>
<td>264.9</td>
<td>287.9</td>
</tr>
<tr>
<td>benzo[k]fluoranthene</td>
<td>19.19</td>
<td>23</td>
<td>363.9</td>
<td>443.3</td>
<td>367</td>
<td>447.1</td>
<td>18.98</td>
<td>367</td>
<td>441.5</td>
</tr>
<tr>
<td>2-methylnaphthalene</td>
<td>7.843</td>
<td>35</td>
<td>61.07</td>
<td>92.93</td>
<td>63.77</td>
<td>97.03</td>
<td>7.806</td>
<td>61.81</td>
<td>92.57</td>
</tr>
<tr>
<td>chrysene</td>
<td>30.94</td>
<td>12</td>
<td>441.9</td>
<td>577.4</td>
<td>449.5</td>
<td>587.7</td>
<td>30.69</td>
<td>451.1</td>
<td>571.2</td>
</tr>
<tr>
<td>benzo[e]pyrene</td>
<td>24.25</td>
<td>23</td>
<td>563.2</td>
<td>663.4</td>
<td>566.6</td>
<td>668.2</td>
<td>24.2</td>
<td>565.3</td>
<td>661</td>
</tr>
<tr>
<td>benzo[a]anthracene</td>
<td>15.45</td>
<td>29</td>
<td>356.4</td>
<td>419.7</td>
<td>358.8</td>
<td>422.7</td>
<td>15.36</td>
<td>358</td>
<td>418.4</td>
</tr>
<tr>
<td>benzo[ghi]perylene</td>
<td>34.47</td>
<td>37</td>
<td>387.8</td>
<td>527.5</td>
<td>397.1</td>
<td>540</td>
<td>34.38</td>
<td>389.6</td>
<td>525.1</td>
</tr>
<tr>
<td>indeno[1,2,3-cd]pyrene</td>
<td>50.56</td>
<td>39</td>
<td>474.2</td>
<td>678.7</td>
<td>489.2</td>
<td>700.4</td>
<td>50.35</td>
<td>477.7</td>
<td>675.5</td>
</tr>
<tr>
<td>benzo[b]pyrene</td>
<td>15.37</td>
<td>26</td>
<td>598</td>
<td>661.1</td>
<td>599.6</td>
<td>662.8</td>
<td>15.33</td>
<td>599.2</td>
<td>665.9</td>
</tr>
<tr>
<td>perylene</td>
<td>10.64</td>
<td>29</td>
<td>155.4</td>
<td>198.9</td>
<td>157.7</td>
<td>201.6</td>
<td>10.61</td>
<td>156.1</td>
<td>198</td>
</tr>
<tr>
<td>phenanthrene</td>
<td>29.64</td>
<td>41</td>
<td>445.1</td>
<td>564.9</td>
<td>451.1</td>
<td>572.8</td>
<td>29.44</td>
<td>447.7</td>
<td>563.3</td>
</tr>
<tr>
<td>benzo[b]fluoranthene</td>
<td>31.06</td>
<td>12</td>
<td>998.1</td>
<td>1133</td>
<td>1001</td>
<td>1137</td>
<td>30.65</td>
<td>1007</td>
<td>1127</td>
</tr>
<tr>
<td>pyrene</td>
<td>43.83</td>
<td>36</td>
<td>1071</td>
<td>1249</td>
<td>1077</td>
<td>1256</td>
<td>43.43</td>
<td>1077</td>
<td>1248</td>
</tr>
<tr>
<td>fluoranthene</td>
<td>56.87</td>
<td>12</td>
<td>992.6</td>
<td>1240</td>
<td>1005</td>
<td>1255</td>
<td>56.19</td>
<td>1010</td>
<td>1231</td>
</tr>
<tr>
<td>naphthalene</td>
<td>6.941</td>
<td>41</td>
<td>103.7</td>
<td>131.8</td>
<td>105.2</td>
<td>133.6</td>
<td>6.908</td>
<td>104.2</td>
<td>131.5</td>
</tr>
</tbody>
</table>

Table 4.3. Example of study values and variances corrected for proportional bias using control data when errors are correlated.

For each compound, $\hat{\sigma}_{dDM} = \text{the estimated standard error of } \hat{d} \text{ using the Delta Method}; \nu_d = \text{the degrees of freedom of the variance of } \hat{d}; (lcl_{DM}, ucl_{DM}) = 95\% \text{ confidence interval for } \hat{d} \text{ using the Delta Method}; (lcl_F, ucl_F) = 95\% \text{ confidence interval for } \hat{d} \text{ using the Fieller’s Method}; \hat{\sigma}_{dB} = \text{bootstrap estimate of standard error of } \hat{d}; (lcl_B, ucl_B) = \text{the 95\% bootstrap confidence interval.}$
Table 4.4. Example of study values corrected for proportional bias using control data, contrasting when errors are assumed independent or correlated.

For each compound, $\hat{\rho}_{YW} = $ estimated correlation between the control and study values; $(lcl_{F,ind}, ucl_{F,ind}) = $ a 95% confidence interval for the observed mean study value using Fieller’s method and assuming independent errors; $(lcl_{F,corr}, ucl_{F,corr}) = $ a 95% confidence interval for the corrected study value using the Fieller’s method and assuming correlated errors.
Table 4.5. Example of estimated bias of the study values corrected for proportional bias using control data, contrasting when errors are assumed independent or correlated.

For each compound, $\hat{\rho}_{YW}$ = estimated correlation coefficient of the control data and study data. Assuming independent errors: $\hat{\text{bias}}_{DM,\text{ind}} = \text{the estimated bias of } \hat{d}$ using the delta method and $\hat{\text{bias}}_{B,\text{ind}} = \text{bootstrap estimate of bias in } \hat{d}$. Assuming correlated errors: $\hat{\text{bias}}_{DM,\text{corr}} = \text{the estimated bias of } \hat{d}$ using the delta method and $\hat{\text{bias}}_{B,\text{corr}} = \text{bootstrap estimate of bias in } \hat{d}$.

<table>
<thead>
<tr>
<th>Compound</th>
<th>$\hat{\rho}_{YW}$</th>
<th>$\hat{\text{bias}}_{DM,\text{ind}}$</th>
<th>$\hat{\text{bias}}_{B,\text{ind}}$</th>
<th>$\hat{\text{bias}}_{DM,\text{corr}}$</th>
<th>$\hat{\text{bias}}_{B,\text{corr}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2,3,5-trimethylnaphthalene</td>
<td>0.69</td>
<td>0.0207</td>
<td>0.0217</td>
<td>0.0140</td>
<td>0.0142</td>
</tr>
<tr>
<td>acenaphthene</td>
<td>-0.89</td>
<td>0.0459</td>
<td>0.0471</td>
<td>0.0548</td>
<td>0.0548</td>
</tr>
<tr>
<td>dibenz[a,h]anthracene</td>
<td>0.28</td>
<td>0.0053</td>
<td>0.0176</td>
<td>0.0009</td>
<td>0.0120</td>
</tr>
<tr>
<td>acenaphthylene</td>
<td>-0.52</td>
<td>0.0123</td>
<td>0.0179</td>
<td>0.0203</td>
<td>0.0253</td>
</tr>
<tr>
<td>biphenyl</td>
<td>-1.00</td>
<td>0.0087</td>
<td>0.0118</td>
<td>0.0159</td>
<td>0.0157</td>
</tr>
<tr>
<td>1-methylphenanthrene</td>
<td>0.92</td>
<td>0.0020</td>
<td>0.0038</td>
<td>0.0002</td>
<td>-0.0004</td>
</tr>
<tr>
<td>2,6-dimethylnaphthalene</td>
<td>0.52</td>
<td>0.0086</td>
<td>0.0111</td>
<td>0.0064</td>
<td>0.0083</td>
</tr>
<tr>
<td>fluorene</td>
<td>-0.95</td>
<td>0.0377</td>
<td>0.0374</td>
<td>0.0467</td>
<td>0.0444</td>
</tr>
<tr>
<td>triphenylene</td>
<td>0.49</td>
<td>0.0260</td>
<td>0.0701</td>
<td>-0.0062</td>
<td>0.0287</td>
</tr>
<tr>
<td>1-methylnaphthalene</td>
<td>-0.43</td>
<td>0.0500</td>
<td>0.0538</td>
<td>0.0573</td>
<td>0.0609</td>
</tr>
<tr>
<td>anthracene</td>
<td>1.00</td>
<td>0.0640</td>
<td>0.0744</td>
<td>0.0359</td>
<td>0.0321</td>
</tr>
<tr>
<td>benzo[j]fluoranthene</td>
<td>-0.01</td>
<td>0.0044</td>
<td>0.0492</td>
<td>0.0047</td>
<td>0.0495</td>
</tr>
<tr>
<td>benzo[k]fluoranthene</td>
<td>-0.87</td>
<td>0.0665</td>
<td>0.1066</td>
<td>0.1222</td>
<td>0.1458</td>
</tr>
<tr>
<td>2-methylnaphthalene</td>
<td>-0.57</td>
<td>0.0442</td>
<td>0.0431</td>
<td>0.0504</td>
<td>0.0487</td>
</tr>
<tr>
<td>chrysene</td>
<td>-0.97</td>
<td>0.2722</td>
<td>0.2813</td>
<td>0.3178</td>
<td>0.3154</td>
</tr>
<tr>
<td>benzo[e]pyrene</td>
<td>0.92</td>
<td>0.0927</td>
<td>0.1508</td>
<td>0.0111</td>
<td>0.0208</td>
</tr>
<tr>
<td>benz[a]anthracene</td>
<td>0.96</td>
<td>0.1182</td>
<td>0.1380</td>
<td>0.0780</td>
<td>0.0794</td>
</tr>
<tr>
<td>benzo[ghi]perylene</td>
<td>0.45</td>
<td>0.0432</td>
<td>0.0467</td>
<td>0.0312</td>
<td>0.0311</td>
</tr>
<tr>
<td>indeno[1,2,3-cd]pyrene</td>
<td>-0.98</td>
<td>0.0970</td>
<td>0.1045</td>
<td>0.1524</td>
<td>0.1332</td>
</tr>
<tr>
<td>benzo[a]pyrene</td>
<td>0.99</td>
<td>0.0406</td>
<td>0.0684</td>
<td>0.0148</td>
<td>0.0116</td>
</tr>
<tr>
<td>perylene</td>
<td>0.68</td>
<td>0.0321</td>
<td>0.0614</td>
<td>-0.0037</td>
<td>0.0136</td>
</tr>
<tr>
<td>phenanthrene</td>
<td>-0.76</td>
<td>0.0899</td>
<td>0.1112</td>
<td>0.1285</td>
<td>0.1413</td>
</tr>
<tr>
<td>benzo[b]fluoranthene</td>
<td>-0.61</td>
<td>0.0795</td>
<td>0.1783</td>
<td>0.1373</td>
<td>0.2269</td>
</tr>
<tr>
<td>pyrene</td>
<td>-0.65</td>
<td>0.1263</td>
<td>0.1903</td>
<td>0.1817</td>
<td>0.2361</td>
</tr>
<tr>
<td>fluoranthene</td>
<td>-0.38</td>
<td>1.0210</td>
<td>1.0740</td>
<td>1.0870</td>
<td>1.1380</td>
</tr>
<tr>
<td>naphthalene</td>
<td>-0.26</td>
<td>0.0034</td>
<td>0.0119</td>
<td>0.0053</td>
<td>0.0139</td>
</tr>
</tbody>
</table>
Table 4.6. Example of estimated within material variance for study data, under proportional bias.

For each compound, $\hat{\sigma}_V^2$ = the estimated within material variance; $(\text{lcl}(\sigma_V^2), \text{ucl}(\sigma_V^2))$ = a 95% confidence interval for the within material variance; $\hat{Pr}(\hat{\sigma}_V^2 < 0)$ = the estimated probability of getting a negative estimate of the within material variance. All under the assumption of constant measurement error variance for each compound across samples.
4.2 One lab, multiple compounds at one point in time

Often all or many compounds in a material undergo the same processes and transformations in order to be measured and thus any error related to the measurement procedure is similar across compounds. In such cases, the assumption of sharing parameters among a whole set of compounds appears to be reasonable and the information across compounds can be used to estimate the parameters. Our discussion here assumes a linear bias model is shared across compounds. However, even if all the compounds in a material undergo the same transformations, each may be calibrated separately using a specific reference or standard. In such cases the common linear model is hard to justify, since a different linear model may be required for each compound.

4.2.1 Working with the Control Data Only

As before, there is no interest on pooling information across different participants. However, there is interest on testing whether a participant’s method is biased.

Additional caution must be exercised here in order to handle the measurement error for the predictor. That is, we are now using a model similar to (4.1) but we only have $\hat{x}$ rather than $x$, as discussed below.

4.2.1.1 The Model

Depending on the specific design of the experiment this problem can be addressed by working directly with the detailed data (using the per replicate model) or with the summary data (using the means model).

The per replicate model

The model for the replicate observed values for compound $a$ in the control material is
\[ Y_{ak} | x_a = \beta_0 + \beta_1 x_a + (\beta_1 U_{ak} + \epsilon_{ak}), \]  

(4.51)

where \( a = 1 \) to \( A \), \( k = 1 \) to \( K_a \), \( N = \sum_a K_a \), \( E[U_{ak}] = 0 \), \( V[U_{ak}] = \sigma_{Ua}^2 \), \( E[\epsilon_{ak}] = 0 \), \( V[\epsilon_{ak}] = \sigma_{\epsilon a}^2 \) and \( (\beta_0, \beta_1, x_a, \sigma_{\epsilon a}^2, \sigma_{Ua}^2) \) are unknown parameters. The \((U_{ak}, \epsilon_{ak})\) are independent for a given \((a, k)\) and are independent among \((a, k)\). The line \( \beta_0 + \beta_1 x_a \) is the expected response at level \( x_a \), \( \beta_1 U_{ak} \) is the error due to heterogeneity in the material and \( \epsilon_{ak} \) is the error in the equation. There is no measurement error in the response. Assume that \( \hat{x}_a, \sigma_{\epsilon a}^2 \) and \( \sigma_{Ua}^2 \) are known unbiased estimators for \( x_a, \sigma_{Ua}^2 \), and \( V[\hat{x}_a] \).

The reference value \( \hat{x}_a \) is an estimator of \( x_a \), plus some additive measurement error; that is

\[ \hat{x}_a | x_a = x_a + e_{\hat{x}a}, \]  

(4.52)

\[ e_{\hat{x}a} | x_a \sim (0, \sigma_{\hat{x}a}^2), \]

where \( e_{\hat{x}a} | x_a \) is the error in \( \hat{x}_a \) as an estimator of \( x_a \).

Note that the same reference value is used for all the replicates of a compound \( \hat{x}_{ak} = \hat{x}_a \), hence the measurement error associated with the replicates on a compound are correlated; that is \( Cov[\hat{x}_{ak}, \hat{x}_{ak'}] = \sigma_{\hat{x}a}^2 \). This implies that the covariance matrix of \( \hat{x}_{ak} \) is a block diagonal matrix \( = diag(\sigma_{\hat{x}1}^2, \ldots, \sigma_{\hat{x}A}^2) \), with \( J_a \) a \( K_a \times K_a \) square matrix with all its entries equal to 1.

**The mean model**

The mean model for the control data becomes:

\[ \bar{Y}_a | x_a = \beta_0 + \beta_1 x_a + (\beta_1 \bar{U}_a + \bar{\epsilon}_a), \]  

(4.53)

where \( a = 1 \) to \( A \), \( \bar{U}_a = \frac{1}{K_a} \sum_{k=1}^{K_a} U_{ak} \), \( E[\bar{U}_a] = 0 \), \( V[\bar{U}_a] = \sigma_{\bar{U}a}^2 = \sigma_{Ua}^2 / K_a \), \( \bar{\epsilon}_a = \frac{1}{K_a} \sum_{k=1}^{K_a} \epsilon_{ak} \), \( E[\bar{\epsilon}_a] = 0 \), \( V[\bar{\epsilon}_a] = \sigma_{\bar{\epsilon}a}^2 = \sigma_{\epsilon a}^2 / K_a \).

Each reference value \( \hat{x}_a \) is used just once for each compound. Hence working with the means model avoids the issue of correlated measurement errors over replicates.
This is the model that we use for the remainder of the section. Note that with respect to the measurement error literature $\beta_1 \bar{U}_a + \bar{\epsilon}_a$ plays the role of error in the equation and that the variances can change across compounds. However, this is a combination of error in the equation $\bar{\epsilon}_a$ and error due to heterogeneity in the material $\bar{U}_a$. Let us refer to it as a pseudo-error in the equation $\bar{\eta}_a$, with variance

$$
\sigma^2_{\bar{\eta}_a} = \beta_1^2 \sigma^2_{\bar{U}_a} + \sigma^2_{\bar{\epsilon}_a},
$$

(4.54)

**Comment:** It may be tempting to use the model by conditioning on $\bar{X}_a$, after all these are the true allocated values. This yields

$$
\bar{Y}_a | \bar{X}_a = \beta_0 + \beta_1 \bar{X}_a + \bar{\epsilon}_a
$$

(4.55)

and this suggests using the reference value $\hat{x}_a$ as an estimator of $\bar{X}_a$, plus some additive measurement error. However,

$$
\hat{x}_a | \bar{X}_a = \bar{X}_a - \bar{U}_a + \epsilon_{\hat{x}_a},
$$

$$
(-\bar{U}_a + \epsilon_{\hat{x}_a}) | \bar{X}_a \sim (x_a - \bar{X}_a, \sigma^2_{\hat{x}_a}),
$$

(4.56)

where $(-\bar{U}_a + \epsilon_{\hat{x}_a})$ is the error in $\hat{x}_a$ as an estimator of $\bar{X}_a$. The new measurement error term has conditional mean 0 only if $\bar{X}_a = x_a$, since conditioning in $\bar{X}_a$ fixes $\bar{U}_a$. This violates the assumption of additive measurement error.

### 4.2.1.2 Coefficient Estimation

It has long been known that a regression based on observed values with error produces biased estimators of the regression coefficients (Fuller (1987) [22]). The problem of fitting a line with one or both variables measured with error has been studied and revisited by numerous researchers. The approaches to solve it also have been labeled with different names over time: error in variables regression, measurement error regression, orthogonal least squares, total least squares (Casella and Berger (2002) [11], Fuller (1987) [22]).

Ignoring any variance structure and assuming balancedness, the naive estimators are obtained by regressing the observed data ($\bar{Y}_a, \hat{x}_a$):
\[ \hat{\beta}_{1,\text{naive}} = \frac{S_{XY}}{S_X^2}, \quad (4.57) \]

where \( \bar{Y} = \frac{1}{A} \sum_{a=1}^{A} Y_a, \bar{x} = \frac{1}{A} \sum_{a=1}^{A} \hat{x}_a, \]
\[ S_{XY} = \frac{1}{A-1} \sum_{a=1}^{A} (Y_a - \bar{Y})(\hat{x}_a - \bar{x}), \quad \hat{S}_Y^2 = \frac{1}{A-1} \sum_{a=1}^{A} (\hat{Y}_a - \bar{Y})^2 \]
and \( S_X^2 = \frac{1}{A-1} \sum_{a=1}^{A} (\hat{x}_a - \bar{x})^2. \) It is well known that the naive estimators are both biased and inconsistent. It is straightforward to show that

\[ E[\bar{Y}] = \beta_0 + \beta_1 \bar{x}, \]
\[ E[S_X^2] = \sigma_x^2 + \sigma_x^2, \]
\[ E[S_{XY}] = \beta_1 \sigma_x^2, \] so \( \beta_1 = E[S_{XY}]/\sigma_x^2, \) and

\[ E[S_Y^2] = \frac{1}{A-1} \left( \sum_a E[\hat{Y}_a^2] - AE[\hat{Y}^2] \right) \]
\[ = \frac{1}{A-1} \left( \beta_1^2 \left( \sum_a x_a^2 - A\bar{x}^2 \right) + \sum_a V[Y_a] - \frac{1}{A} \sum_a V[Y_a] \right) \]
\[ = \beta_1^2 \sigma_x^2 + \beta_1 \overline{\sigma_\epsilon^2} + \sigma_\epsilon^2, \]

with \( \sigma_x^2 = \frac{1}{A-1} \sum_a (x_a - \bar{x})^2, \overline{\sigma_\epsilon^2} = \frac{1}{A} \sum_a \sigma_{\epsilon a}^2, \overline{\sigma_\epsilon^2} = \frac{1}{A} \sum_a \sigma_{\epsilon a}^2 \) and \( \overline{\sigma_\epsilon^2} = \frac{1}{A} \sum_a \sigma_{\epsilon a}^2. \)

The exact expressions for the true coefficients can be obtained from (4.58) as

\[ \beta_1 = \frac{\sigma_{xy}}{\sigma_x^2} = \frac{E[S_{XY}]}{E[S_X^2] - \overline{\sigma_x^2}}, \quad (4.59) \]
\[ \beta_0 = E[\bar{Y}] - \beta_1 E[\bar{X}]. \]

Biased but consistent moment corrected estimators (Buonaccorsi(1985) [6], Buonaccorsi(2010) [9]) are obtained by plugging unbiased estimators in (4.58),

\[ \hat{\beta}_{1,\text{mm}} = \frac{S_{XY}}{\hat{\sigma}_x^2} = \frac{S_{XY}}{S_X^2 - \overline{\sigma_x^2}}, \quad (4.60) \]
\[ \hat{\beta}_{0,\text{mm}} = \bar{Y} - \hat{\beta}_{1,\text{mm}} \bar{x}, \]

where \( \overline{\sigma_x^2} = \frac{1}{A} \sum_{a=1}^{A} \hat{\sigma}_{xa}^2 \) is the mean of the measurement error variances and \( \hat{\sigma}_{xa}^2 \) can change across compounds.
The variances of the error in the equation are $\sigma^2_{\epsilon_a}$, however only the related quantity $\bar{\sigma}^2_{\epsilon} = \sum_a \frac{\sigma^2_{\epsilon_a}}{AK_a}$ can be estimated from (4.58)

$$\bar{\sigma}^2_{\epsilon} = S^2_Y - \hat{\beta}^2_{1,mm}(S^2_{\hat{x}} - \bar{\sigma}^2_{\hat{x}} + \bar{\sigma}^2_{\hat{U}}).$$  (4.61)

For the quantities in (4.60) to be proper estimators, $\hat{\sigma}^2_x = S^2_{\hat{x}} - \bar{\sigma}^2_{\hat{x}}$ and $\bar{\sigma}^2_{\epsilon}$ should be non-negative. If $\hat{\sigma}^2_x$ is less than zero or the estimates of the variance of the error in the equation are less than zero then some modification to the estimator of $\beta_1$ can be considered as discussed in Section 4.2.1.6.

Note that the difference between the naive estimators and the moment corrected estimators depends on the denominator $S^2_{\hat{x}} - \bar{\sigma}^2_{\hat{x}}$. The effect of measurement error becomes more important when the variance of the measurement error $\bar{\sigma}^2_{\hat{x}}$ gets larger with respect to the dispersion on the predictor $\hat{x}_a$. Hence, when we have different compounds with a shorter range of values or the same compound with similar values from different CRMs the measurement error becomes more important.

### 4.2.1.3 Interval Estimators and Estimators for the Coefficient Covariance

**Wald type confidence intervals**

Under general conditions $\hat{\beta} \sim N(\beta, \Sigma_{\hat{\beta}})$ approximately and if an estimator of $\Sigma_{\hat{\beta}}$ is available,

$$\hat{\Sigma}_{\hat{\beta}} = \begin{bmatrix} v_{00} & v_{01} \\ v_{01} & v_{11} \end{bmatrix},$$  (4.62)

then an approximate confidence interval for $\hat{\beta}_i$ is

$$(\hat{\beta}_i - z_{\alpha/2}\sqrt{v_{ii}}, \hat{\beta}_i + z_{\alpha/2}\sqrt{v_{ii}}).$$  (4.63)

There are several ways for calculating the estimator $\hat{\Sigma}_{\hat{\beta}}$ (see for example Fuller(1987) [22] or Buonaccorsi(2010) [9]). Specifically we point to the robust and the asymptotic normal-based estimators presented in (5.12) and (5.15) in Buonaccorsi(2010) [9].
respectively. The robust estimator allows either $\sigma^2_{\hat{x}_a}$ or $\beta_1^2 \sigma_\epsilon^2 + \sigma^2_{\hat{x}_a}$ to change with $a$, while the so called normal-based estimator assumes the measurement error variances are constant. Notationally $\hat{\Sigma}_{\beta,R}$ stands for the robust estimate and $\hat{\Sigma}_{\beta,N}$ stands for the “normal” estimate. Here, for simple linear regression the robust estimator is

$$\hat{\Sigma}_{\beta,R} = M_{XX}^{-1} \hat{H} R M_{XX}^{-1},$$

$$\hat{H} = \frac{1}{A(A-2)} \sum_a \hat{\Delta}_a \hat{\Delta}_a',$$

$$\hat{\Delta}_a = \begin{bmatrix} Y_a - \hat{\beta}_{0,\text{mm}} - \hat{\beta}_{1,\text{mm}} \hat{x}_a \\ \hat{x}_a (Y_a - \hat{\beta}_{0,\text{mm}} - \hat{\beta}_{1,\text{mm}} \hat{x}_a) + \hat{\beta}_{1,\text{mm}} \hat{\sigma}^2_{\hat{x}_a} \end{bmatrix},$$

(4.64)

$$M_{XX} = \begin{bmatrix} 1 & \bar{x} \\ \bar{x} & \bar{x}^2 - \hat{\sigma}^2_{\hat{x}}. \end{bmatrix}$$

and

$$\hat{\Sigma}_{\beta,N} = M_{XX}^{-1} \hat{H} N M_{XX}^{-1},$$

$$\hat{H} = \frac{1}{A^2} \sum_a (\hat{X}_a \hat{X}_a' v_a^2 + \hat{Z}_a \hat{Z}_a'),$$

$$\hat{Z}_a = \begin{bmatrix} 0 \\ -\hat{\beta}_{1,\text{mm}} \hat{\sigma}^2_{\hat{x}_a} \end{bmatrix}, \hat{X}_a = \begin{bmatrix} 1 \\ \hat{x}_a \end{bmatrix},$$

(4.65)

$$v_a^2 = \hat{\sigma}^2_{\hat{\epsilon}} + \hat{\beta}_{1,\text{mm}} \hat{\sigma}^2_{\hat{x}_a}.$$

The robust estimator allows for changing measurement error variances with no explicit error in the equation variances required, this information is hidden within the use of the residuals in the $\hat{\Delta}_a$ terms. The normal based estimator requires an explicit estimate of the variance of the pseudo error in the equation $(\hat{\sigma}^2_{\hat{\epsilon}} = \beta_1^2 \sigma_\epsilon^2 + \sigma^2_{\hat{x}})$ and it assumes it is constant across compounds. Under the presence of heterogeneity in the material it is hard to justify the use of the normal-based estimator and hence the robust estimator is preferable.
Under normality and the variance in (4.54) being constant over \( a \) there are alternative ways to estimate \( \text{Cov}[\hat{\beta}] \), we refer for example to Buonaccorsi(2010) [9] section 5.4.2.

An approximate Z-test of \( H_0 : \beta_i = b \) uses \( Z = \frac{(\hat{\beta}_i - b)}{\sqrt{\sigma^2_i}} \) assuming \( Z \) is distributed as a standard normal.

### 4.2.1.4 Testing for unbiasedness of the method

A general linear hypothesis test can be conducted (Ravishanker and Dey (2002) [43]) to test the hypothesis of an unbiased model \((H_0 : \beta_0 = 0, \beta_1 = 1)\). Consider \( H_0 : \beta = d \), where \( \beta = [\beta_0, \beta_1]^T \) and \( d = [0, 1]^T \). Then, under normality and known covariance matrix \( \Sigma_{\beta} \) we have: \((\hat{\beta} - d)^T \Sigma_{\beta}^{-1} (\hat{\beta} - d) \sim \chi^2_2\), under \( H_0 \). In practice \( \Sigma_{\beta} \) is unknown, instead we use the estimated covariance matrix \( \hat{\Sigma}_{\beta} \) and we define

\[
\chi^{2*} = (\hat{\beta} - d)^T \hat{\Sigma}_{\beta}^{-1} (\hat{\beta} - d)
\]  

(4.66)

where \( \chi^{2*} \) is distributed approximately as a \( \chi^2 \) random variable, with 2 degrees of freedom, under \( H_0 \). Hence, \( H_0 \) can be rejected at a significance level of approximately \( \alpha \) if

\[
\chi^{2*} > \chi^2_{(1-\alpha;2)}.
\]  

(4.67)

### 4.2.1.5 Estimation of the Variance of the Error in Equation

Let us assume first that the variance of the error in the equation is constant, so \( \sigma^2_{ea} = \sigma^2_\epsilon, \forall a \). Using (4.58) we can write

\[
\overline{\sigma}_\epsilon^2 = \frac{\sum_{a=1}^{A} \sigma^2_{ea}}{A} = \sum_{a=1}^{A} \frac{\sigma^2_{ea}}{AK_a} = \sigma^2_\epsilon \sum_{a=1}^{A} \frac{1}{AK_a}
\]

\[
= E[S_Y^2] - \beta_1^2 \left( E[S_X^2] - \overline{\sigma}_x^2 + \overline{\sigma}_U^2 \right)
\]

(4.68)

and a biased but consistent method of moments estimator of \( \sigma^2_\epsilon \) is
\[ \hat{\sigma}_{\epsilon,mm}^2 = \frac{S_Y^2 - \hat{\beta}_{1,mm}^2 \left( S_X^2 - \bar{\sigma}_x^2 + \bar{\sigma}_U^2 \right)}{\sum_{a=1}^A \frac{1}{K_a}}. \] (4.69)

This extends a special case of equation (4.9) in Buonaccorsi(2010) [9] when there is no measurement error in the response but there is some heterogeneity in the material. This estimator is consistent since \( \hat{\beta}_{1,mm} \) is consistent for \( \beta_1 \).

Another approach is to consider the mean square error while regressing the error-prone data using the corrected coefficients

\[
MSE_c = \frac{1}{A-2} \sum_a \bar{Y}_a = \frac{1}{A-2} \sum_a (\bar{Y}_a - (\hat{\beta}_{0,mm} + \hat{\beta}_{1,mm} \hat{x}_a))^2.
\] (4.70)

then

\[
MSE_c = \frac{A-1}{A-2} \left( S_Y^2 + \hat{\beta}_{1,mm}^2 S_X^2 - 2\hat{\beta}_{1,mm} S_{XY} \right)
= \frac{A-1}{A-2} \left( S_Y^2 - \hat{\beta}_{1,mm}^2 (S_X^2 - 2\bar{x}^2) \right)
= \frac{A-1}{A-2} \left( \hat{\sigma}_{\epsilon,mm}^2 + \hat{\beta}_{1,mm}^2 (\bar{\sigma}_x^2 + \bar{\sigma}_U^2) \right)
\rightarrow \sigma_{\epsilon}^2 + \beta_1^2 (\bar{\sigma}_x^2 + \bar{\sigma}_U^2),
\] (4.71)

obtained by replacing \( \hat{\beta}_{0,mm} \) and \( S_{XY} \) using (4.60) and the identity in (4.69). Hence a biased but consistent estimator for \( \sigma_{\epsilon}^2 \) is

\[
\hat{\sigma}_{\epsilon}^2 = \frac{MSE_c - \hat{\beta}_{1,mm}^2 \left( \bar{\sigma}_x^2 + \bar{\sigma}_U^2 \right)}{\sum_{a=1}^A \frac{1}{K_a}}.
\] (4.72)

Since we are using (4.69) to derive (4.72) the assumption of constant variance of the error in the equation (\( \sigma_{\epsilon a}^2 = \sigma_{\epsilon}^2, \forall a \)) holds. We can rewrite (4.69) as

\[
\hat{\sigma}_{\epsilon,mm}^2 = \frac{(\frac{A-2}{A-1}) MSE_c - \hat{\beta}_{1,mm}^2 \left( \bar{\sigma}_x^2 + \bar{\sigma}_U^2 \right)}{\sum_{a=1}^A \frac{1}{K_a}}.
\]
This makes clear that (4.69) and (4.72) are almost the same for large samples and each could be negative in which case the estimate is set equal to 0.

A third approach is obtained by using the replicated data and defining the sample variance for compound $a$

$$s^2_{Ya} = \frac{1}{K_a - 1} \sum_k (Y_{ak} - \bar{Y}_a)^2. \quad (4.73)$$

It is straightforward to show that

$$E[s^2_{Ya}] = \beta_1^2 \sigma^2_{Ua} + \sigma^2_{ea}. \quad (4.74)$$

Hence we can first obtain point estimates of the variance of the error in the equation for each compound as

$$\hat{\sigma}^2_{ea} = s^2_{Ya} - \hat{\beta}_1^2 \hat{\sigma}^2_{Ua}. \quad (4.75)$$

This is slightly different from what we obtained in (3.5) from Chapter 3, since we are exploiting the linear regression model. However it requires $\hat{\sigma}^2_{Ua}$ to be known. We could use $\hat{\beta}_1^2 - \hat{V}[\hat{\beta}_1]$ as estimate of $\beta_1^2$. If there is no heterogeneity in the material ($\sigma^2_{Ua} = 0$) then we always obtain a non-negative estimate $s^2_{Ya}$ of the individual variances of errors in the equation and same as in previous models. If the estimate in (4.75) is negative it must be set to zero.

In addition if it could be assumed that all $\sigma^2_{ea} = \sigma^2_{e}$ then we could use individual truncation as

$$\hat{\sigma}^2_{e,rep} = \frac{1}{A} \sum_{a=1}^{A} max(0, \hat{\sigma}^2_{ea}). \quad (4.76)$$

The individual estimators $\hat{\sigma}^2_{ea}$ are biased but consistent. Introducing truncation on the individual estimates brings additional bias in the estimated constant variance for the error in the equation $\hat{\sigma}^2_{e,rep}$.
What would happen if we assume the among subsample heterogeneity variance $\sigma_{Ua}^2$ is zero but it is not? Basically the variance of the error in the equation $\sigma_{ea}^2$ is overestimated individually or on average as described by (4.69), (4.72) and (4.76).

4.2.1.6 Modified Estimators

In the case of negative estimates of $\sigma_x^2$ or negative estimates of the average variance of the error in the equation $\sigma_{\bar{e}}^2$, then the estimators of the coefficients in (4.60) need some modification.

If $\hat{\sigma}_x^2 = S_X^2 - \bar{\sigma}_x^2 > 0$, $\bar{\sigma}_U^2 > 0$ (there is heterogeneity in the material) and $\bar{\sigma}_{\bar{e}}^2 \leq 0$ (there is no error in the equation) in (4.61) then, setting $\bar{\sigma}_{\bar{e}}^2 = 0$ in (4.58) lead us to consider

$$\hat{\beta}_{1,ICH} = \frac{S_Y^2}{S_{XY}} \left( \frac{S_X^2 - \bar{\sigma}_x^2}{S_X^2 - \bar{\sigma}_x^2 + \bar{\sigma}_U^2} \right).$$

(4.77)

If $\hat{\sigma}_x^2 > 0$, $\bar{\sigma}_U^2 = 0$ and $\bar{\sigma}_{\bar{e}}^2 \leq 0$ in (4.60) then there is no heterogeneity in the material and no error in the equation. In that case (4.77) reduces to

$$\hat{\beta}_{1,IC} = \frac{S_Y^2}{S_{XY}}.$$

(4.78)

This is the inverse calibration estimator (equation 1.2.4 in Fuller(1987) [22]).

There is potential for problems if the denominator of $\hat{\beta}_{1,mm}$ in (4.60) is non-positive. This condition will be used as described in the following algorithms. Recall that the expression in the denominator is an estimator of $\sigma_x^2 \geq 0$. Hence a negative estimate is not acceptable since it is outside the parameter space. On the other hand, a zero estimate implies the predictor values are statistically the same quantity and the data does not support the model (4.51). Under no replicated data one option is to consider a reduced model such as (3.71) or (4.5).

One way to deal with negative estimates of $\sigma_x^2$ is by individually examining each compound. Note that $S_a^2$ and $\bar{X}^2$ can be estimated unbiasedly by $\hat{S}_a^2 - \hat{\sigma}_{xa}^2$ and
\( \bar{x}^2 - \frac{1}{A} \hat{\sigma}^2_{\bar{x}} \), respectively. Hence we can apply individual truncation to zero in order to ensure non negative terms

\[
\hat{\sigma}^2 = \frac{1}{A-1} \sum_a \max \left( (\hat{x}_a - \bar{x})^2 - \frac{A-2}{A} \hat{\sigma}^2_{\bar{x}a} - \frac{1}{A} \hat{\sigma}^2_{\bar{x}}, 0 \right).
\]  

(4.79)

Then we can substitute this estimator in (4.60). This estimator is consistent but it has augmented bias by construction when compared to \( S^2_{\bar{x}} - \hat{\sigma}^2_{\bar{x}} \).

Modified estimators were suggested by Fuller(1987) [22] where the negative estimate is attributed to a small sample size. We refer to Section 2.5 in Fuller(1987) [22] for details on a technique to handle negative estimates of \( \sigma^2_{\bar{x}} \), specifically equation 2.5.3. This strategy considers a modified denominator in (4.60) as

\[
\tilde{\beta}_1 = \frac{S_{\bar{X}Y}}{\tilde{H}_{xx} + \frac{\alpha}{A-1} \hat{\sigma}^2_{\bar{x}}}
\]  

(4.80)

where

\[
\tilde{H}_{xx} = \begin{cases} 
S^2_{\bar{X}} - \hat{\sigma}^2_{\bar{x}} & \text{if } \lambda \geq 1 + 1/(A-1), \\
S^2_{\bar{X}} - (\lambda - \frac{1}{A-1}) \hat{\sigma}^2_{\bar{x}} & \text{if } \lambda < 1 + 1/(A-1).
\end{cases}
\]  

(4.81)

and \( \lambda \) is the root of

\[
\begin{vmatrix}
S_{\bar{Y}^2} & S_{\bar{X}Y} \\
S_{\bar{X}Y} & S^2_{\bar{X}} - \lambda \hat{\sigma}^2_{\bar{x}}
\end{vmatrix} = 0.
\]

The parameter \( \alpha \) is determined by minimizing the mean square error of \( \tilde{\beta}_1 \).

4.2.1.7 Bootstrapping for inference

An alternative to the Wald method is the bootstrap toolbox. The so called two-stage parametric bootstrap technique will be used for this situation.

Using the two-stage parametric bootstrap is not an option when the estimate of variance \( \hat{\sigma}^2_{\bar{x}} \) is negative or when the distribution of the error in the equation is not specified. Here we assume normality of the error terms and \( \hat{\sigma}^2_{\bar{x}a} > 0 \) for at least one \( a \). The basic bootstrap cycle is described in Algorithm 1.
Algorithm 1 Bootstrap algorithm for linear coefficients under constant $\sigma^2_{\epsilon}$

for $b = 1$ to $B$ (large) do
  $\hat{x}_{ab} \leftarrow \hat{x}_a + \hat{\sigma}_x Z_{a1b}$,
  $\hat{\sigma}^2_{xab} \leftarrow \hat{\sigma}^2_x \chi^2(\nu_{\hat{\sigma}_x}) / \nu_{\hat{\sigma}_x}$,
  $\hat{Y}_{ab} \leftarrow \hat{\beta}_{0,mm} + \hat{\beta}_{1,mm} \hat{x}_a + Z_{a2b} \sqrt{\hat{\beta}^2_{1,mm} \hat{\sigma}^2_{\hat{U}_a} + \hat{\sigma}^2_{\epsilon}}$,
  where $(Z_{a1b}, Z_{a2b})$ are independent, each distributed as a standard normal, and $\hat{\sigma}^2_{\epsilon} = \hat{\sigma}_{\epsilon}^2$ as given in (4.75) (if we assume to change across compounds and replicates are available) or $\hat{\sigma}^2_{\epsilon} = \hat{\sigma}^2_{\epsilon} / K_a$ with $\hat{\sigma}^2_{\epsilon}$ given in (4.69) or (4.72) if we assume equal variance.

Under normality this is the same as using $\bar{Y}_{ab} = \hat{\beta}_{0,mm} + \hat{\beta}_{1,mm} \hat{x}_a + \hat{\beta}_{1,mm} \hat{U}_{ab} + \hat{\epsilon}_{ab}$, where $\hat{U}_{ab} \sim N(0, \hat{\sigma}^2_{\hat{U}_a} / K_a)$ and $\hat{\epsilon}_{ab} \sim N(0, \hat{\sigma}^2_{\epsilon} / K_a)$.

Apply (4.60) to $(\hat{x}_{ab}, \hat{Y}_{ab})$ to get $(\hat{\beta}_{0,b}, \hat{\beta}_{1,b})$.

Apply (4.69) or (4.72) to obtain $\hat{\sigma}^2_{\epsilon,b}$.

If $S^2_{\chi b} > \hat{\sigma}^2_{\hat{U}_a,b}$ and $\hat{\sigma}^2_{\epsilon,b} \leq 0$ then apply (4.77) to $(\hat{x}_{ab}, \hat{Y}_{ab})$ to get a modified estimator $(\hat{\beta}_{0,b}, \hat{\beta}_{1,b})$.

If $S^2_{\chi b} \leq \hat{\sigma}^2_{\hat{U}_a,b}$ and $\hat{\sigma}^2_{\epsilon,b} > 0$ then apply (4.80) to $(\hat{x}_{ab}, \hat{Y}_{ab})$ to get a modified estimator $(\hat{\beta}_{0,b}, \hat{\beta}_{1,b})$.

If $S^2_{\chi b} \leq \hat{\sigma}^2_{\hat{U}_a,b}$ and $\hat{\sigma}^2_{\epsilon,b} \leq 0$ then warning: model (4.51) is not supported by the data and drop the estimate.

end for

Bootstrap confidence intervals can be computed via the percentile method. For example, consider $B_0 = \{ \hat{\beta}_{0,b}, b = 1 \ to \ B \}$. A $(1 - \alpha)100\%$ CI for $\beta_0$ is obtained by computing

$\left( Q_{(\alpha/2)}(B_0), Q_{(1-\alpha/2)}(B_0) \right)$, \hspace{1cm} (4.82)

where $Q_{(1-\alpha/2)}(B_0)$ is the $\alpha^{th}$ quantile of $B_0$. The confidence interval for $\beta_1$ can be obtained in a similar way.

The bootstrap estimates of the biases in the coefficients are
\[
\text{bias}[\hat{\beta}_0]^* = \bar{\beta}_0^* - \hat{\beta}_{0,mm},
\]
\[
\text{bias}[\hat{\beta}_1]^* = \bar{\beta}_1^* - \hat{\beta}_{1,mm},
\]
\[
\bar{\beta}_0^* = \frac{1}{B} \sum_{b=1}^{B} \hat{\beta}_{0,b},
\]
\[
\bar{\beta}_1^* = \frac{1}{B} \sum_{b=1}^{B} \hat{\beta}_{1,b},
\]
(4.83)

and the bootstrap standard errors are

\[
\text{se}[\hat{\beta}_0]^* = \sqrt{\frac{1}{B - 1} \sum_{b=1}^{B} (\hat{\beta}_{0,b} - \bar{\beta}_0^*)^2},
\]
\[
\text{se}[\hat{\beta}_1]^* = \sqrt{\frac{1}{B - 1} \sum_{b=1}^{B} (\hat{\beta}_{1,b} - \bar{\beta}_1^*)^2},
\]
(4.84)

\[
\text{Cov}[\hat{\beta}_0, \hat{\beta}_1]^* = \frac{1}{B - 1} \sum_{b=1}^{B} (\hat{\beta}_{0,b} - \bar{\beta}_0^*) (\hat{\beta}_{1,b} - \bar{\beta}_1^*).
\]

Only resampling the original data to estimate the coefficients of the linear regression is not an option since we need to assume the \((\hat{x}_a, \bar{Y}_a)\) are independent and identically distributed.

Bootstrap inferences for \(\sigma^2_{\epsilon}\) could be carried out in a similar manner using the \((\hat{x}_{ab}, \bar{Y}_{ab})\) values and using (4.72) if constant variance of the error in the equations is assumed.

For methods using replicate values we can use (4.75) to obtaining CI for \(\sigma^2_{\epsilon a}\) for each compound by modifying the bootstrap cycle to include bootstrap replicates as described in the Algorithm 2.

Then we can proceed to obtain a percentile CI from the set \(\{\hat{\sigma}_{cab}^2, b = 1 to B\}\) for each compound. It is clear that using either the means bootstrap or the replicates bootstrap allows unequal variances and/or unequal number of replicates as needed.
Algorithm 2 Bootstrap algorithm for linear coefficients using replicates and changing error variances

\begin{algorithm}
\begin{algorithmic}
  \State $b = 1$ to $B$(large) \Do
    \State $\hat{x}_{ab} \Leftarrow \hat{x}_a + \hat{\sigma}_{xa} Z_{a1b}$,
    \State $\hat{\sigma}_{xab}^2 \Leftarrow \hat{\sigma}_{xa}^2 \chi^2(\nu_{xa}) / \nu_{xa}$,
    \State $\hat{Y}_{a,kb} \Leftarrow \hat{\beta}_{0,mm} + \hat{\beta}_{1,mm} \hat{x}_a + Z_{a2,kb} \sqrt{\hat{\beta}_{1,mm}^2 \hat{\sigma}^2_{Ua} + \hat{\sigma}^2_{xa}}$,
  \EndFor
  \where $k = 1, \ldots, K_a$, and $(Z_{a1b}, Z_{a2,kb})$ are independent, each distributed as a standard normal, $\forall a, k, b$,
  \State $\bar{Y}_{ab} \Leftarrow \frac{1}{K_a} \sum_k Y_{a,kb}$,
  \State $s^2_{\bar{Y}_{ab}} \Leftarrow \frac{1}{K_a - 1} \sum_k (Y_{a,kb} - \bar{Y}_{ab})^2$, \End
  \State Apply (4.60) to $(\hat{x}_{ab}, \bar{Y}_{ab})$ to get $(\hat{\beta}_{0,b}, \hat{\beta}_{1,b})$.
  \State Apply (4.75) to obtain $\hat{\sigma}^2_{\epsilon,ab}$.
  \If $S^2_{X \hat{b}} > \hat{\sigma}^2_{x \hat{b}}$ and $\hat{\sigma}^2_{\epsilon,ab} \leq 0 \forall a$ then apply (4.77) to $(\hat{x}_{ab}, \bar{Y}_{ab})$ to get a modified estimator $(\hat{\beta}_{0,b}, \hat{\beta}_{1,b})$.
  \If $S^2_{X \hat{b}} \leq \hat{\sigma}^2_{x \hat{b}}$ and $\hat{\sigma}^2_{\epsilon,ab} > 0$ for at least one $a$ then apply (4.80) to $(\hat{x}_{ab}, \bar{Y}_{ab})$ to get a modified estimator $(\hat{\beta}_{0,b}, \hat{\beta}_{1,b})$.
  \If $S^2_{X \hat{b}} \leq \hat{\sigma}^2_{x \hat{b}}$ and $\hat{\sigma}^2_{\epsilon,ab} \leq 0 \forall a$ then warning: model (4.51) is not supported by the data and drop this estimate.
\end{algorithmic}
\end{algorithm}

The bootstrap CIs can also be used to test unbiasedness for example using the two CIs and applying Bonferroni’s method. We can accept that the method/laboratory is unbiased ($H_0 : \beta_0 = 0, \beta_1 = 1$) at a significance level $\alpha$ if

$$0 \in \left( Q_{(\alpha/4)}(B_0), Q_{(1-\alpha/4)}(B_0) \right) \quad \text{and} \quad 1 \in \left( Q_{(\alpha/4)}(B_1), Q_{(1-\alpha/4)}(B_1) \right).$$

Another way to test for unbiasedness is by resampling under the null hypothesis and computing a bootstrap $p - value$ associated with the $\chi^2$ test statistic described in (4.66). This alternative requires modifying slightly the bootstrap cycle by adding a couple of instructions to obtain the quantity of interest. First we need the estimate of the covariance of the bootstrap coefficients, then we can obtain the test statistic from each resample under the null hypothesis $H_0 : (\hat{\beta}_{0,mm}, \hat{\beta}_{1,mm}) = (0, 1)$. This is detailed in the Algorithm 3.
Algorithm 3 Bootstrap algorithm for linear coefficients when testing for unbiasedness with constant variance $\sigma^2$

(Resample as earlier but with $\hat{\beta}_{0,mm} = 0$ and $\hat{\beta}_{1,mm} = 1$).

for $b = 1$ to $B$ (large) do
  $\hat{x}_a \leftarrow \hat{x}_a + \hat{\sigma}_2 Z_{a1b}$,
  $\hat{\sigma}_{2ab} \leftarrow \hat{\sigma}_{2a} \chi^2(\nu_{2a})/\nu_{2a}$,
  $\hat{Y}_{ab} \leftarrow \hat{x}_a + Z_{a2b} \sqrt{\hat{\beta}_{0,mm}^2 + \hat{\sigma}_{2a}^2}$,
  where $(Z_{a1b}, Z_{a2b})$ are independent, each distributed as a standard normal.
  Apply (4.60) to $(\hat{x}_{ab}, \hat{Y}_{ab})$ to get $(\hat{\beta}_{0,b}, \hat{\beta}_{1,b})$.
  Apply (4.69) or (4.72) to obtain $\hat{\sigma}_{r,b}^2$.
  If $S_{Xb}^2 > \hat{\sigma}_{X,b}^2$ and $\hat{\sigma}_{r,b}^2 \leq 0$ then apply (4.77) to $(\hat{x}_{ab}, \hat{Y}_{ab})$ to get a modified estimator $(\hat{\beta}_{0,b}, \hat{\beta}_{1,b})$.
  If $S_{Xb}^2 \leq \hat{\sigma}_{X,b}^2$ and $\hat{\sigma}_{r,b}^2 > 0$ then apply (4.80) to $(\hat{x}_{ab}, \hat{Y}_{ab})$ to get a modified estimator $(\hat{\beta}_{0,b}, \hat{\beta}_{1,b})$.
  If $S_{Xb}^2 \leq \hat{\sigma}_{X,b}^2$ and $\hat{\sigma}_{r,b}^2 \leq 0$ then warning: model (4.51) is not supported by the data and drop this estimate.
  Apply (4.84), (4.64) or (4.65) to obtain $\hat{\Sigma}_{\hat{\beta}b}^{-1}$.
  $\chi_b^2 \leftarrow (\hat{\beta}_{b} - \beta^*\hat{\Sigma}_{\hat{\beta}b}^{-1}(\hat{\beta}_{b} - \beta^*),$
  where $\beta^* = (\beta^*_{0,mm}, \beta^*_{1,mm})'$ is defined using (4.83).
end for

Consider the estimated test statistic from the observed data $\hat{\chi}^2 = (\hat{\beta} - (0,1)')^t\hat{\Sigma}_{\hat{\beta}}^{-1}(\hat{\beta} - (0,1))$. A bootstrap estimate of the $p-value$ is

$$ p-value^* = 1 + \frac{\sum_{b=1}^B I(\chi_b^2 \geq \hat{\chi}^2)}{B + 1}, \quad (4.86) $$

where $I$ is the indicator function. The +1 in the numerator and denominator of the $p-value$ makes the estimator more stable when $B$ is relatively small and also it keeps from reporting a $p-value$ of exactly zero.

4.2.1.8 Performance Evaluation

Under normality and assuming the true variances and coefficients are known the residuals of regressing the error-prone observations and the true coefficients can be used to build a set of scores as

$$ t_{score,a} = \frac{\bar{Y}_a - (\beta_0 + \beta_1\bar{x}_a)}{\sqrt{MSE_c}} \sim N(0, 1). \quad (4.87) $$
However, the true variance and coefficients are unknown and instead estimates are used, such that an approximate $\hat{t}$-score with approximate distribution becomes

$$\hat{t}_{score,a} = \frac{\bar{Y}_a - (\hat{\beta}_{0,mm} + \hat{\beta}_{1,mm} \hat{x}_a)}{\sqrt{s_{\bar{Y}_a}^2 + \hat{\beta}_{1,mm}^2 \hat{\sigma}_{\hat{x}_a}^2}} \sim T_{\nu_a}. \quad (4.88)$$

We can proceed as in the previous section. By using the upgraded $\hat{t}_{score,a}$ from (4.88), with critical values as described in Section 3.1.2.3 and adjusted by the adequate degrees of freedom

$$\nu_a \approx \frac{(s_{\bar{Y}_a}^2 + \hat{\beta}_{1,mm}^2 \hat{\sigma}_{\hat{x}_a}^2)^2}{\frac{s_{\bar{Y}_a}^4}{K_a - 1} + \frac{\beta_{1,mm}^4 V[\bar{r}_a]}{\nu_{\bar{r}_a}}}. \quad (4.89)$$

The use of the method of moment estimators leads to scores of this form.

The evaluation of the participant’s method variability can be done by comparing it against the within variability of the methods used during the certification as described in Section 3.1.2.3.

### 4.2.1.9 Residual Analysis

#### Linearity assumption

The analysis of the residuals is a qualitative way to assess the linearity assumption. Consider the residuals from the observed values and the corrected coefficients

$$\bar{r}_a = \bar{Y}_a - (\hat{\beta}_{0,mm} + \hat{\beta}_{1,mm} \hat{x}_a). \quad (4.90)$$

This appears to be an estimator for the error in the equation $\bar{r}_a$. However, even if the model is linear $E[\bar{r}_a|x_a] = \beta_1(x_a - x_a)$ since $\bar{r}_a$ and $\hat{x}_a$ are correlated. We use the idea of Fuller [22] to obtain a modified value $\hat{x}_a^*$ instead of $\hat{x}_a$ for which $E[\bar{r}_a|x_a^*] = 0$ when linearity holds. This modified predictor is defined as

$$\hat{x}_a^* = \hat{x}_a + \frac{\hat{\beta}_{1,mm} \hat{\sigma}_{\hat{x}_a}^2}{V[\bar{r}_a]} \bar{r}_a. \quad (4.91)$$
Using (4.91) requires an estimator for \( V[\bar{r}_a] \) to be readily available. Under replicated data this is not a problem, since we can obtain individual estimates for each compound as
\[
\hat{V}[\bar{r}_a] = \hat{\beta}^2_{1,mm}(\hat{\sigma}^2_{Ua} + \hat{\sigma}^2_{\epsilon a}) + \hat{\sigma}^2_{\epsilon a}.
\] (4.92)

However, with no replicated data, some constraints are required. The usual option is to consider the variance of the error in the equation to be constant and then some estimators already found in (4.69) and (4.72) can be used. There is some risk involved in using a constant variance for the error in the equation when in fact it is not constant.

The new estimator in (4.91) arises by considering the \( \beta \)'s as known and obtaining a generalized least squares estimator of \( x_a \) (Fuller (1987) [22], equation 1.2.27). This suggests to plot \( \bar{r}_a \) versus \( \hat{x}_a^* \) to assess the linearity assumption.

A non random distribution centered around zero of the residuals is indicative of violations of the linearity assumption.

**Constancy of the variance of the error in the equation**

We note that under replication the variance of the error in the equation can be estimated for each compound and we can obtain CIs without assuming the linear model for bias over compounds, as demonstrated in Section 4.1.2.1, and specifically, as described in (4.6), (4.20) and (4.21), requiring indexing by compound. In this sense we can plot the individual estimates with CIs to qualitatively assess constant variance over compounds.

Note also that under replication and model (4.51), the modified squared residual technique in a typical ME problem is not adequate (Buonaccorsi (2010) [9]). Using the variance of the regular residuals (with the fitted coefficients) will account for changes in the measurement error variance but also changes due to the heterogeneity of the material (\( \sigma^2_{Ua} \)). It is preferable to consider a new modified squared residual
\[ msr_a = \max(0, r_a^2 - \hat{\beta}_{1,mm}^2(\hat{\sigma}_{2a}^2 + \hat{\sigma}_{Ua}^2)), \]

\[ V[\bar{\epsilon}_a] = E[msr_a|\beta_0, \beta_1, \sigma_{Ua}^2, \sigma_{\bar{x}a}^2]. \] (4.93)

This estimator converges to the typical modified squared residual as \( \hat{\sigma}_{Ua}^2 \) tends to zero.

A trend in the plot of \( \hat{\sigma}^2_{\bar{x}a} \), \( msr_a \) or \( \sqrt{msr_a} \) versus \( \hat{x}_a^* \) is suggestive of changing variance for the error in the equation (Buonaccorsi (2010) [9]).

Note that \( msr_a \) can be used to build a consistent estimator for \( \sigma_{\bar{x}a}^2 \) that does not require replicates and neither does it require assuming constant variance of the error in the equation. We consider this as a fourth alternative for estimating the variance of the error in the equation

\[ \hat{\sigma}^2_{\bar{x}a}(msr) = K_a \times \max(0, r_a^2 - \hat{\beta}_{1,mm}^2(\hat{\sigma}_{2a}^2 + \hat{\sigma}_{Ua}^2)). \] (4.94)

### 4.2.1.10 Fieller’s confidence intervals

Fieller’s method (Fieller (1950) [21]) is based on finding a pivotal quantity to obtain a confidence set for a ratio. Buonaccorsi (2001) [7] summarizes both the theorem and its applications including inverse prediction, estimation of the point of intersection of two linear regressions, estimation of relative potency of a new treatment to that of a standard treatment, estimation of extremum in a quadratic regression, all expressed as a ratio of random variables.

Lyles and Kupper (1999) [36] proposed an adjustment to the coefficient of a regression model due to measurement error by means of a variance stabilizing transformation. They noted that this is a ratio of random variables but they did not pursue Fieller’s method.

The idea of this technique is to use the well-known theory for ratio estimation and to take advantage of the relationship between the naive estimators and the moment
corrected estimators. This relationship can be clarified using (4.57) and (4.59), in general
\[ \gamma_1 = E[\hat{\beta}_{1,\text{naive}}] \approx \frac{\sigma_{X,Y}}{\sigma_X^2 + \sigma_U^2 + \sigma_\epsilon^2} = \beta_1 \kappa, \]  
(4.95)
where \( \kappa = \frac{\sigma_X^2}{\sigma_X^2 + \sigma_U^2 + \sigma_\epsilon^2} \) is the reliability coefficient. Then
\[ \beta_1 = \frac{\gamma_1}{\kappa}, \]
(4.96)
We also know that
\[ \hat{\beta}_{1,\text{mm}} = \frac{\hat{\beta}_{1,\text{naive}}}{\hat{\kappa}} = \frac{\hat{\gamma}_1}{\kappa}, \]
(4.97)
where \( \hat{\kappa} = \frac{S_X^2 - \hat{\sigma}_Y^2 - \hat{\sigma}_\epsilon^2}{S_X^2} \). Assuming \( \hat{\beta}_{1,\text{mm}} \) is approximately normal and \( \hat{V}[\hat{\beta}_{1,\text{mm}}] \) is distributed proportional to a chi-square(\( \nu \)) distribution.

\[
(1 - \alpha) \approx Pr \left( \frac{|\hat{\beta}_{1,\text{mm}} - \beta_1|}{\sqrt{\hat{V}[\hat{\beta}_{1,\text{mm}}]}} \leq T_{(1-\alpha/2;\nu)} \right) \\
\approx Pr \left( \frac{|S_{X,Y} - \beta_1 \left(S_X^2 - \hat{\sigma}_\epsilon^2\right)|}{\sqrt{\hat{V} \left[S_{X,Y} - \beta_1 \left(S_X^2 - \hat{\sigma}_\epsilon^2\right)\right]}} \leq T_{(1-\alpha/2;\nu)} \right) \\
= Pr \left( q(\beta_1) \leq 0 \right),
\]
(4.98)
where \( q(\beta_1) \) is a quadratic function on \( \beta_1 \), and \( \nu \) can be estimated by \( A - 1 \), obtained by the Welch-Satterthwaite approximation or just assume it is large enough to use the normal distribution. Assuming \( \hat{\sigma}_\epsilon^2 \) is uncorrelated with \((\hat{x}, \hat{Y})\) data
\[
V \left[S_{X,Y} - \beta_1 \left(S_X^2 - \hat{\sigma}_\epsilon^2\right)\right] = V[S_{X,Y}] + \beta_1^2 \left(V[S_X^2] + V[\hat{\sigma}_\epsilon^2]\right) \\
- 2\beta_1 Cov[S_{X,Y}, S_X^2],
\]
(4.99)
Using estimates of the variance and covariance terms and leaving \( \beta_1 \) unestimated then (4.99) is quadratic in \( \beta_1 \). Using \( \nu \) approximate degrees of freedom, \( q(\beta_1) \) in (4.98) becomes
\[ q(\beta_1) = f_2 \beta_1^2 - 2f_1 \beta_1 + f_0 \]
\[ = \left( \left( S_2 X - \tilde{\sigma}^2 X \right)^2 - t_{(1-\alpha/2, \nu)}^2 \hat{V}[S_2 X] + \hat{V}[\tilde{\sigma}^2 X] \right) \beta_1^2 \]
\[ - 2 \left( S_{XY} \left( S_2 X - \tilde{\sigma}^2 X \right) - t_{(1-\alpha/2, \nu)}^2 \hat{Cov}[S_{XY}, S_2 X] \right) \beta_1 \]
\[ + \left( S_2^2 X - t_{(1-\alpha/2, \nu)} \hat{V}[S_{XY}] \right). \]

Hence the Fieller’s method leads to an approximate confidence set for \( \beta_1 \) of the form \( \{ \beta_1 : q(\beta_1) \leq 0 \} \). Defining \( D = f_2^2 - f_0 f_2 \), \( r_1 = (f_1 - D^{1/2})/f_2 \), and \( r_2 = (f_1 + D^{1/2})/f_2 \), the approximate \((1 - \alpha)100\%\) confidence set for \( \beta_1 \) can be expressed as

\[
\begin{cases}
[r_1, r_2] & \text{if } f_2 > 0, \\
(\infty, r_2] \cup [r_1, \infty) & \text{if } f_2 \leq 0 \text{ and } D > 0, \\
(\infty, \infty) & \text{if } f_2 \leq 0 \text{ and } D \leq 0.
\end{cases}
\]  \hspace{1cm} (4.101)

Under model (4.51) the exact expressions for \( V[S_2^2 X], V[S_{XY}], Cov[S_2^2 X, S_{XY}] \) can be found by extending the theory of quadratic forms. The variance of the sample variance under normality or assuming common third and fourth moments is well known (Seber and Lee (2003) [48] Theorem 1.6 and exercise 17, Ravishanker and Dey (2002)[43] exercise 5.23). We give an extension to this result without assuming normality below:

**Theorem 1** Let \( X_i \) be \( n \) independent random variables with \( E[X_i] = \mu_i \), \( V[X_i] = \mu_{2,i} \), \( E[(X_i - \mu_i)^3] = \mu_{3,i} \) and \( E[(X_i - \mu_i)^4] = \mu_{4,i} \), the second, third and fourth central moments, respectively, assuming they all are finite. Also define \( \Sigma = \text{diag}(\mu_{2,i}) \), \( M_3 = \text{diag}(\mu_{3,i}) \) and \( M_4 = \text{diag}(\mu_{4,i}) \). If \( A \) is any \( n \times n \) symmetric matrix and \( a \) is the column vector of the diagonal elements of \( A \), then

\[
V[X'AX] = a' (M_4 - 3\Sigma^2) a + 2tr((A\Sigma)^2) + 4\mu' A \Sigma A \mu + 4\mu' AM_3 a. \]  \hspace{1cm} (4.102)
Also we provide expressions for the other terms which involve quadratic forms and bilinear forms.

**Theorem 2** Let $X_i, Y_i$ be $n$ independent random variables with expected values $E[X] = \mu_X = (\mu_{X1}, \ldots, \mu_{Xn})$, $E[Y] = \mu_Y$, and $V[X] = \Sigma_{XX} = \text{diag}(\mu_{X,i})$ and $V[Y] = \Sigma_{YY} = \text{diag}(\mu_{Y,i})$, the second finite central moments. If $A$ is any $n \times n$ symmetric matrix, then

$$V[X'AY] = \mu_X' A \Sigma_{YY} A \mu_X + \mu_Y' A \Sigma_{XX} A \mu_Y + \text{tr}(A \Sigma_{XX} A \Sigma_{YY}).$$

(4.103)

**Theorem 3** Let $X_i, Y_i$ be $n$ independent random variables with $E[X_i] = \mu_{Xi}$, and $\Sigma_{XX}, M_3$ and $M_4$ defined as in Theorem 1. Similarly define $\mu_Y, \Sigma_{YY}, M_3, M_4$ using the respective first moment and the second, third and fourth central moments of $Y$. If $A, B$ are any $n \times n$ symmetric matrices and $b$ is the column vector of the diagonal elements of $B$ then

$$\text{Cov}[X'AY, Y'BY] = 2\mu_X' A \Sigma_{YY} B \mu_Y + \mu_X' A M_3 b.$$  

(4.104)

The proofs can be found in Appendix C.

The sample variances and covariance of the control data can be rewritten as

$$S_{XY} = \bar{X}' Q \bar{Y},$$

$$S_{X} = \bar{X}' Q \bar{X},$$

$$S_{Y} = \bar{Y}' Q \bar{Y},$$

(4.105)

where $Q = \frac{1}{A-I} \left(I_A - \frac{1}{A} J_A \right)$ is a symmetric matrix.

Applying Theorems 1, 2, 3, assuming normality and assuming $\hat{\sigma}_{2a}^2 \nu_a / \sigma_{2a}^2$ is distributed as a chi-square($\nu_a$) distribution we obtain the method of moments estimators
\[ \hat{V}[S_X^2] = 2tr((\hat{Q}\hat{\Sigma}_x^2\hat{Q})^2) + 4\hat{x}'\hat{Q}\hat{\Sigma}_x^2\hat{Q}\hat{x}, \]

\[ \hat{V}[S_{XY}] = \hat{X}'\hat{Q}\hat{\Sigma}_Y^2\hat{Q}\hat{X} + \hat{Y}'\hat{Q}\hat{\Sigma}_Y^2\hat{Q}\hat{Y} + tr(\hat{Q}\hat{\Sigma}_x^2\hat{Q}\hat{\Sigma}_Y^2), \]

\[ \hat{Cov}[S_{XY}, S_X^2] = 2\hat{Y}'\hat{Q}\hat{\Sigma}_Y^2\hat{Q}\hat{x}, \]

\[ \hat{V}[\hat{\sigma}_x^2] = \frac{2}{\hat{A}^2} \sum_a \frac{\hat{\sigma}_{xa}^4}{\nu_a + 2}, \]

where \( \hat{\Sigma}_x^2 = \text{diag}(\hat{\sigma}_{x1}^2, \ldots, \hat{\sigma}_{xA}^2), \hat{\Sigma}_Y^2 = \text{diag}(s_{Y1}^2, \ldots, s_{YA}^2). \)

If \( \hat{\sigma}_x^2 > 0 \) and \( \hat{\sigma}_e^2 \leq 0 \) then we can use the modified estimators in (4.77) or (4.78) to update the relationship (4.98) as

\[ (1 - \alpha) \approx Pr \left( \left| \hat{\beta}_{1,IC} - \beta_1 \right| \leq T_{(1 - \alpha/2, \nu)} \right) \]

\[ \approx Pr \left( \left| \frac{S_Y^2 - \beta_1 S_{XY}}{\hat{V}[S_Y^2 - \beta_1 S_{XY}]} \right| \leq T_{(1 - \alpha/2, \nu)} \right) \]

\[ = Pr(q(\beta_1) \leq 0), \]

and we proceed as described above.

4.2.1.11 Simulation Study

To compare the Fieller’s CIs in (4.101) versus the Wald’s CIs using the robust covariance estimator and the Bootstrap CIs, we simulate data under model (4.51) and (4.52) over a range of conditions and the different methods are applied to obtain the estimated CI sets, then the coverage of each option was computed. The conditions domain must mimic the parameter space in a realistic way and some caution must be exercised. In particular, we are interested on levels of \( x_a \) different from zero, so the variance \( \sigma_{xa}^2 \) must be bounded from above, for example, \( cv_{\hat{x}_a} = \sigma_{\hat{x}_a}/x_a < 0.3 \). Since small variance of the reference value and small variance of the error in the
equation should lead to a even smaller estimated variance due to heterogeneity, a simple way to model $\sigma^2_U$ is as a fraction of the variance of the measurement error $\sigma^2_{\hat{x}}$. The predictor was assigned three sets of fixed values. The degrees of freedom of the estimated measurement error variance and the estimated heterogeneity variance were fixed at 30. The values assigned to the parameters are summarized in Table 4.7. These values are based in part on the example that follows later in Section 4.2.1.13. The simulation is described in Algorithm 4.

<table>
<thead>
<tr>
<th>parameter</th>
<th>values</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{x}$</td>
<td>(25.5, 38.4, 53.0, 53.3), (25.5, 38.4, 53.0, 53.3, 74.0, 73.2, 75.9, 85.0, 108.0, 127.0) and (25.5, 38.4, 53.0, 53.3, 74.0, 73.2, 75.9, 85.0, 108.0, 127.0, 184.0, 217.0, 225.0, 276.0, 291.0, 325.0, 335.0, 307.0, 341.0, 358.0, 397.0, 406.0, 453.0, 581.0, 651.0)</td>
</tr>
<tr>
<td>$\beta_0$</td>
<td>20, 900</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>0.9, 0.4</td>
</tr>
<tr>
<td>$K$</td>
<td>3, 5</td>
</tr>
<tr>
<td>$cv_{\hat{x}}$</td>
<td>0.01, 0.10, 0.20</td>
</tr>
<tr>
<td>$\sigma_\epsilon$</td>
<td>5%, 10% of $x$ and the constant values 20 and 90</td>
</tr>
<tr>
<td>$\sigma_U/\sigma_{\hat{x}}$</td>
<td>0.0, 0.2, 0.5</td>
</tr>
</tbody>
</table>

Table 4.7. Summary of parameter values used during the simulation study.

The computation of the observed value $Y_a$ is the linear response $\beta_0 + \beta_1 \hat{x}_a$ plus some error $\beta_1 U_a + \epsilon_a$ with the constraint of being a non-negative value. Hence if a negative value is obtained, it is set to zero.

Confidence Interval Performance

The results for confidence intervals for $\beta_1$ are plotted in Figures 4.1, 4.2, 4.3, 4.4, 4.5 and 4.6. For each method (Wald, Bootstrap and Fieller) we present two panel charts. The first ones using the first three CIs obtained in algorithm 4, based on $\hat{\sigma}^2_{\epsilon,mm}$ and assuming it is constant. The second ones using the last three CIs obtained in algorithm 4, based on $\hat{\sigma}^2_{\epsilon,a}$ and assuming it can change over compounds and it is based on replicated measurements. Each panel chart is organized horizontally by increasing
Algorithm 4 Simulation study

for each combination of parameters and $x$ values do
  for $s = 1$ to $S(large)$ do
    Resample from the target sample parametrically using the current combination of parameters to build the $s^{th}$ resampled data $(\hat{x}_s, \bar{Y}_s, \hat{\sigma}^2_{x_s}, \hat{\sigma}^2_{U_s}, \hat{\sigma}^2_{\epsilon_s})$, with
    \[
    \hat{x}_s \leftarrow \hat{x} + \hat{\sigma}_x Z_{1s},
    \hat{\sigma}^2_{x_s} \leftarrow \sigma^2_x \chi^2(\nu_{\hat{\sigma}_x})/\nu_{\hat{\sigma}_x},
    \hat{\sigma}^2_{U_s} \leftarrow \sigma^2_U \chi^2(\nu_{\hat{\sigma}_U})/\nu_{\hat{\sigma}_U},
    \hat{\sigma}^2_{\epsilon_s} \leftarrow \sigma^2_\epsilon \chi^2(\nu_{\hat{\sigma}_\epsilon})/\nu_{\hat{\sigma}_\epsilon},
    s^2_{Y_s} \leftarrow \beta_0^2 + \beta_1 \hat{x}_s + Z_{2s}^2 \sqrt{s^2_{Y_s}/K},
    \]
    where $(Z_{1s}, Z_{2s})$ are vectors of $A$ independent standard normal distributed.
    Obtain the CIs using $\sigma^2_{\epsilon,mm}$ from (4.69).
    Compute the Wald’s CI for $\beta_1$ using normal based covariance matrix.
    Compute the Bootstrap CI for $\beta_1$ using $B(large)$ cycles.
    Compute the Fieller’s CI for $\beta_1$.
    Obtain the CIs using $\sigma^2_{\epsilon a}$ from the replicated data by applying (4.75).
    Compute the Wald’s CI under replication for $\beta_1$ using robust covariance matrix, which allows for changing $\sigma^2_{\epsilon a}$.
    Compute the Bootstrap CI under replication for $\beta_1$ using $B(large)$ cycles.
    Compute the Fieller’s CI under replication for $\beta_1$.
  end for
  Record the current combination of parameters and the computed CIs.
end for

Record the current combination of parameters and the proportion of times the computed CIs contained $\beta_1$. 
end for
the variance of the error in the equation \( \sigma^2 \) and vertically by increasing the sample size \( A \). The first and second columns show estimates under non constant variance of the error in the equation and they are expressed in terms of the coefficient of variation of the error in the equation. The first column for \( cve = cv_\epsilon = 0.05 \), the second column for \( cve = cv_\epsilon = 0.10 \). The third and fourth columns show estimates under constant variance of the error in the equation and they are expressed in terms of the standard deviation of the error in the equation. The third column for \( sde = \sigma_\epsilon = 20 \) and the fourth column for \( sde = \sigma_\epsilon = 90 \). The first row for \( A = 25 \), the second row for \( A = 10 \) and the third row for \( A = 4 \). Within each panel we plot the estimated coverage versus the coefficient of variation of the measurement error while varying \( \sigma^2_U \) = the variance due to heterogeneity of the material and \( K \) = the number of replicates. The data points with \( \sigma_U/\sigma_\hat{x} = 0 \) are coded as dots, \( \sigma_U/\sigma_\hat{x} = 0.2 \) as triangles and \( \sigma_U/\sigma_\hat{x} = 0.5 \) as squares; the data points with number of replicates \( K = 3 \) are coded as empty symbols and \( K = 5 \) as solid symbols.

The variance \( \sigma^2_\epsilon \) was estimated by using (4.69) when assumed constant and (4.75) when assumed to be changing.

We observe that the number of replicates, the true value of the coefficients and the variance due to heterogeneity have little impact on the performance of the CIs, at least for the predefined values (see Table 4.7). Larger number of replicates and larger true value for the slope tend to improve the performance minimally. The low impact of the variance due to heterogeneity of the material can be explained since the variance due to heterogeneity is modeled as a fraction of the variance of the measurement error. However, larger variance due to heterogeneity may have important effects in the performance of the CIs.

The sample size appears to be more important in order for the CIs to have a good performance, a small sample size \( A = 4 \) leads to the lowest performance of the bootstrap and Fieller’s methods. Surprisingly, the Wald’s CI performance appears to
Figure 4.1. Estimated coverage of the Wald’s CI for $\beta_1$.

Each cell shows the estimated coverage of the CI vs the coefficient of variation of the measurement error. The panel is organized by $A =$ the sample size and $cve = \sigma_{\epsilon,a}/x_a =$ the coefficient of variation of the error in the equation when its variance is assumed to be changing and $sde = \sigma_\epsilon =$ the standard deviation of the error in the equation when its variance is assumed constant. The data points with $cv_U = \sigma_U/\sigma_\hat{x} = 0$ are coded as dots, $cv_U = 0.2$ as triangles and $cv_U = 0.5$ as squares; the data points with number of replicates $K = 3$ are coded as empty symbols and $K = 5$ as solid symbols.

be better than that of the bootstrap and Fieller’s CI when assuming the variance of the error in the equation is constant. However, the Wald’s CI better performance is
Figure 4.2. Estimated coverage of the Bootstrap CI for $\beta_1$.

Each cell shows the estimated coverage of the CI vs the coefficient of variation of the measurement error. The panel is organized by $A$ = the sample size and $cve = \sigma_{\epsilon,a}/x_a$ = the coefficient of variation of the error in the equation when its variance is assumed to be changing and $sde = \sigma_\epsilon$ = the standard deviation of the error in the equation when its variance is assumed constant. The data points with $cv_U = \sigma_U/\hat{\sigma}_x = 0$ are coded as dots, $cv_U = 0.2$ as triangles and $cv_U = 0.5$ as squares; the data points with number of replicates $K = 3$ are coded as empty symbols and $K = 5$ as solid symbols.

only apparent. Table 4.9 shows the mean length of the CI assuming constant variance for the error in the equation (right section) and constant coefficient of variation of
Figure 4.3. Estimated coverage of the Fieller’s CI for $\beta_1$.

Each cell shows the estimated coverage of the CI vs the coefficient of variation of the measurement error. The panel is organized by $A = \text{the sample size}$ and $cve = \sigma_{e,a}/x_a = \text{the coefficient of variation of the error in the equation when its variance is assumed to be changing}$ and $sde = \sigma_e = \text{the standard deviation of the error in the equation when its variance is assumed constant}$. The data points with $cv_U = \sigma_U/\sigma_\hat{x} = 0$ are coded as dots, $cv_U = 0.2$ as triangles and $cv_U = 0.5$ as squares; the data points with number of replicates $K = 3$ are coded as empty symbols and $K = 5$ as solid symbols.

the error in the equation (left section). We can see that for small sample sizes all the methods provide meaningless CIs. As the sample size decreases, the Wald’s and
Figure 4.4. Estimated coverage of the Wald’s CI for $\beta_1$ using replicates.

Each cell shows the estimated coverage of the CI vs the coefficient of variation of the measurement error. The panel is organized by $A =$ the sample size and $cve = \sigma_{e,a}/x_a =$ the coefficient of variation of the error in the equation when its variance is assumed to be changing and $sde = \sigma_e =$ the standard deviation of the error in the equation when its variance is assumed constant. The data points with $cv_U = \sigma_U/\sigma_x = 0$ are coded as dots, $cv_U = 0.2$ as triangles and $cv_U = 0.5$ as squares; the data points with number of replicates $K = 3$ are coded as empty symbols and $K = 5$ as solid symbols.

Bootstrap CIs become too conservative in such a way that the estimated slope can be statistically equal to zero, while the Fieller’s method tends to produce non-finite
Each cell shows the estimated coverage of the CI vs the coefficient of variation of the measurement error. The panel is organized by $A =$ the sample size and $cve = \sigma_{e,a}/x_a =$ the coefficient of variation of the error in the equation when its variance is assumed to be changing and $sde = \sigma_e =$ the standard deviation of the error in the equation when its variance is assumed constant. The data points with $cv_U = \sigma_U/\sigma_\tilde{x} = 0$ are coded as dots, $cv_U = 0.2$ as triangles and $cv_U = 0.5$ as squares; the data points with number of replicates $K = 3$ are coded as empty symbols and $K = 5$ as solid symbols.

CIs. We can also observe that the methods perform similarly when the variance of the error in the equation is constant and when the coefficient of variation is constant.
Figure 4.6. Estimated coverage of the Fieller’s CI for $\beta_1$ using replicates.

Each cell shows the estimated coverage of the CI vs the coefficient of variation of the measurement error. The panel is organized by $A =$ the sample size and $cve = \sigma_{e,a}/x_a =$ the coefficient of variation of the error in the equation when its variance is assumed to be changing and $sde = \sigma_e =$ the standard deviation of the error in the equation when its variance is assumed constant. The data points with $cv_U = \sigma_U/x_\hat{z} = 0$ are coded as dots, $cv_U = 0.2$ as triangles and $cv_U = 0.5$ as squares; the data points with number of replicates $K = 3$ are coded as empty symbols and $K = 5$ as solid symbols.

Fieller’s intervals outperforms the bootstrap intervals and the performance is similar to that of the Wald’s intervals when the sample size is not extremely small. The
bootstrap method is very sensitive to small sample sizes. This is due to the fact that a small sample, used to resample from, is hardly representative of the entire population.

Table 4.8 shows a summary of the coverage of the confidence sets obtained by the different methods. The top section shows the coverage when using only the means or the single observed values and assuming the variance of the error in the equation is constant. The bottom section shows the respective estimates under the assumption of changing variance of the error in the equation and using the sample mean and variance of the replicates. The estimated values are as follow: the rows labeled as Wald list the estimated $Pr(\beta_1 \in CI_{Wald})$ where $CI_{Wald}$ was obtained by using (4.63) and using the normal-based estimator for the covariance matrix in (4.65), the rows labeled as Bootstrap list the estimated $Pr(\beta_1 \in CI_{Bootstrap})$ where $CI_{Bootstrap}$ was obtained by using algorithm 1 and the rows labeled as Fieller list the estimated $Pr(\beta_1 \in CI_{Fieller})$ where $CI_{Fieller}$ was obtained by using (4.101). The cells show the average and standard deviation of the estimates over the number of replicates ($K$), the variance due to heterogeneity of the material ($\sigma_U$) and the value of the true coefficients ($\beta_0, \beta_1$), all constrained to the values considered in Table 4.7.

The length of the CIs, say $length(CI_{Wald}), length(CI_{Bootstrap})$ and $length(CI_{Fieller})$, were obtained by the difference of the upper and lower confidence limits recorded from algorithm 4. Table 4.9 lists the mean and standard error of the length of the CIs obtained by the three methods. In the case of the Fieller’s method the means and standard errors were computed using the finite length CIs only. The proportion of finite and infinite length CIs complements this information.

Table 4.10 lists the estimated probability of getting a finite length Feiller’s CI and the estimated probability of getting the whole real line as a the Feiller’s CI while varying the sample size and the variance of the measurement error. The length of the Fieller’s CIs are denoted by $\lambda$. The Feiller’s method produces finite length CIs with about the same probability either using single observations data or replicated data.
<table>
<thead>
<tr>
<th>( A )</th>
<th>method</th>
<th>constant ( cv_\epsilon )</th>
<th>( cv_\hat{x} )</th>
<th>( \hat{x} )</th>
<th>constant ( \sigma_\epsilon )</th>
<th>( cv_\hat{x} )</th>
<th>( \hat{x} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>Wald</td>
<td>0.836(0.015)</td>
<td>0.948(0.023)</td>
<td>0.970(0.006)</td>
<td>0.946(0.006)</td>
<td>0.962(0.008)</td>
<td>0.972(0.006)</td>
</tr>
<tr>
<td></td>
<td>Bootstrap</td>
<td>0.812(0.015)</td>
<td>0.929(0.029)</td>
<td>0.949(0.012)</td>
<td>0.936(0.006)</td>
<td>0.942(0.010)</td>
<td>0.943(0.017)</td>
</tr>
<tr>
<td></td>
<td>Fieller</td>
<td>0.829(0.015)</td>
<td>0.938(0.018)</td>
<td>0.970(0.010)</td>
<td>0.942(0.005)</td>
<td>0.955(0.006)</td>
<td>0.970(0.010)</td>
</tr>
<tr>
<td>10</td>
<td>Wald</td>
<td>0.923(0.015)</td>
<td>0.978(0.012)</td>
<td>0.983(0.006)</td>
<td>0.949(0.009)</td>
<td>0.961(0.009)</td>
<td>0.974(0.006)</td>
</tr>
<tr>
<td></td>
<td>Bootstrap</td>
<td>0.842(0.017)</td>
<td>0.923(0.032)</td>
<td>0.946(0.014)</td>
<td>0.903(0.009)</td>
<td>0.890(0.013)</td>
<td>0.857(0.043)</td>
</tr>
<tr>
<td></td>
<td>Fieller</td>
<td>0.907(0.016)</td>
<td>0.957(0.017)</td>
<td>0.994(0.004)</td>
<td>0.941(0.007)</td>
<td>0.949(0.007)</td>
<td>0.949(0.027)</td>
</tr>
<tr>
<td>4</td>
<td>Wald</td>
<td>0.975(0.011)</td>
<td>0.998(0.003)</td>
<td>0.996(0.002)</td>
<td>0.952(0.004)</td>
<td>0.977(0.004)</td>
<td>0.991(0.005)</td>
</tr>
<tr>
<td></td>
<td>Bootstrap</td>
<td>0.795(0.007)</td>
<td>0.927(0.033)</td>
<td>0.954(0.007)</td>
<td>0.748(0.010)</td>
<td>0.710(0.029)</td>
<td>0.734(0.105)</td>
</tr>
<tr>
<td></td>
<td>Fieller</td>
<td>0.954(0.012)</td>
<td>0.997(0.006)</td>
<td>1.000(0.000)</td>
<td>0.933(0.004)</td>
<td>0.684(0.104)</td>
<td>0.889(0.082)</td>
</tr>
<tr>
<td>25</td>
<td>Wald(rep)</td>
<td>0.863(0.007)</td>
<td>0.872(0.007)</td>
<td>0.904(0.008)</td>
<td>0.913(0.008)</td>
<td>0.900(0.012)</td>
<td>0.906(0.009)</td>
</tr>
<tr>
<td></td>
<td>Bootstrap(rep)</td>
<td>0.907(0.019)</td>
<td>0.949(0.007)</td>
<td>0.947(0.006)</td>
<td>0.937(0.004)</td>
<td>0.945(0.005)</td>
<td>0.941(0.008)</td>
</tr>
<tr>
<td></td>
<td>Fieller(rep)</td>
<td>0.918(0.014)</td>
<td>0.953(0.004)</td>
<td>0.976(0.006)</td>
<td>0.949(0.007)</td>
<td>0.962(0.007)</td>
<td>0.977(0.006)</td>
</tr>
<tr>
<td>10</td>
<td>Wald(rep)</td>
<td>0.811(0.006)</td>
<td>0.850(0.021)</td>
<td>0.895(0.009)</td>
<td>0.875(0.004)</td>
<td>0.862(0.013)</td>
<td>0.893(0.022)</td>
</tr>
<tr>
<td></td>
<td>Bootstrap(rep)</td>
<td>0.904(0.020)</td>
<td>0.945(0.012)</td>
<td>0.946(0.004)</td>
<td>0.929(0.025)</td>
<td>0.927(0.023)</td>
<td>0.910(0.025)</td>
</tr>
<tr>
<td></td>
<td>Fieller(rep)</td>
<td>0.942(0.015)</td>
<td>0.975(0.006)</td>
<td>0.996(0.001)</td>
<td>0.960(0.016)</td>
<td>0.975(0.010)</td>
<td>0.988(0.010)</td>
</tr>
<tr>
<td>4</td>
<td>Wald(rep)</td>
<td>0.900(0.008)</td>
<td>0.966(0.010)</td>
<td>0.970(0.005)</td>
<td>0.883(0.007)</td>
<td>0.878(0.016)</td>
<td>0.944(0.010)</td>
</tr>
<tr>
<td></td>
<td>Bootstrap(rep)</td>
<td>0.916(0.010)</td>
<td>0.947(0.009)</td>
<td>0.940(0.014)</td>
<td>0.928(0.013)</td>
<td>0.911(0.016)</td>
<td>0.900(0.029)</td>
</tr>
<tr>
<td></td>
<td>Fieller(rep)</td>
<td>1.000(0.000)</td>
<td>1.000(0.000)</td>
<td>1.000(0.000)</td>
<td>0.998(0.001)</td>
<td>1.000(0.000)</td>
<td>1.000(0.000)</td>
</tr>
</tbody>
</table>

Table 4.8. Summary for the mean and standard deviation of the estimated coverage of the different CIs for \( \beta_1 \).

Top block using only the means or the single observed values, bottom block using the sample mean and variance of the replicates. Both blocks varying \( K \) = the number of replicates, (\( \beta_0, \beta_1 \)) = the true values of the coefficients and \( \sigma_U^2 \) = variance due to heterogeneity of the material.
<table>
<thead>
<tr>
<th></th>
<th>method</th>
<th>(A)</th>
<th>constant (cv_\epsilon)</th>
<th>(cv_\hat{x})</th>
<th>constant (\sigma_\epsilon)</th>
<th>(cv_\hat{x})</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.01</td>
<td>0.10</td>
<td>0.20</td>
<td>0.01</td>
</tr>
<tr>
<td>25</td>
<td>Wald</td>
<td></td>
<td>0.055(0.019)</td>
<td>0.155(0.052)</td>
<td>0.331(0.124)</td>
<td>0.135(0.087)</td>
</tr>
<tr>
<td></td>
<td>Bootstrap</td>
<td></td>
<td>0.052(0.017)</td>
<td>0.141(0.048)</td>
<td>0.268(0.099)</td>
<td>0.126(0.081)</td>
</tr>
<tr>
<td></td>
<td>Fieller</td>
<td></td>
<td>0.054(0.018)</td>
<td>0.168(0.057)</td>
<td>0.495(0.189)</td>
<td>0.132(0.085)</td>
</tr>
<tr>
<td>10</td>
<td>Wald</td>
<td></td>
<td>0.144(0.050)</td>
<td>0.385(0.128)</td>
<td>1.050(0.393)</td>
<td>1.332(0.865)</td>
</tr>
<tr>
<td></td>
<td>Bootstrap</td>
<td></td>
<td>0.116(0.040)</td>
<td>0.296(0.098)</td>
<td>0.597(0.219)</td>
<td>1.075(0.699)</td>
</tr>
<tr>
<td></td>
<td>Fieller</td>
<td></td>
<td>0.136(0.046)</td>
<td>0.540(0.189)</td>
<td>19.65(12.74)</td>
<td>1.258(0.817)</td>
</tr>
<tr>
<td>4</td>
<td>Wald</td>
<td></td>
<td>0.579(0.200)</td>
<td>1.661(0.545)</td>
<td>439.9(557.3)</td>
<td>8.722(5.595)</td>
</tr>
<tr>
<td></td>
<td>Bootstrap</td>
<td></td>
<td>0.216(0.075)</td>
<td>0.656(0.220)</td>
<td>10.35(4.242)</td>
<td>3.267(2.097)</td>
</tr>
<tr>
<td></td>
<td>Fieller</td>
<td></td>
<td>0.478(0.164)</td>
<td>6.006(4.990)</td>
<td>(\infty)</td>
<td>7.211(4.626)</td>
</tr>
<tr>
<td>25</td>
<td>Wald(rep)</td>
<td></td>
<td>0.064(0.022)</td>
<td>0.129(0.038)</td>
<td>0.254(0.091)</td>
<td>0.122(0.079)</td>
</tr>
<tr>
<td></td>
<td>Bootstrap(rep)</td>
<td></td>
<td>0.075(0.026)</td>
<td>0.154(0.045)</td>
<td>0.270(0.093)</td>
<td>0.130(0.084)</td>
</tr>
<tr>
<td></td>
<td>Fieller(rep)</td>
<td></td>
<td>0.078(0.027)</td>
<td>0.181(0.055)</td>
<td>0.499(0.183)</td>
<td>0.136(0.088)</td>
</tr>
<tr>
<td>10</td>
<td>Wald(rep)</td>
<td></td>
<td>0.120(0.042)</td>
<td>0.265(0.080)</td>
<td>0.720(0.266)</td>
<td>1.044(0.677)</td>
</tr>
<tr>
<td></td>
<td>Bootstrap(rep)</td>
<td></td>
<td>0.148(0.051)</td>
<td>0.315(0.093)</td>
<td>0.587(0.198)</td>
<td>1.181(0.778)</td>
</tr>
<tr>
<td></td>
<td>Fieller(rep)</td>
<td></td>
<td>0.173(0.060)</td>
<td>0.560(0.180)</td>
<td>19.63(12.73)</td>
<td>1.383(0.911)</td>
</tr>
<tr>
<td>4</td>
<td>Wald(rep)</td>
<td></td>
<td>0.355(0.122)</td>
<td>0.940(0.301)</td>
<td>129.3(136.3)</td>
<td>5.237(3.358)</td>
</tr>
<tr>
<td></td>
<td>Bootstrap(rep)</td>
<td></td>
<td>0.261(0.091)</td>
<td>0.662(0.199)</td>
<td>9.758(4.287)</td>
<td>4.642(3.062)</td>
</tr>
<tr>
<td></td>
<td>Fieller(rep)</td>
<td></td>
<td>0.577(0.200)</td>
<td>5.810(5.176)</td>
<td>(\infty)</td>
<td>10.27(6.771)</td>
</tr>
</tbody>
</table>

**Table 4.9.** Summary for the mean and standard deviation of the estimated CI length, while varying \(K\) = the number of replicates and \(\sigma^2_U\) = variance due to heterogeneity of the material. The Fieller’s CI estimates are based on the finite length CIs only when possible.
However, when using the replicated data, the method tends to produce complements of finite length CIs rather than infinite length CIs. It is clear that under moderate or large variance of the measurement error, the Fieller’s method fails to produce finite length CI for $A = 4$. The issue when $A = 4$ is not the small sample size, rather it is the variance of the measurement error being large relative to the dispersion of the true values. From (4.100) and (4.101) we can obtain a Fieller’s CI with finite length if and only if $\left| S^2_\hat{z} - \hat{\sigma}^2_\hat{z} \right| > t_{\nu,1-\alpha/2} \sqrt{\hat{V}[S^2_\hat{z}] + \hat{V}[\hat{\sigma}^2_\hat{z}]}$.

In order to observe the effects of assuming constant variance when the true variance of the error in the equation is not constant, consider the panels for sample size $A = 25$ from Figures 4.1, 4.2, 4.3, 4.4, 4.5 and 4.6. For $A = 25$ and the specific levels of the predictor, a coefficient of variation $cv_\epsilon = 0.10$ corresponds to a standard deviation $\sigma_\epsilon = 24$ on average. Hence comparing the panels for $(A = 25, cv_\epsilon = 0.10)$ and $(A = 25, \sigma_\epsilon = 20)$, we conclude that the departure from the constant variance assumption for the error in the equation has a large effect on the performance of the CIs of the three methods.

We can conclude that the Wald method is preferable if the sample size is large and no replicates are available. It produces CIs with similar coverage to the other methods and is the simplest method.

Fieller’s method outperforms the Wald’s and bootstrap methods when the sample size is not extremely small ($A = 10, A = 25$) and measurement error is present, specially when replicates are available. In this situation, both the Wald’s and bootstrap methods tend to produce narrower CIs but both likely fail on producing CI with the nominal coverage.

The coverage of the CIs obtained by three methods are more sensitive to some of the parameters. AIC criterion was used for model selection using the parameters are predictors for the coverage, and ANOVA was conducted for the more promising models in order to evaluate the contribution of the predictors. Under constant variance of
Left block under constant coefficient of variation of the error in the equation and right block under constant variance of the error in the equation. The third column on each block uses the sample mean and variance of the replicates. Both blocks varying $K =$ the number of replicates, $(\beta_0, \beta_1) =$ the true values of the coefficients and $\sigma^2_U =$ variance due to heterogeneity of the material.
the error in the equation, the CIs obtained by Wald’s method tend to perform better than those obtained by the bootstrap and Fieller’s method. In that case the coverage of the CIs is more sensitive to the sample size, the variance of the measurement error, the variance of the error in the equation and the number of replicates. Under changing variance of the error in the equation, the CIs obtained by the Fieller’s method tend to perform better than those obtained by the bootstrap and Wald’s method. In that case the coverage of the CIs is more sensitive to the sample size, the variance of the measurement error and the number of replicates.

Estimating the Laboratory Variance

In addition, this simulation gave information about the performance of the different estimators for the variance of the error in the equation. For this purpose, the variance of the error in the equation was modeled as described above (some proportional values and some constant values were considered). The different estimators were obtained: \( \hat{\sigma}^2_{\epsilon,mm} \) using (4.69), \( \hat{\sigma}^2_{\epsilon} \) using (4.72), \( \hat{\sigma}^2_{\epsilon,rep} \) using (4.76) and \( \hat{\sigma}^2_{\epsilon,msr} \) by taking the average over the separate estimates using (4.94). Then the mean and standard deviation of the estimates over the simulation were obtained and the ratio of the mean estimated values divided by the true value (\( \sigma^2_{\epsilon} \)) were obtained. The ideal estimator should lead to a ratio near to one and should have small variance. Table 4.11 lists the estimated variances from the simulation under the assumption of constant variance of the error in the equation.

Table 4.12 lists the estimates when the variance of the error in the equation changes with constant coefficient of variation relative to the predictor. The estimated coefficients of variation were computed as
The simulation results suggest that under constant variance of the error in the equation the estimator based on the modified squared residuals $\hat{\sigma}^2_{e(msr)}$ outperforms the other three estimators. When the variance is changing with constant coefficient of variation relative to the predictor level $x$, the estimates based on the modified squared residuals are always underestimating the true coefficient of variation, however it also performs better than the other three estimators.

4.2.1.12 Weighted estimators

If changing error in equation variance is determined, Fuller’s weighted estimators (Fuller (1987) [22], Buonaccorsi (2010) [9]) can be obtained by considering weighted versions of $S^2_X, S_{X\bar{Y}}$.

When the measurement error variances change considerably across compounds, weighting may be a real benefit. Fuller (1987) [22] and Buonaccorsi (2010) [9] approaches are based on the assumption of constant variance of the error in the equation. This assumption translates as constant variance of the pseudo-error in the equation in (4.54).

The weights can be defined in terms of the true pseudo error in equation variance as $\pi_a = (\sigma^2_{\eta a} + \beta^2_1 \sigma^2_{xa})^{-1}$, and these can be estimated by $\hat{\pi}_a = (\hat{\sigma}^2_{\eta a} + \hat{\beta}^2_1 \hat{\sigma}^2_{xa})^{-1}$, then the weighted version of $S^2_X, S_{X\bar{Y}}$ are
<table>
<thead>
<tr>
<th>$\sigma_e^2$</th>
<th>$A$</th>
<th>$cv_e$</th>
<th>$\hat{\sigma}_{e,mm}^2$</th>
<th>$\hat{\sigma}_{e,mm}/\sigma_e^2$</th>
<th>$\hat{\sigma}_e^2$</th>
<th>$\hat{\sigma}_e^2/\sigma_e^2$</th>
<th>$\hat{\sigma}_{e,rep}^2$</th>
<th>$\hat{\sigma}_{e,rep}/\sigma_e^2$</th>
<th>$\hat{\sigma}_{e,msr}^2$</th>
<th>$\hat{\sigma}_{e,msr}/\sigma_e^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>400</td>
<td>25</td>
<td>0.01</td>
<td>383.6(6.9)</td>
<td>0.959</td>
<td>401.9(6.6)</td>
<td>1.005</td>
<td>371.9(5.2)</td>
<td>0.930</td>
<td>399.6(2.9)</td>
<td>0.999</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>0.01</td>
<td>360.4(4.2)</td>
<td>0.901</td>
<td>405.7(4.7)</td>
<td>1.014</td>
<td>324.4(3.8)</td>
<td>0.811</td>
<td>398.4(4.0)</td>
<td>0.996</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.01</td>
<td>251.4(1.4)</td>
<td>0.629</td>
<td>377.4(2.1)</td>
<td>0.944</td>
<td>188.5(1.0)</td>
<td>0.471</td>
<td>397.4(2.3)</td>
<td>0.993</td>
</tr>
<tr>
<td>25</td>
<td>0.10</td>
<td>381.8(101.6)</td>
<td>0.954</td>
<td>480.5(129.3)</td>
<td>1.201</td>
<td>995.3(494.9)</td>
<td>2.488</td>
<td>404.3(104.4)</td>
<td>1.011</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>0.10</td>
<td>345.2(15.3)</td>
<td>0.863</td>
<td>409.5(4.6)</td>
<td>1.024</td>
<td>347.0(19.9)</td>
<td>0.867</td>
<td>398.2(4.0)</td>
<td>0.996</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.10</td>
<td>230.3(4.4)</td>
<td>0.576</td>
<td>371.7(6.8)</td>
<td>0.929</td>
<td>175.6(1.5)</td>
<td>0.439</td>
<td>396.7(2.4)</td>
<td>0.992</td>
</tr>
<tr>
<td>25</td>
<td>0.20</td>
<td>823.7(518.5)</td>
<td>2.059</td>
<td>1092(664)</td>
<td>2.730</td>
<td>3007(2012)</td>
<td>7.517</td>
<td>434.7(67.7)</td>
<td>1.087</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>0.20</td>
<td>319.4(21.6)</td>
<td>0.799</td>
<td>421.1(32.8)</td>
<td>1.053</td>
<td>465.3(117)</td>
<td>1.163</td>
<td>396.9(4.2)</td>
<td>0.992</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.20</td>
<td>360.7(163.5)</td>
<td>0.902</td>
<td>166186(441080)</td>
<td>415.5</td>
<td>286157(400951)</td>
<td>715.4</td>
<td>386.2(11.5)</td>
<td>0.965</td>
</tr>
<tr>
<td>8100</td>
<td>25</td>
<td>0.01</td>
<td>7560(346)</td>
<td>0.933</td>
<td>7890(361)</td>
<td>0.974</td>
<td>7258(332)</td>
<td>0.896</td>
<td>8091.0(59.5)</td>
<td>0.999</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>0.01</td>
<td>6733(707)</td>
<td>0.831</td>
<td>7575(796)</td>
<td>0.935</td>
<td>6059(636)</td>
<td>0.748</td>
<td>8067.7(80.6)</td>
<td>0.996</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.01</td>
<td>4493(695)</td>
<td>0.555</td>
<td>6741(1042)</td>
<td>0.832</td>
<td>3369(521)</td>
<td>0.416</td>
<td>8046.8(45.8)</td>
<td>0.993</td>
</tr>
<tr>
<td>25</td>
<td>0.10</td>
<td>7232(367)</td>
<td>0.893</td>
<td>7714(343)</td>
<td>0.952</td>
<td>7573(541)</td>
<td>0.935</td>
<td>8091(60)</td>
<td>0.999</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>0.10</td>
<td>6724(707)</td>
<td>0.830</td>
<td>7595(803)</td>
<td>0.938</td>
<td>6069(642)</td>
<td>0.749</td>
<td>8067(81)</td>
<td>0.996</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.10</td>
<td>4213(633)</td>
<td>0.520</td>
<td>6510(980)</td>
<td>0.804</td>
<td>3168(477)</td>
<td>0.391</td>
<td>8038(48)</td>
<td>0.993</td>
</tr>
<tr>
<td>25</td>
<td>0.20</td>
<td>6511(668)</td>
<td>0.804</td>
<td>7364(341)</td>
<td>0.909</td>
<td>9377(2010)</td>
<td>1.158</td>
<td>8086(59)</td>
<td>0.998</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>0.20</td>
<td>6471(671)</td>
<td>0.799</td>
<td>7441(787)</td>
<td>0.919</td>
<td>6070(673)</td>
<td>0.749</td>
<td>8057(83)</td>
<td>0.995</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.20</td>
<td>4360(991)</td>
<td>0.538</td>
<td>207520(431180)</td>
<td>25.62</td>
<td>297094(405385)</td>
<td>36.68</td>
<td>7914(150)</td>
<td>0.977</td>
</tr>
</tbody>
</table>

Table 4.11. Results for the estimated variance of the error in the equation.
<table>
<thead>
<tr>
<th>$cv_e$</th>
<th>$A$</th>
<th>$\sigma_z$</th>
<th>$\hat{cv}_{e,mm}/cv_e$</th>
<th>$\hat{cv}_{e,mm}/cv_e$</th>
<th>$\hat{cv}_e/cv_e$</th>
<th>$\hat{cv}_{e,rep}/cv_e$</th>
<th>$\hat{cv}_{e,rep}/cv_e$</th>
<th>$\hat{cv}_{e,msr}/cv_e$</th>
<th>$\hat{cv}_{e,msr}/cv_e$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>25</td>
<td>0.01</td>
<td>0.173(0.003)</td>
<td>3.467</td>
<td>0.178(0.003)</td>
<td>3.557</td>
<td>0.044(0.001)</td>
<td>0.888</td>
<td>0.046(0.002)</td>
</tr>
<tr>
<td>0.05</td>
<td>10</td>
<td>0.01</td>
<td>0.065(0.001)</td>
<td>1.305</td>
<td>0.070(0.001)</td>
<td>1.397</td>
<td>0.037(0.001)</td>
<td>0.740</td>
<td>0.046(0.002)</td>
</tr>
<tr>
<td>0.05</td>
<td>4</td>
<td>0.01</td>
<td>0.041(0.002)</td>
<td>0.811</td>
<td>0.052(0.001)</td>
<td>1.033</td>
<td>0.025(0.002)</td>
<td>0.503</td>
<td>0.046(0.002)</td>
</tr>
<tr>
<td>0.05</td>
<td>25</td>
<td>0.10</td>
<td>0.125(0.020)</td>
<td>2.494</td>
<td>0.153(0.018)</td>
<td>3.053</td>
<td>0.075(0.023)</td>
<td>1.498</td>
<td>0.044(0.003)</td>
</tr>
<tr>
<td>0.05</td>
<td>10</td>
<td>0.10</td>
<td>0.050(0.009)</td>
<td>1.010</td>
<td>0.069(0.013)</td>
<td>1.388</td>
<td>0.053(0.015)</td>
<td>1.061</td>
<td>0.044(0.003)</td>
</tr>
<tr>
<td>0.05</td>
<td>4</td>
<td>0.10</td>
<td>0.037(0.010)</td>
<td>0.742</td>
<td>0.070(0.020)</td>
<td>1.400</td>
<td>0.025(0.006)</td>
<td>0.507</td>
<td>0.043(0.003)</td>
</tr>
<tr>
<td>0.05</td>
<td>25</td>
<td>0.20</td>
<td>0.170(0.058)</td>
<td>3.391</td>
<td>0.217(0.068)</td>
<td>4.339</td>
<td>0.136(0.052)</td>
<td>2.715</td>
<td>0.043(0.003)</td>
</tr>
<tr>
<td>0.05</td>
<td>10</td>
<td>0.20</td>
<td>0.073(0.026)</td>
<td>1.451</td>
<td>0.105(0.036)</td>
<td>2.098</td>
<td>0.094(0.035)</td>
<td>1.882</td>
<td>0.043(0.003)</td>
</tr>
<tr>
<td>0.05</td>
<td>4</td>
<td>0.20</td>
<td>0.116(0.054)</td>
<td>2.321</td>
<td>0.679(0.524)</td>
<td>13.58</td>
<td>0.493(0.414)</td>
<td>9.858</td>
<td>0.042(0.004)</td>
</tr>
<tr>
<td>0.10</td>
<td>25</td>
<td>0.01</td>
<td>0.350(0.004)</td>
<td>3.503</td>
<td>0.358(0.004)</td>
<td>3.582</td>
<td>0.091(0.001)</td>
<td>0.907</td>
<td>0.091(0.003)</td>
</tr>
<tr>
<td>0.10</td>
<td>10</td>
<td>0.01</td>
<td>0.132(0.001)</td>
<td>1.319</td>
<td>0.140(0.001)</td>
<td>1.402</td>
<td>0.076(0.001)</td>
<td>0.762</td>
<td>0.091(0.003)</td>
</tr>
<tr>
<td>0.10</td>
<td>4</td>
<td>0.01</td>
<td>0.085(0.001)</td>
<td>0.850</td>
<td>0.105(0.001)</td>
<td>1.064</td>
<td>0.054(0.002)</td>
<td>0.540</td>
<td>0.091(0.003)</td>
</tr>
<tr>
<td>0.10</td>
<td>25</td>
<td>0.10</td>
<td>0.266(0.044)</td>
<td>2.660</td>
<td>0.299(0.028)</td>
<td>2.995</td>
<td>0.102(0.015)</td>
<td>1.019</td>
<td>0.090(0.004)</td>
</tr>
<tr>
<td>0.10</td>
<td>10</td>
<td>0.10</td>
<td>0.100(0.013)</td>
<td>0.999</td>
<td>0.124(0.006)</td>
<td>1.237</td>
<td>0.076(0.008)</td>
<td>0.765</td>
<td>0.090(0.004)</td>
</tr>
<tr>
<td>0.10</td>
<td>4</td>
<td>0.10</td>
<td>0.061(0.004)</td>
<td>0.611</td>
<td>0.102(0.009)</td>
<td>1.016</td>
<td>0.040(0.002)</td>
<td>0.401</td>
<td>0.089(0.004)</td>
</tr>
<tr>
<td>0.10</td>
<td>25</td>
<td>0.20</td>
<td>0.248(0.040)</td>
<td>2.476</td>
<td>0.303(0.037)</td>
<td>3.026</td>
<td>0.153(0.045)</td>
<td>1.531</td>
<td>0.088(0.006)</td>
</tr>
<tr>
<td>0.10</td>
<td>10</td>
<td>0.20</td>
<td>0.098(0.017)</td>
<td>0.981</td>
<td>0.136(0.025)</td>
<td>1.357</td>
<td>0.110(0.029)</td>
<td>1.095</td>
<td>0.087(0.006)</td>
</tr>
<tr>
<td>0.10</td>
<td>4</td>
<td>0.20</td>
<td>0.130(0.050)</td>
<td>1.295</td>
<td>0.695(0.514)</td>
<td>6.954</td>
<td>0.496(0.406)</td>
<td>4.963</td>
<td>0.083(0.007)</td>
</tr>
</tbody>
</table>

Table 4.12. Results for the ratio of the estimated standard deviation (CV) of the error in the equation over the true standard deviation (CV), for $A=10$, $K=3$, $\beta_0 = 20$, $\beta_1 = 0.9$. 
\[ S^2_{\hat{X}_\pi} = \frac{1}{A-1} \sum_a \hat{\pi}_a (\hat{x}_a - \bar{\hat{x}})^2 - \frac{1}{A} \sum_a \hat{\pi}_a \hat{\sigma}^2_{\hat{x}_a}, \]  
\[ S_{\hat{X}\hat{Y}_\pi} = \frac{1}{A-1} \sum_a \hat{\pi}_a (\hat{x}_a - \bar{\hat{x}})(\hat{Y}_a - \bar{\hat{Y}}). \]  
(4.109)

Finally the weighted estimates of the coefficients become:

\[ \hat{\beta}_{1w} = \frac{S_{\hat{X}\hat{Y}_\pi}}{S^2_{\hat{X}_\pi}}, \]
\[ \hat{\beta}_{0w} = \hat{Y}_w - \hat{\beta}_{1w} \bar{\hat{x}}_w, \]  
(4.110)

This procedure can be iterated. Often a few iterations are required to converge.

The covariance matrix must be amended accordingly. For the robust estimate (Buonaccorsi (2010) [9])

\[ \hat{\Sigma}_{\beta,R,\pi} = \hat{M}^{-1}_{XX\pi} \hat{H}_{R\pi} \hat{M}^{-1}_{XX\pi}, \]
\[ \hat{H}_{R\pi} = \frac{1}{A(A-2)} \sum_a \hat{\Delta}_{\pi a} \hat{\Delta}'_{\pi a}, \]
\[ \hat{\Delta}_{\pi a} = \hat{\pi}_a \left[ \begin{array}{c} \hat{Y}_a - \hat{\beta}_{0,mm} - \hat{\beta}_{1,mm} \hat{x}_a \\ \hat{x}_a (\hat{Y}_a - \hat{\beta}_{0,mm} - \hat{\beta}_{1,mm} \hat{x}_a) + \hat{\beta}_{1,mm} \hat{\sigma}^2_{\hat{x}_a} \end{array} \right], \]  
(4.111)
\[ \hat{M}_{XX\pi} = \left[ \begin{array}{cc} \hat{\pi} \hat{x} & \hat{\pi} \hat{x}^2 \\ \hat{\pi} \hat{x} & \hat{\pi} \hat{x}^2 - \hat{\pi} \hat{\sigma}^2_{\hat{z}} \end{array} \right] \]

and for the normal-based estimator

\[ \hat{\Sigma}_{\beta,N,\pi} = \hat{M}^{-1}_{XX\pi} \hat{H}_{N\pi} \hat{M}^{-1}_{XX\pi}, \]
\[ \hat{H}_{N\pi} = \frac{1}{A^2} \sum_a \hat{\pi}_a^2 (\hat{X}_{a*} \hat{X}'_{a*}/\hat{\pi}_a + \hat{Z}_{a} \hat{Z}'_{a}). \]  
(4.112)
Note that the test for constant variance of the pseudo error in the equation also needs to be modified, for weighted estimators we need to test if \( \sigma^2_{\eta a} \pi_a \) is constant for all \( a \). If it is constant then we can use \( \hat{\Sigma}_{\beta, N, \pi} \) otherwise we should use \( \hat{\Sigma}_{\beta, R, \pi} \).

### 4.2.1.13 Example

Table 4.13 contains a copy of the summary data showed in Table 3.1. For comparison purposes the summary data of a second participating laboratory from the same intercomparison of sediments in 2003 is included in Table 4.14. Both participants are using the same reference material and it is a homogeneous material so that \( \hat{\sigma}^2_{Ua} = 0, \forall a \).

Table 4.15 shows the estimated coefficients for a linear bias model for the data showed in Tables 4.13 and 4.14, applying (4.57) for naive point estimators, (4.60), (4.77), (4.78) and (4.80), if required, for corrected point estimators, respectively. The estimated standard error and the covariance were obtained by using (4.64), (4.65) for the robust and normal asymptotic standard errors and covariance of the coefficients. The mean square error was calculated with (4.70) using the corresponding estimated coefficients for each case. The estimated intercept does not differ from zero significantly in both cases, however the estimated slope does differ from identity for Laboratory 2. Also the corrected mean square error of Laboratory 2 is much larger than the one from Laboratory 1.

Figures 4.7 and 4.8 show the observed control data from laboratory 1 and 2 respectively, the fitted line and the confidence band for the fitted line based on the normal theory. The confidence intervals for each point were obtained by using the variance of the reference value and the variance of the replicates for the response. This suggests that laboratory 2 is biased while laboratory 1 is unbiased. Laboratory 2 shows a larger variability both within replicates and about the predicted value. Qualita-
<table>
<thead>
<tr>
<th>Compound</th>
<th>$\bar{Y}$</th>
<th>$s_Y$</th>
<th>K</th>
<th>$\bar{W}$</th>
<th>$s_W$</th>
<th>M</th>
<th>$\rho_{YW}$</th>
<th>$\hat{x}$</th>
<th>$\hat{\sigma}_x$</th>
<th>$\nu_x$</th>
<th>$\hat{\sigma}_w$</th>
<th>$\nu_w$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2,3,5-trimethylnaphthalene</td>
<td>22.7</td>
<td>1.05</td>
<td>3</td>
<td>25.7</td>
<td>0.56</td>
<td>3</td>
<td>0.69</td>
<td>25.5</td>
<td>2.60</td>
<td>29</td>
<td>3.18</td>
<td>9</td>
</tr>
<tr>
<td>acenaphthene</td>
<td>36.0</td>
<td>2.19</td>
<td>3</td>
<td>34.7</td>
<td>0.46</td>
<td>3</td>
<td>-0.89</td>
<td>38.4</td>
<td>2.65</td>
<td>30</td>
<td>3.23</td>
<td>13</td>
</tr>
<tr>
<td>dibenz[a,h]anthracene</td>
<td>56.3</td>
<td>0.69</td>
<td>3</td>
<td>112.0</td>
<td>4.16</td>
<td>3</td>
<td>0.28</td>
<td>53.0</td>
<td>5.10</td>
<td>10</td>
<td>3.20</td>
<td>4</td>
</tr>
<tr>
<td>acenaphthylene</td>
<td>56.8</td>
<td>2.07</td>
<td>3</td>
<td>29.6</td>
<td>1.36</td>
<td>3</td>
<td>-0.52</td>
<td>53.3</td>
<td>3.27</td>
<td>28</td>
<td>4.80</td>
<td>12</td>
</tr>
<tr>
<td>1-methylphenanthrene</td>
<td>71.8</td>
<td>0.66</td>
<td>3</td>
<td>69.4</td>
<td>0.62</td>
<td>3</td>
<td>0.92</td>
<td>73.2</td>
<td>3.01</td>
<td>26</td>
<td>5.38</td>
<td>12</td>
</tr>
<tr>
<td>biphenyl</td>
<td>70.0</td>
<td>2.03</td>
<td>3</td>
<td>29.2</td>
<td>0.71</td>
<td>3</td>
<td>-1.00</td>
<td>74.0</td>
<td>4.08</td>
<td>28</td>
<td>6.29</td>
<td>13</td>
</tr>
<tr>
<td>2,6-dimethylnaphthalene</td>
<td>74.0</td>
<td>1.98</td>
<td>3</td>
<td>35.2</td>
<td>0.46</td>
<td>3</td>
<td>0.52</td>
<td>75.9</td>
<td>2.30</td>
<td>23</td>
<td>3.91</td>
<td>14</td>
</tr>
<tr>
<td>fluorene</td>
<td>85.6</td>
<td>3.57</td>
<td>3</td>
<td>65.5</td>
<td>0.69</td>
<td>3</td>
<td>-0.95</td>
<td>85.0</td>
<td>7.65</td>
<td>39</td>
<td>7.24</td>
<td>9</td>
</tr>
<tr>
<td>triphenylene</td>
<td>111.0</td>
<td>2.65</td>
<td>3</td>
<td>141.0</td>
<td>8.50</td>
<td>3</td>
<td>0.49</td>
<td>108.0</td>
<td>2.55</td>
<td>11</td>
<td>4.46</td>
<td>3</td>
</tr>
<tr>
<td>1-methylnaphthalene</td>
<td>120.7</td>
<td>6.51</td>
<td>3</td>
<td>49.0</td>
<td>0.91</td>
<td>3</td>
<td>-0.43</td>
<td>127.0</td>
<td>7.14</td>
<td>30</td>
<td>13.89</td>
<td>12</td>
</tr>
<tr>
<td>anthracene</td>
<td>181.7</td>
<td>6.03</td>
<td>3</td>
<td>172.3</td>
<td>2.52</td>
<td>3</td>
<td>1.00</td>
<td>184.0</td>
<td>9.18</td>
<td>39</td>
<td>10.74</td>
<td>9</td>
</tr>
<tr>
<td>benzo[j]fluoranthene</td>
<td>220.7</td>
<td>1.53</td>
<td>3</td>
<td>281.0</td>
<td>8.50</td>
<td>3</td>
<td>-0.01</td>
<td>217.0</td>
<td>2.55</td>
<td>5</td>
<td>5.00</td>
<td>4</td>
</tr>
<tr>
<td>benzo[k]fluoranthene</td>
<td>226.3</td>
<td>5.03</td>
<td>3</td>
<td>406.0</td>
<td>8.72</td>
<td>3</td>
<td>-0.87</td>
<td>225.0</td>
<td>9.18</td>
<td>24</td>
<td>10.38</td>
<td>7</td>
</tr>
<tr>
<td>2-methylnaphthalene</td>
<td>284.0</td>
<td>11.79</td>
<td>3</td>
<td>79.2</td>
<td>0.81</td>
<td>3</td>
<td>-0.57</td>
<td>276.0</td>
<td>27.04</td>
<td>32</td>
<td>30.26</td>
<td>12</td>
</tr>
<tr>
<td>chrysene</td>
<td>312.3</td>
<td>12.50</td>
<td>3</td>
<td>547.0</td>
<td>3.79</td>
<td>3</td>
<td>-0.97</td>
<td>291.0</td>
<td>15.82</td>
<td>9</td>
<td>17.96</td>
<td>5</td>
</tr>
<tr>
<td>benzo[ghi]perylene</td>
<td>327.3</td>
<td>5.51</td>
<td>3</td>
<td>488.0</td>
<td>5.00</td>
<td>3</td>
<td>0.45</td>
<td>307.0</td>
<td>22.96</td>
<td>36</td>
<td>23.08</td>
<td>11</td>
</tr>
<tr>
<td>benzo[e]pyrene</td>
<td>312.7</td>
<td>6.66</td>
<td>3</td>
<td>590.0</td>
<td>12.01</td>
<td>3</td>
<td>0.92</td>
<td>325.0</td>
<td>12.76</td>
<td>28</td>
<td>21.91</td>
<td>14</td>
</tr>
<tr>
<td>benzo[a]anthracene</td>
<td>326.3</td>
<td>9.87</td>
<td>3</td>
<td>378.0</td>
<td>4.04</td>
<td>3</td>
<td>0.96</td>
<td>335.0</td>
<td>12.76</td>
<td>39</td>
<td>20.49</td>
<td>10</td>
</tr>
<tr>
<td>indeno[1,2,3-cd]pyrene</td>
<td>359.7</td>
<td>8.08</td>
<td>3</td>
<td>608.0</td>
<td>7.94</td>
<td>3</td>
<td>-0.98</td>
<td>341.0</td>
<td>29.08</td>
<td>36</td>
<td>32.45</td>
<td>11</td>
</tr>
<tr>
<td>benzo[a]pyrene</td>
<td>361.7</td>
<td>5.03</td>
<td>3</td>
<td>636.0</td>
<td>5.69</td>
<td>3</td>
<td>0.99</td>
<td>358.0</td>
<td>8.67</td>
<td>30</td>
<td>19.72</td>
<td>12</td>
</tr>
<tr>
<td>perylene</td>
<td>412.3</td>
<td>9.61</td>
<td>3</td>
<td>184.0</td>
<td>7.00</td>
<td>3</td>
<td>0.68</td>
<td>397.0</td>
<td>22.96</td>
<td>33</td>
<td>35.29</td>
<td>11</td>
</tr>
<tr>
<td>phenanthrene</td>
<td>433.3</td>
<td>10.02</td>
<td>3</td>
<td>539.0</td>
<td>7.00</td>
<td>3</td>
<td>-0.76</td>
<td>406.0</td>
<td>22.45</td>
<td>39</td>
<td>25.97</td>
<td>9</td>
</tr>
<tr>
<td>benzo[b]fluoranthene</td>
<td>463.6</td>
<td>6.93</td>
<td>3</td>
<td>1089.0</td>
<td>19.50</td>
<td>3</td>
<td>-0.61</td>
<td>453.0</td>
<td>10.71</td>
<td>10</td>
<td>19.17</td>
<td>4</td>
</tr>
<tr>
<td>pyrene</td>
<td>565.3</td>
<td>10.21</td>
<td>3</td>
<td>1129.0</td>
<td>13.75</td>
<td>3</td>
<td>-0.65</td>
<td>581.0</td>
<td>19.90</td>
<td>38</td>
<td>40.48</td>
<td>10</td>
</tr>
<tr>
<td>fluoranthene</td>
<td>638.0</td>
<td>33.42</td>
<td>3</td>
<td>1094.0</td>
<td>9.64</td>
<td>3</td>
<td>-0.38</td>
<td>651.0</td>
<td>25.51</td>
<td>38</td>
<td>41.96</td>
<td>10</td>
</tr>
<tr>
<td>naphthalene</td>
<td>873.7</td>
<td>8.08</td>
<td>3</td>
<td>121.3</td>
<td>2.52</td>
<td>3</td>
<td>-0.26</td>
<td>848.0</td>
<td>48.47</td>
<td>38</td>
<td>57.54</td>
<td>10</td>
</tr>
</tbody>
</table>

**Table 4.13.** Example of summary data and reference values for laboratory 1.

For each compound, $\bar{Y} =$ mean of the control value, $s_Y =$ standard deviation of the control value, $K =$ sample size of the control sample, $\bar{W} =$ mean of the study value, $s_W =$ standard deviation of the study value, $M =$ sample size of the study sample, $\rho_{YW} =$ the correlation coefficient between the control and study samples, $\hat{x} =$ reference value estimate, $\hat{\sigma}_x =$ the uncertainty of the reference value, $\nu_x =$ the degrees of freedom of the reference value uncertainty, $\hat{\sigma}_w =$ the within lab uncertainty and $\nu_w =$ degrees of freedom of the within lab uncertainty.
<table>
<thead>
<tr>
<th>Compound</th>
<th>( \bar{Y} )</th>
<th>( s_Y )</th>
<th>( K )</th>
<th>( W )</th>
<th>( s_W )</th>
<th>( M )</th>
<th>( \rho_{YW} )</th>
<th>( \hat{x} )</th>
<th>( \hat{\sigma}_x )</th>
<th>( \nu_x )</th>
<th>( \hat{\sigma}_w )</th>
<th>( \nu_w )</th>
</tr>
</thead>
<tbody>
<tr>
<td>2,3,5-trimethylnaphthalene</td>
<td>25.7</td>
<td>6.24</td>
<td>3</td>
<td>24.4</td>
<td>1.30</td>
<td>3</td>
<td>0.36</td>
<td>25.5</td>
<td>2.60</td>
<td>29</td>
<td>3.18</td>
<td>9</td>
</tr>
<tr>
<td>acenaphthene</td>
<td>21.7</td>
<td>2.20</td>
<td>3</td>
<td>26.0</td>
<td>1.42</td>
<td>3</td>
<td>0.77</td>
<td>38.4</td>
<td>2.65</td>
<td>30</td>
<td>3.23</td>
<td>13</td>
</tr>
<tr>
<td>dibenz[a,h]anthracene</td>
<td>50.3</td>
<td>14.82</td>
<td>3</td>
<td>110.9</td>
<td>11.63</td>
<td>3</td>
<td>0.98</td>
<td>53.0</td>
<td>5.10</td>
<td>10</td>
<td>3.20</td>
<td>4</td>
</tr>
<tr>
<td>acenaphthylene</td>
<td>39.1</td>
<td>14.32</td>
<td>3</td>
<td>33.1</td>
<td>14.38</td>
<td>3</td>
<td>1.00</td>
<td>53.3</td>
<td>3.27</td>
<td>28</td>
<td>4.80</td>
<td>12</td>
</tr>
<tr>
<td>1-methylphenanthrene</td>
<td>50.6</td>
<td>2.31</td>
<td>3</td>
<td>61.6</td>
<td>7.18</td>
<td>3</td>
<td>0.19</td>
<td>73.2</td>
<td>3.01</td>
<td>26</td>
<td>5.38</td>
<td>12</td>
</tr>
<tr>
<td>biphenyl</td>
<td>38.1</td>
<td>2.99</td>
<td>3</td>
<td>17.2</td>
<td>1.04</td>
<td>3</td>
<td>1.00</td>
<td>74.0</td>
<td>4.08</td>
<td>28</td>
<td>6.29</td>
<td>13</td>
</tr>
<tr>
<td>2,6-dimethylnaphthalene</td>
<td>60.6</td>
<td>5.44</td>
<td>3</td>
<td>27.6</td>
<td>1.73</td>
<td>3</td>
<td>1.00</td>
<td>75.9</td>
<td>2.30</td>
<td>23</td>
<td>3.91</td>
<td>14</td>
</tr>
<tr>
<td>fluorene</td>
<td>50.7</td>
<td>9.46</td>
<td>3</td>
<td>41.5</td>
<td>2.16</td>
<td>3</td>
<td>0.32</td>
<td>85.0</td>
<td>7.65</td>
<td>39</td>
<td>7.24</td>
<td>9</td>
</tr>
<tr>
<td>1-methylnaphthalene</td>
<td>59.6</td>
<td>3.70</td>
<td>3</td>
<td>26.8</td>
<td>1.50</td>
<td>3</td>
<td>1.00</td>
<td>127.0</td>
<td>7.14</td>
<td>30</td>
<td>13.89</td>
<td>12</td>
</tr>
<tr>
<td>anthracene</td>
<td>132.3</td>
<td>14.74</td>
<td>3</td>
<td>103.8</td>
<td>4.76</td>
<td>3</td>
<td>0.58</td>
<td>184.0</td>
<td>9.18</td>
<td>39</td>
<td>10.74</td>
<td>9</td>
</tr>
<tr>
<td>benzo[k]fluoranthene</td>
<td>225.7</td>
<td>72.25</td>
<td>3</td>
<td>474.7</td>
<td>200.46</td>
<td>3</td>
<td>0.97</td>
<td>225.0</td>
<td>9.18</td>
<td>24</td>
<td>10.38</td>
<td>7</td>
</tr>
<tr>
<td>2-methylnaphthalene</td>
<td>116.3</td>
<td>8.74</td>
<td>3</td>
<td>35.9</td>
<td>3.23</td>
<td>3</td>
<td>1.00</td>
<td>276.0</td>
<td>27.04</td>
<td>32</td>
<td>30.26</td>
<td>12</td>
</tr>
<tr>
<td>chrysene</td>
<td>339.0</td>
<td>40.73</td>
<td>3</td>
<td>715.0</td>
<td>45.13</td>
<td>3</td>
<td>-0.04</td>
<td>291.0</td>
<td>15.82</td>
<td>9</td>
<td>17.96</td>
<td>5</td>
</tr>
<tr>
<td>benzo[ghi]perylene</td>
<td>224.7</td>
<td>51.05</td>
<td>3</td>
<td>524.0</td>
<td>37.99</td>
<td>3</td>
<td>0.99</td>
<td>307.0</td>
<td>22.96</td>
<td>36</td>
<td>23.08</td>
<td>11</td>
</tr>
<tr>
<td>benzo[e]pyrene</td>
<td>292.0</td>
<td>38.22</td>
<td>3</td>
<td>660.0</td>
<td>63.38</td>
<td>3</td>
<td>-0.19</td>
<td>325.0</td>
<td>12.76</td>
<td>28</td>
<td>21.91</td>
<td>14</td>
</tr>
<tr>
<td>benz[a]anthracene</td>
<td>237.0</td>
<td>27.87</td>
<td>3</td>
<td>369.3</td>
<td>22.37</td>
<td>3</td>
<td>0.29</td>
<td>335.0</td>
<td>12.76</td>
<td>39</td>
<td>20.49</td>
<td>10</td>
</tr>
<tr>
<td>indeno[1,2,3-cd]pyrene</td>
<td>240.3</td>
<td>49.03</td>
<td>3</td>
<td>555.0</td>
<td>29.82</td>
<td>3</td>
<td>0.79</td>
<td>341.0</td>
<td>29.08</td>
<td>36</td>
<td>32.45</td>
<td>11</td>
</tr>
<tr>
<td>benzo[a]pyrene</td>
<td>232.3</td>
<td>27.74</td>
<td>3</td>
<td>610.0</td>
<td>57.03</td>
<td>3</td>
<td>0.26</td>
<td>358.0</td>
<td>8.67</td>
<td>30</td>
<td>19.72</td>
<td>12</td>
</tr>
<tr>
<td>perylene</td>
<td>273.3</td>
<td>42.57</td>
<td>3</td>
<td>185.3</td>
<td>13.87</td>
<td>3</td>
<td>0.09</td>
<td>397.0</td>
<td>22.96</td>
<td>33</td>
<td>35.29</td>
<td>11</td>
</tr>
<tr>
<td>phenanthrene</td>
<td>298.3</td>
<td>30.92</td>
<td>3</td>
<td>418.3</td>
<td>8.50</td>
<td>3</td>
<td>-0.68</td>
<td>406.0</td>
<td>22.45</td>
<td>39</td>
<td>25.97</td>
<td>9</td>
</tr>
<tr>
<td>benzo[b]fluoranthene</td>
<td>405.3</td>
<td>69.87</td>
<td>3</td>
<td>912.3</td>
<td>55.52</td>
<td>3</td>
<td>0.67</td>
<td>453.0</td>
<td>10.71</td>
<td>10</td>
<td>19.17</td>
<td>4</td>
</tr>
<tr>
<td>pyrene</td>
<td>467.3</td>
<td>47.43</td>
<td>3</td>
<td>1049.7</td>
<td>66.03</td>
<td>3</td>
<td>0.26</td>
<td>581.0</td>
<td>19.90</td>
<td>38</td>
<td>40.48</td>
<td>10</td>
</tr>
<tr>
<td>fluoranthene</td>
<td>513.3</td>
<td>52.54</td>
<td>3</td>
<td>983.7</td>
<td>50.12</td>
<td>3</td>
<td>-0.11</td>
<td>651.0</td>
<td>25.51</td>
<td>38</td>
<td>41.96</td>
<td>10</td>
</tr>
<tr>
<td>naphthalene</td>
<td>356.7</td>
<td>25.32</td>
<td>3</td>
<td>49.5</td>
<td>6.09</td>
<td>3</td>
<td>0.98</td>
<td>848.0</td>
<td>48.47</td>
<td>38</td>
<td>57.54</td>
<td>10</td>
</tr>
</tbody>
</table>

**Table 4.14.** Example of summary data and reference values for laboratory 2.

For each compound, \( \bar{Y} \) = mean of the control value, \( s_Y \) = standard deviation of the control value, \( K \) = sample size of the control sample, \( W \) = mean of the study value, \( s_W \) = standard deviation of the study value, \( M \) = sample size of the study sample, \( \rho_{YW} \) = the correlation coefficient between the control and study samples, \( \hat{x} \) = reference value estimate, \( \hat{\sigma}_x \) = the uncertainty of the reference value, \( \nu_x \) = the degrees of freedom of the reference value uncertainty, \( \hat{\sigma}_w \) = the within lab uncertainty and \( \nu_w \) = degrees of freedom of the within lab uncertainty.
Laboratory 1

<table>
<thead>
<tr>
<th>Method</th>
<th>$\hat{\beta}_0$</th>
<th>$se(\hat{\beta}_0)$</th>
<th>$\hat{\beta}_1$</th>
<th>$se(\hat{\beta}_1)$</th>
<th>$\hat{v}_{01}$</th>
<th>$MSE$</th>
</tr>
</thead>
<tbody>
<tr>
<td>NAIVE</td>
<td>-0.2873</td>
<td>3.7325</td>
<td>1.0148</td>
<td>0.0111</td>
<td>-0.0330</td>
<td>134.24</td>
</tr>
<tr>
<td>CORR-R</td>
<td>-1.0677</td>
<td>2.6483</td>
<td>1.0178</td>
<td>0.0143</td>
<td>-0.0314</td>
<td>134.63</td>
</tr>
<tr>
<td>CORR-N</td>
<td>-1.0677</td>
<td>5.8553</td>
<td>1.0178</td>
<td>0.0295</td>
<td>-0.1580</td>
<td>134.63</td>
</tr>
</tbody>
</table>

Laboratory 2

<table>
<thead>
<tr>
<th>Method</th>
<th>$\hat{\beta}_0$</th>
<th>$se(\hat{\beta}_0)$</th>
<th>$\hat{\beta}_1$</th>
<th>$se(\hat{\beta}_1)$</th>
<th>$\hat{v}_{01}$</th>
<th>$MSE$</th>
</tr>
</thead>
<tbody>
<tr>
<td>NAIVE</td>
<td>25.3365</td>
<td>22.6773</td>
<td>0.6292</td>
<td>0.0657</td>
<td>-1.1847</td>
<td>4543</td>
</tr>
<tr>
<td>CORR-R</td>
<td>24.1266</td>
<td>24.6586</td>
<td>0.6336</td>
<td>0.1188</td>
<td>-2.7267</td>
<td>4544</td>
</tr>
<tr>
<td>CORR-N</td>
<td>24.1266</td>
<td>22.7767</td>
<td>0.6336</td>
<td>0.0678</td>
<td>-1.2316</td>
<td>4544</td>
</tr>
</tbody>
</table>

**Table 4.15.** Coefficients estimate assuming linear bias and constant variance.

Tively the response of the method/laboratory 1 is likely linear but the response of the method/laboratory 2 shows too much noise to conclude anything with certainty.

In order to test if the model is the same for all the compounds within the range of levels of the predictor, the reference values $\hat{x}_a$ in Table 4.13 were split in lower values ($\hat{x}_a \leq \bar{x}$) and higher values ($\hat{x}_a > \bar{x}$). We do not set a fixed value since laboratories can use different CRMs. If the model is the same for all the range then the coefficients of the model fitted for the lower levels and the coefficient of the model fitted for the higher levels should be similar, otherwise there is some risk in assuming the model is the same over the whole range of levels of the predictor. Figures 4.9, 4.10 and 4.11 show the confidence regions for the linear coefficients for three representative cases, obtained using all the data points (dotted ellipse), the lower level data points (solid ellipse) and the high level data points (dashed ellipse), the estimated coefficients and the robust covariance matrix estimator. We can observe that there is larger variance at the higher levels of concentration (the area of the dashed ellipse is larger than the area of the solid ellipse). Laboratory 1 appears to be unbiased since in all the three cases the estimated confidence regions of the coefficients contain the point (0, 1). Laboratory 2 appears to be biased since in all the cases the confidence regions of the coefficients do not contain the point (0, 1). Finally, laboratory 6 appears to be non
linear since at lower levels appear to be unbiased but at higher levels it underestimate
the true value (biased downwards) and the estimated confidence regions are totally
disconnected. Using all the data points the estimates are highly influenced by the
higher levels.

Figures 4.9, 4.10 and 4.11 also show the proportion of one estimated confidence
region covering the other. For laboratory 1 the estimated confidence region of the
coefficients obtained at high levels of concentration contains about 0.966 of the esti-
mated confidence region of the coefficients obtained at lower levels of concentration,
while in the opposite direction it is only about 0.0915 of the area. The coverage
quantities for laboratory 2 are 0.356 and 0.0411, respectively. This is evidence that
the model may be changing over the level of concentration, the area of the ellipses
suggests that the variance is larger at higher levels and we should consider weighted
estimators. The estimated confidence regions for laboratory 6 unlikely intersect with
each other, even though the variability is larger at higher levels.

Table 4.16 shows the estimated 95% confidence intervals respectively, applying
(4.63) for Wald type CIs using the naive estimator, robust estimator and large sample
normal estimator for the covariance matrix of the coefficients using (4.64) and (4.65),
respectively. It also shows the Fieller’s confidence intervals obtained by using (4.100).

The robust interval estimate tends to be wider than the naive confidence interval.
There is no clear pattern when comparing the Fieller’s CIs against the Wald’s CIs.
In general the Fieller’s CIs tend to be slightly wider than the bootstrap confidence
intervals and are similar in width to the Wald’s confidence intervals using the normal-
based covariance matrix.

Table 4.17 shows the results of testing for unbiasedness of the method using (4.66)
and the different covariance matrices estimators of the coefficients. Since $\chi^2_{0.95,2} = 5.99$
there is no strong evidence of bias for laboratory 1 and there is strong evidence of bias
Figure 4.7. Observed data for laboratory 1.

for laboratory 2. However, this diagnostic is not definitive since the test is sensitive to
the presence of contaminating outliers in the sample and the normality assumption.

Table 4.18 shows the estimated variance of the error in the equation for a linear
bias model applying (4.69), (4.72), (4.76) and (4.94) for estimating the variance of
the error in the equation, assumed to be constant over the compounds using the
summary data and (4.75) and (4.94) for estimating the mean coefficient of variation
Table 4.19 shows the bootstrap confidence interval estimates for the linear coefficients while assuming constant variance for the error in the equation. The constant variance was estimated using (4.69). The untrimmed bootstrap estimators were obtained as described by the algorithm 1. Trimmed bootstrap estimators were obtained by discarding the lower and upper 2% of the bootstrap estimates.
Table 4.20 shows the bootstrap estimates for the bias and standard error in the coefficients using (4.83) and (4.84).

Table 4.21 shows the results of testing for unbiasedness of the method using (4.66) and the bootstrap estimates of the covariance matrices of the coefficients. Since $\chi^2_{0.95,2} = 5.99$ there is no strong evidence of bias for laboratory 1 and there is strong evidence of bias for laboratory 2. This is consistent with the previous conclusions.

Figure 4.9. Confidence regions for the linear coefficients for laboratory 1.
Figure 4.10. Confidence regions for the linear coefficients for laboratory 2.

Table 4.22 shows the bootstrapped \( p \) values for testing unbiasedness of the method by using several of the covariance matrix estimators. These \( p \)-values suggest the method of Laboratory 1 is unbiased while the method of Laboratory 2 is biased.

Model Assessment

Figures 4.12 and 4.13 show the residuals versus the adjusted true value \( \hat{x}_a^* \) for laboratory 1 and 6, respectively. These use (4.90) and (4.91) with error bars of

\[
\pm t_{(K_a-1,1-\alpha/2)} \sqrt{\hat{V}[\bar{r}_a]}, \quad \text{with } \hat{V}[\bar{r}_a] \text{ from (4.92)}.
\]

There is no evidence that the linear

\[
\]
model is not adequate over the range of $\hat{x}_a$ for laboratory 1. However, Figure 4.13 shows a quadratic model is needed for laboratory 6.

Figure 4.14 shows the estimated variance of the error in the equation under replication $\sigma^2_{\epsilon a}$ versus the predictor $\hat{x}_a$ for laboratory 1, using (4.75) for the point estimates and (4.20) and (4.21) for the confidence interval estimates. The estimated variances are shown in log scale for clarity. This suggests that the variance of the error in the equation is not constant. In contrast, Figure 4.15 shows the square root of the

Figure 4.11. Confidence regions for the linear coefficients for laboratory 6.
Table 4.16. Parameter confidence intervals assuming linear bias and constant variance.

<table>
<thead>
<tr>
<th>Method</th>
<th>$lcl(\hat{\beta}_0)$</th>
<th>$ucl(\hat{\beta}_0)$</th>
<th>$lcl(\hat{\beta}_1)$</th>
<th>$ucl(\hat{\beta}_1)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>NAIVE</td>
<td>-7.9907</td>
<td>7.4161</td>
<td>0.9918</td>
<td>1.0378</td>
</tr>
<tr>
<td>CORR-R</td>
<td>-6.2584</td>
<td>4.1229</td>
<td>0.9897</td>
<td>1.0458</td>
</tr>
<tr>
<td>CORR-R(T)</td>
<td>-6.5336</td>
<td>4.3982</td>
<td>0.9882</td>
<td>1.0473</td>
</tr>
<tr>
<td>CORR-N</td>
<td>-12.5439</td>
<td>10.4084</td>
<td>0.9599</td>
<td>1.0756</td>
</tr>
<tr>
<td>CORR-N(T)</td>
<td>-13.1524</td>
<td>11.0169</td>
<td>0.9569</td>
<td>1.0786</td>
</tr>
<tr>
<td>FIELLER</td>
<td>-20.5454</td>
<td>12.3559</td>
<td>0.9672</td>
<td>1.0911</td>
</tr>
</tbody>
</table>

Table 4.17. Test for unbiasedness of the method assuming linear bias model and constant variance.

<table>
<thead>
<tr>
<th>Method</th>
<th>$\chi^2$</th>
<th>$p$-value</th>
<th>$\chi^2$</th>
<th>$p$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>NAIVE</td>
<td>4.35</td>
<td>0.1136</td>
<td>62.64</td>
<td>2.5e-14</td>
</tr>
<tr>
<td>CORR-R</td>
<td>2.79</td>
<td>0.2481</td>
<td>36.44</td>
<td>1.2e-8</td>
</tr>
<tr>
<td>CORR-N</td>
<td>1.20</td>
<td>0.5501</td>
<td>58.08</td>
<td>2.4e-13</td>
</tr>
</tbody>
</table>

Table 4.18. Estimated variance of the error in equation assuming linear bias.

<table>
<thead>
<tr>
<th>Laboratory</th>
<th>Method</th>
<th>$\hat{\sigma}^2_{\epsilon,mm}$</th>
<th>$\hat{\sigma}^2_{\epsilon}$</th>
<th>$\hat{\sigma}^2_{\epsilon,rep}$</th>
<th>$\hat{\sigma}^2_{\epsilon(msr)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>NAIVE</td>
<td>386.6</td>
<td>402.7</td>
<td>85.79</td>
<td>38.68</td>
</tr>
<tr>
<td>1</td>
<td>CORR</td>
<td>0.0</td>
<td>0.0</td>
<td>85.79</td>
<td>42.71</td>
</tr>
<tr>
<td>2</td>
<td>NAIVE</td>
<td>13036</td>
<td>13629</td>
<td>1213</td>
<td>12212</td>
</tr>
<tr>
<td>2</td>
<td>CORR</td>
<td>12656</td>
<td>13248</td>
<td>1213</td>
<td>12211</td>
</tr>
</tbody>
</table>

modified squared residual $msr_a$ versus the modified predictor $\hat{x}_a^*$, using (4.91) and (4.93). The constant variance of the error in equation and the 95% upper confidence limit are shown as dotted and dashed lines respectively. This shows the problem of
Table 4.19. Parameter confidence interval estimates using bootstrap and assuming linear bias and constant variance for the error in the equation.

<table>
<thead>
<tr>
<th>Laboratory</th>
<th>$lcl(\beta_0)^*$</th>
<th>$ucl(\beta_0)^*$</th>
<th>$lcl(\beta_1)^*$</th>
<th>$ucl(\beta_1)^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-12.8680</td>
<td>10.0225</td>
<td>0.9631</td>
<td>1.0767</td>
</tr>
<tr>
<td>1</td>
<td>trimmed</td>
<td>-11.3035</td>
<td>8.5293</td>
<td>0.9693</td>
</tr>
<tr>
<td>2</td>
<td>untrimmed</td>
<td>-19.77</td>
<td>66.89</td>
<td>0.508</td>
</tr>
<tr>
<td>2</td>
<td>trimmed</td>
<td>-14.39</td>
<td>60.76</td>
<td>0.522</td>
</tr>
</tbody>
</table>

Table 4.20. Parameter bias and standard error estimates using bootstrap and assuming linear bias and constant variance for the error in the equation.

<table>
<thead>
<tr>
<th>Laboratory</th>
<th>bias$(\beta_0)^*$</th>
<th>se$(\beta_0)^*$</th>
<th>bias$(\beta_1)^*$</th>
<th>se$(\beta_1)^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.1097</td>
<td>5.7865</td>
<td>0.0009</td>
<td>0.0293</td>
</tr>
<tr>
<td>1</td>
<td>trimmed</td>
<td>-0.1216</td>
<td>5.1573</td>
<td>0.0009</td>
</tr>
<tr>
<td>2</td>
<td>untrimmed</td>
<td>-0.2165</td>
<td>22.1174</td>
<td>0.0014</td>
</tr>
<tr>
<td>2</td>
<td>trimmed</td>
<td>-0.2083</td>
<td>19.6725</td>
<td>0.0012</td>
</tr>
</tbody>
</table>

Table 4.21. Test for unbiasedness of the method assuming linear bias model and constant variance.

<table>
<thead>
<tr>
<th>Method</th>
<th>Laboratory 1 $\chi^2$</th>
<th>Laboratory 2 $\chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Untrimmed bootstrap</td>
<td>1.14</td>
<td>48.65</td>
</tr>
<tr>
<td>Trimmed bootstrap</td>
<td>1.19</td>
<td>51.42</td>
</tr>
</tbody>
</table>

Table 4.22. Bootstrapped p-values for testing unbiasedness of the method assuming linear bias model.

<table>
<thead>
<tr>
<th>Method</th>
<th>Laboratory 1 $p-value$</th>
<th>Laboratory 2 $p-value$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Robust</td>
<td>0.6661</td>
<td>0.0006</td>
</tr>
<tr>
<td>Normal</td>
<td>0.2474</td>
<td>0.0002</td>
</tr>
<tr>
<td>Bootstrap</td>
<td>0.2454</td>
<td>0.0002</td>
</tr>
<tr>
<td>Robust Bootstrap</td>
<td>0.2442</td>
<td>0.0002</td>
</tr>
<tr>
<td>Normal Bootstrap</td>
<td>0.2524</td>
<td>0.0002</td>
</tr>
</tbody>
</table>

Trying to assess for changing variance using $msr$ when a large number of the estimates $msr_a$ are zero. The evidence for a linear relationship is poor and it is consistent with
the initial estimate of a constant zero variance of the error in the equation found previously (see Table 4.18). Here, the use of the modified squared residual is just for illustration since we can use variance estimates from replicates.

![Figure 4.12. Example of residual versus adjusted true value estimate.](image)

This plot shows the residuals versus the modified predictor ($\hat{x}_a$) for each compound from the PAHs family in Sediments 2003 exercise for laboratory 1. No pattern is evidently present.

Table 4.23 shows the weighted point estimates and the estimated robust covariance for laboratory 2 for illustration purposes only. Table 4.24 shows the weighted
Figure 4.13. Example of residual versus adjusted true value estimate.

This plot shows the residuals versus the modified predictor ($\hat{x}_{a}$) for each compound from the PAHs family in Sediments 2003 exercise for laboratory 6. A probable quadratic trend is shown.

interval estimates, applying (4.110), (4.111) and (4.112). The weighted estimates using the normal based weighted covariance matrix suggests the method is still biased while using the robust weighted covariance matrix suggests the method is apparently unbiased. However due to the correlation between the coefficients, it suggests the method is still biased, furthermore the robust estimates suggest the true slope might
Figure 4.14. Example of the estimated variance of the error in the equation versus adjusted true value estimate.

This plot shows the estimated variance of the error in the equation ($\hat{\sigma}_{ea}$) versus the modified predictor ($\hat{x}^*_a$) in a log-log scale for each compound from the PAHs family in Sediments 2003 exercise for laboratory 1. The dashed line is the mean fitted variance, the dotted lines are the 95% confidence intervals for the estimated variance, respectively.

be zero and there is some risk on assuming a linear relationship does exist between the reference values and the observed control values.
Figure 4.15. Example of the modified squared residual versus adjusted true value estimate.

This plot shows the square root of the modified squares residuals ($\sqrt{msr_{a}}$) versus the modified predictor ($\hat{x}_{a}^{*}$) for each compound from the PAHs family in Sediments 2003 exercise for laboratory 1. The dashed line is the predicted mean standard deviation of the residuals, the dotted lines are the 95% confidence intervals for the predicted standard deviation of the residuals.

Performance Evaluation

The performance can be evaluated by using (4.88) and (3.23). Figures 4.16 and 4.17 show the $t_{score}$ and $\sqrt{f_{score}}$ obtained while using the corrected estimators for the linear bias model coefficients for the laboratory 1 and laboratory 2, respectively.
Table 4.23. Parameter weighted estimates assuming linear bias and changing variance of the error in the equation.

<table>
<thead>
<tr>
<th>Method</th>
<th>$\hat{\beta}_0$</th>
<th>se($\hat{\beta}_0$)</th>
<th>$\hat{\beta}_1$</th>
<th>se($\hat{\beta}_1$)</th>
<th>$\hat{\sigma}_{01}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>CORR-R</td>
<td>-1.4019</td>
<td>29.7326</td>
<td>0.6755</td>
<td>0.4036</td>
<td>-11.948</td>
</tr>
<tr>
<td>CORR-N</td>
<td>-1.4019</td>
<td>1.7864</td>
<td>0.6755</td>
<td>0.0185</td>
<td>-0.0249</td>
</tr>
</tbody>
</table>

Table 4.24. Parameter weighted confidence intervals assuming linear bias and changing variance of the error in the equation.

<table>
<thead>
<tr>
<th>Method</th>
<th>lcl($\hat{\beta}_0$)</th>
<th>ucl($\hat{\beta}_0$)</th>
<th>lcl($\hat{\beta}_1$)</th>
<th>ucl($\hat{\beta}_1$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CORR-R</td>
<td>-72.92</td>
<td>70.12</td>
<td>-0.2953</td>
<td>1.6464</td>
</tr>
<tr>
<td>CORR-N</td>
<td>-5.699</td>
<td>2.895</td>
<td>0.6309</td>
<td>0.7201</td>
</tr>
</tbody>
</table>

Figures 4.16 and 4.17 show the individual scores of the linear bias expressed as $t_{score}$ and the individual uncertainty ratio relative to the reference uncertainty for each compound from the PAHs in the Sediments 2003 exercise for laboratory 1 (Figure 4.16) and 2 (Figure 4.17). The black dotted semicircles represent the approximate boundaries for satisfactory and unsatisfactory results. The black dot represents the average of the scores. The red triangle represents a data point with absolute value larger than 4 in either score.

Table 4.25 lists the estimated bias model for each participating laboratory in the 2003 exercise for PAHs family compound in sediments using the means data only and the robust covariance estimate. Since $\chi^2(2) = 5.99$ only laboratories 1, 1b, 1c and 18 appear to be unbiased. Laboratory 10 appears to have a linear bias model however there is some risk on assuming a linear relationship between the reference values and the observed control data since the slope might be zero (written in parentheses).

Table 4.26 lists the corresponding weighted estimates for the bias model, using the means and variance of the replicated data. The weighted estimates suggest that laboratories 1, 1b, 1c, 7, 11 and 15 are unbiased and laboratory 19 is biased. Also, there is some risk on assuming a linear relationship between the reference values and
Figure 4.16. Example of performance evaluation under the linear bias model assuming constant variance of error in the equation for laboratory 1.

the observed values for laboratories 2, 4, 5, 6, 9, 10, 13, 17 and 18 since the slope could be zero (written in parentheses).

4.2.1.14 Small Range Example

In this section we conduct an analysis of a piece of the data from laboratory 1 and 2 in order to illustrate the impact of the measurement error when it becomes larger with respect to the dispersion of the predictor. Consider the compounds benzo[e]pyrene to perylene from Table 4.13 (rows 17 to 21) and Table 4.14 (rows 15 to 19). Note the reference values $\hat{x}$ are close to each other.

Table 4.27 shows the estimated coefficients for a linear bias model for the subset of the data showed in Tables 4.13 and 4.14, applying (4.57) for naive point estimators, (4.60), (4.77), (4.78) and (4.80), if required, for corrected point estimators respectively. The estimated standard error and the covariance were obtained by using
(4.64), (4.65) for the robust and normal asymptotic standard errors and covariance of the coefficients.

Contrasting with the estimates obtained using the whole dataset shown in Table 4.15 we can see that the difference between the naive and the corrected estimators is larger for the estimates using a subset of the data with a small range (dispersion) of values for the predictor, while the variance of the measurement error grows relative to the dispersion of the predictor values. For laboratory 1, the relative change in the estimated slope is about 0.3% and 14% when using the whole dataset and the subset with small range, respectively. For laboratory 2, the relative change in the estimated slope is even larger, about 0.7% and 81% when using the whole dataset and the subset with small range, respectively. The increase in the standard error can be due to the relatively noisy measurements for these compounds and with the small sample size.
Table 4.25. Parameter pooled estimates using the means data only, assuming linear bias and constant variance of the error in the equation.

<table>
<thead>
<tr>
<th>Laboratory</th>
<th>Equation</th>
<th>$\hat{\beta}_0(se)$</th>
<th>$\hat{\beta}_1(se)$</th>
<th>$\hat{\sigma}_{\beta_0\beta_1}$</th>
<th>$MSE_c$</th>
<th>$\hat{\sigma}_\epsilon^2$</th>
<th>$\chi^2*$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4.77</td>
<td>-1.07(2.65)</td>
<td>1.018(0.014)</td>
<td>-0.03</td>
<td>134.6</td>
<td>0</td>
<td>2.79</td>
</tr>
<tr>
<td>1b</td>
<td>4.59</td>
<td>-2.9(12.1)</td>
<td>1.034(0.055)</td>
<td>-0.55</td>
<td>1848</td>
<td>4521</td>
<td>0.66</td>
</tr>
<tr>
<td>1c</td>
<td>4.59</td>
<td>-6.70(9.33)</td>
<td>1.058(0.036)</td>
<td>-0.15</td>
<td>2448</td>
<td>6338</td>
<td>2.59</td>
</tr>
<tr>
<td>2</td>
<td>4.59</td>
<td>24.1(24.7)</td>
<td>0.634(0.119)</td>
<td>-2.78</td>
<td>4544</td>
<td>13250</td>
<td>36.44</td>
</tr>
<tr>
<td>4</td>
<td>4.59</td>
<td>59.5(23.9)</td>
<td>0.608(0.071)</td>
<td>-1.42</td>
<td>3412</td>
<td>9834</td>
<td>45.62</td>
</tr>
<tr>
<td>5</td>
<td>4.59</td>
<td>931(357)</td>
<td>0.683(0.101)</td>
<td>-28.66</td>
<td>998800</td>
<td>2977000</td>
<td>9.88</td>
</tr>
<tr>
<td>6</td>
<td>4.59</td>
<td>307(102)</td>
<td>0.839(0.026)</td>
<td>-2.08</td>
<td>76350</td>
<td>200400</td>
<td>49.33</td>
</tr>
<tr>
<td>7</td>
<td>4.59</td>
<td>10.8(7.4)</td>
<td>0.733(0.032)</td>
<td>-0.01</td>
<td>3251</td>
<td>9217</td>
<td>72.42</td>
</tr>
<tr>
<td>9</td>
<td>4.59</td>
<td>43.0(34.6)</td>
<td>0.436(0.156)</td>
<td>-4.95</td>
<td>9032</td>
<td>26910</td>
<td>39.32</td>
</tr>
<tr>
<td>10</td>
<td>4.59</td>
<td>822(858)</td>
<td>4.786(3.510)</td>
<td>-2843</td>
<td>3537000</td>
<td>10580000</td>
<td>(37.33)</td>
</tr>
<tr>
<td>11</td>
<td>4.59</td>
<td>15.0(7.2)</td>
<td>0.916(0.034)</td>
<td>-0.12</td>
<td>1936</td>
<td>5004</td>
<td>7.10</td>
</tr>
<tr>
<td>13</td>
<td>4.59</td>
<td>-11.3(15.4)</td>
<td>0.928(0.030)</td>
<td>-0.29</td>
<td>2755</td>
<td>5969</td>
<td>13.74</td>
</tr>
<tr>
<td>15</td>
<td>4.59</td>
<td>26.4(16.6)</td>
<td>1.074(0.033)</td>
<td>-0.35</td>
<td>4261</td>
<td>11680</td>
<td>21.53</td>
</tr>
<tr>
<td>17</td>
<td>4.59</td>
<td>80.5(159)</td>
<td>0.642(0.025)</td>
<td>-3.20</td>
<td>202300</td>
<td>587200</td>
<td>574.1</td>
</tr>
<tr>
<td>18</td>
<td>4.59</td>
<td>-673.2(477)</td>
<td>1.223(0.183)</td>
<td>-80.56</td>
<td>1364000</td>
<td>4031000</td>
<td>2.05</td>
</tr>
<tr>
<td>19</td>
<td>4.59</td>
<td>-2.9(23.2)</td>
<td>1.293(0.102)</td>
<td>-2.14</td>
<td>3951</td>
<td>10110</td>
<td>41.06</td>
</tr>
</tbody>
</table>

Similar problems can rise when working with one compound over time while using similar CRMs.

4.2.2 Exporting the Model to the Study Data

In this section we use the bias model obtained from the control data to correct the study data for linear bias.

4.2.2.1 Parameter Estimation

*Under Independent Errors*

First assume that the error terms are independent for the method measuring the control sample and the study sample; i.e., $\text{Corr}[\epsilon_{ak}, \delta_{ak}] = \rho_a = 0$. Also assume $\sigma^2_{\epsilon a} = \sigma^2_{\delta a}$, i.e. the variance of the method is constant across the control and study samples for each compound. This can be due to the constancy of the method’s
Table 4.26. Parameter pooled weighted estimates, assuming linear bias.

<table>
<thead>
<tr>
<th>Laboratory</th>
<th>$\hat{\beta}_0 (se)$</th>
<th>$\hat{\beta}_1 (se)$</th>
<th>$\hat{\sigma}_{\beta_0 \beta_1}$</th>
<th>$\chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-2.27(21.76)</td>
<td>1.022(0.023)</td>
<td>-4.89</td>
<td>0.09</td>
</tr>
<tr>
<td>1b</td>
<td>1.087(31.19)</td>
<td>0.960(0.324)</td>
<td>-10.05</td>
<td>0.80</td>
</tr>
<tr>
<td>1c</td>
<td>-0.618(22.89)</td>
<td>0.959(0.284)</td>
<td>-6.42</td>
<td>1.08</td>
</tr>
<tr>
<td>2</td>
<td>-1.402(29.73)</td>
<td>0.676(0.404)</td>
<td>-11.95</td>
<td>(83.83)</td>
</tr>
<tr>
<td>4</td>
<td>21.62(51.08)</td>
<td>0.700(0.544)</td>
<td>-27.52</td>
<td>(1.08)</td>
</tr>
<tr>
<td>5</td>
<td>81.81(1256)</td>
<td>0.859(1.150)</td>
<td>-1441.5</td>
<td>(1.08)</td>
</tr>
<tr>
<td>6</td>
<td>121.5(1904)</td>
<td>0.885(1.696)</td>
<td>-3229.7</td>
<td>(0.15)</td>
</tr>
<tr>
<td>7</td>
<td>19.41(25.18)</td>
<td>0.709(0.283)</td>
<td>-6.78</td>
<td>(1.51)</td>
</tr>
<tr>
<td>9</td>
<td>-7.13(23.58)</td>
<td>0.591(0.281)</td>
<td>-6.53</td>
<td>(129.09)</td>
</tr>
<tr>
<td>10</td>
<td>-38.50(311.9)</td>
<td>7.973(3.361)</td>
<td>-1019</td>
<td>(69.39)</td>
</tr>
<tr>
<td>11</td>
<td>9.39(25.07)</td>
<td>1.002(0.317)</td>
<td>-7.801</td>
<td>3.72</td>
</tr>
<tr>
<td>13</td>
<td>-10.83(94.85)</td>
<td>0.918(0.807)</td>
<td>-76.19</td>
<td>(4.71)</td>
</tr>
<tr>
<td>15</td>
<td>4.13(27.49)</td>
<td>1.186(0.347)</td>
<td>-9.10</td>
<td>5.22</td>
</tr>
<tr>
<td>17</td>
<td>-57.16(2974)</td>
<td>0.571(3.500)</td>
<td>-10409</td>
<td>(84.56)</td>
</tr>
<tr>
<td>18</td>
<td>-215.53(1217)</td>
<td>1.058(0.935)</td>
<td>-1133</td>
<td>(1.80)</td>
</tr>
<tr>
<td>19</td>
<td>20.24(29.11)</td>
<td>1.131(0.385)</td>
<td>-10.89</td>
<td>19.67</td>
</tr>
</tbody>
</table>

Table 4.27. Coefficients estimate assuming linear bias and constant variance for a subset with small range of the predictor.

<table>
<thead>
<tr>
<th>Laboratory 1</th>
<th>Method</th>
<th>$\hat{\beta}_0$</th>
<th>$se(\hat{\beta}_0)$</th>
<th>$\hat{\beta}_1$</th>
<th>$se(\hat{\beta}_1)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>NAIVE</td>
<td>13.3380</td>
<td>74.283</td>
<td>0.9900</td>
<td>0.2129</td>
<td></td>
</tr>
<tr>
<td>CORR-R</td>
<td>-34.4180</td>
<td>145.341</td>
<td>1.1274</td>
<td>0.4181</td>
<td></td>
</tr>
<tr>
<td>CORR-N</td>
<td>-34.4180</td>
<td>275.474</td>
<td>1.1274</td>
<td>0.7912</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Laboratory 2</th>
<th>Method</th>
<th>$\hat{\beta}_0$</th>
<th>$se(\hat{\beta}_0)$</th>
<th>$\hat{\beta}_1$</th>
<th>$se(\hat{\beta}_1)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>NAIVE</td>
<td>237.715</td>
<td>188.558</td>
<td>0.0492</td>
<td>0.5355</td>
<td></td>
</tr>
<tr>
<td>CORR-R</td>
<td>223.704</td>
<td>262.524</td>
<td>0.0891</td>
<td>0.7320</td>
<td></td>
</tr>
<tr>
<td>CORR-N</td>
<td>223.704</td>
<td>296.071</td>
<td>0.0891</td>
<td>0.8424</td>
<td></td>
</tr>
</tbody>
</table>

variance over different levels or to a method’s variance dependent on the concentration level but analyzing samples with similar levels for each compound.

Under independent error terms, the number of replicates of the control material and the number of replicates of the study material are not required to be equal.

Define the sample mean $\bar{W}_a = \frac{1}{M_a} \sum_{m=1}^{M_a} W_{am}$ which is unbiased for $E[W_{am}] = \beta_0 + \beta_1 d_a$ and the sample variance $s^2_{W_a} = \frac{1}{M_a-1} \sum_{m=1}^{M_a} (W_{am} - \bar{W}_a)^2$ which is unbiased.
for $V[W_{am}] = \beta_1^2 \sigma_{Va}^2 + \sigma_{ea}^2$. Let

$$\hat{d}_a = \frac{\bar{W}_a - \hat{\beta}_0}{\hat{\beta}_1},$$

(4.113)

and

$$\hat{\sigma}_{Va}^2 = \frac{s_{Wa}^2 - \hat{\sigma}_{ea}^2}{\hat{\beta}_1^2}.$$

(4.114)

The ability to estimate $\sigma_{Va}^2$ depends on the assumption $\sigma_{da}^2 = \sigma_{ea}^2$ and $\hat{\sigma}_{ea}^2$ being readily available from the control data. If either assumption is invalid then neither can be estimated.

These estimators are biased but consistent under the assumptions $M_a \to \infty$ in the study data and $(\hat{\beta}_1, \hat{\sigma}_{ea}^2) \to (\beta_1 \neq 0, \sigma_{ea}^2)$ from the control data, so

$$\frac{\bar{W}_a - \hat{\beta}_0}{\hat{\beta}_1} \overset{p}{\to} \frac{\beta_1 d_a}{\beta_1} = d_a,$$

(4.115)

$$\frac{\hat{\sigma}_{Va}^2}{\hat{\beta}_1^2} \overset{p}{\to} \frac{\beta_1^2 \sigma_{Va}^2 + \sigma_{ea}^2 - \sigma_{ea}^2}{\beta_1^2} = \sigma_{Va}. $$

### 4.2.2.2 Inferences for $d_a$

An approximation of the variance of $\hat{d}_a$ can be obtained from (4.113), by using the delta method

$$\sqrt{V[\hat{d}_a]} \approx \frac{s_{Wa}^2 + \hat{\sigma}_{Va}^2 + \hat{d}_a^2 \hat{\sigma}_{Va}^2 + 2\hat{d}_a \hat{\sigma}_{Va} \hat{\beta}_0 \hat{\beta}_1}{\hat{\beta}_1^2}.$$

(4.116)

Some additional distributional assumptions on the error terms are required in order to get confidence intervals for $d_a$ and $\sigma_{Va}^2$. Assuming normality for $\bar{W}_a - (\hat{\beta}_0 + \hat{\beta}_1 d_a)$, the distribution can be approximated as

$$\frac{\bar{W}_a - \hat{\beta}_0 - \hat{\beta}_1 d_a}{\sqrt{V[\bar{W}_a - \hat{\beta}_0 - \hat{\beta}_1 d_a]}} \sim Z,$$

(4.117)

or

$$\frac{\bar{W}_a - \hat{\beta}_0 - \hat{\beta}_1 d_a}{\sqrt{V[\bar{W}_a - \hat{\beta}_0 - \hat{\beta}_1 d_a]}} \sim T_{(\nu_{da})},$$

(4.118)
where

\[ 
\hat{V}[\hat{W}_a - \hat{\beta}_0 - \hat{\beta}_1 d_a] = s^2_{Wa} + \hat{\sigma}^2_{\beta_0} + \hat{d}^2_a \hat{\sigma}^2_{\beta_1} + 2 \hat{d}_a \hat{\sigma}_{\beta_0 \beta_1},
\]

\[ 
\nu_{d_a} \approx \frac{2 \left( \hat{V}[\hat{W}_a - \hat{\beta}_0 - \hat{\beta}_1 d_a] \right)^2}{\hat{V}[\hat{V}[\hat{W}_a - \hat{\beta}_0 - \hat{\beta}_1 d_a]]}
\]

\[ 
\hat{V} \left[ \hat{V}[\hat{W}_a - \hat{\beta}_0 - \hat{\beta}_1 d_a] \right] = \frac{\nu_{d_a}}{\nu_{d_a}} + \frac{\hat{\sigma}^4_{\beta_0}}{\nu_{\beta_0}} + \frac{\hat{d}^4_a \hat{\sigma}^4_{\beta_1}}{\nu_{\beta_1}} + 4 \hat{d}^2_a \hat{\sigma}_{\beta_0 \beta_1}
\]

\[ 
+ 2 \hat{d}^2_a \text{Cov}[\hat{\sigma}^2_{\beta_0}, \hat{\sigma}^2_{\beta_1}] + 4 \hat{d}_a \text{Cov}[\hat{\sigma}^2_{\beta_0}, \hat{\sigma}_{\beta_0 \beta_1}]
\]

\[ 
+ 4 \hat{d}^3_a \text{Cov}[\hat{\sigma}^2_{\beta_1}, \hat{\sigma}_{\beta_0 \beta_1}].
\]

(4.119)

and the covariance terms can be obtained applying the delta method.

Using \( \hat{d}_a \) for \( d_a \) in the denominator of (4.118), a (1 - \( \alpha \))100\% delta method confidence interval for \( d_a \) then becomes

\[ 
\left( \hat{d}_a + t(\alpha/2, \nu_{d_a}) \sqrt{\hat{V}[\hat{d}_a]}, \hat{d}_a + t(1-\alpha/2, \nu_{d_a}) \sqrt{\hat{V}[\hat{d}_a]} \right),
\]

(4.120)

**Fieller’s method**

We can apply Fieller’s method to obtain a CI for \( d_a \), assuming \( \hat{d}_a \) is approximately normal and \( \hat{V}[\hat{d}_a] = \sqrt{s^2_{Wa} + \hat{\sigma}^2_{\beta_0} + \hat{d}^2_a \hat{\sigma}^2_{\beta_1} + 2 \hat{d}_a \hat{\sigma}_{\beta_0 \beta_1}} \) is distributed proportional to a chi-square(\( \nu_{d_a} \)) distribution, Then for a specific significance level \( \alpha \)

\[ 
(1 - \alpha) \approx Pr \left( \left| \frac{\hat{d}_a - d_a}{\sqrt{\hat{V}[\hat{d}_a]}} \right| \leq T_{(1-\alpha/2, \nu_{d_a})} \right)
\]

\[ 
\approx Pr \left( \left| \frac{W_a - \hat{\beta}_0 - \hat{\beta}_1 d_a}{\sqrt{s^2_{Wa} + \hat{\sigma}^2_{\beta_0} + \hat{d}^2_a \hat{\sigma}^2_{\beta_1} + 2 \hat{d}_a \hat{\sigma}_{\beta_0 \beta_1}}} \right| \leq T_{(1-\alpha/2, \nu_{d_a})} \right)
\]

\[ 
= Pr(q(d_a) \leq 0),
\]

where \( q(x) \) is a quadratic function on \( x \), with
\[ q(x) = f_2 x^2 - 2f_1 x + f_0, \]
\[ f_0 = (\hat{W}_a - \hat{\beta}_0)^2 - t_{(1-\alpha/2,\nu_{da})}^2 (\hat{\sigma}_a^2 + \hat{\sigma}_{\beta_0}^2), \tag{4.122} \]
\[ f_1 = \hat{\beta}_1 (W_a - \hat{\beta}_0) + t_{(1-\alpha/2,\nu_{da})}^2 \hat{\sigma}_{\beta_0} \hat{\beta}_1, \]
\[ f_2 = \hat{\beta}_2 - t_{(1-\alpha/2,\nu_{da})}^2 \hat{\sigma}_{\beta_1}. \]

Hence an approximate confidence interval for \( \hat{d}_a \) is the set \( \{ d : q(d) \leq 0 \} \). Defining \( D = f_1^2 - f_0 f_2, \) \( r_1 = (f_1 - D^{1/2})/f_2, \) and \( r_2 = (f_1 + D^{1/2})/f_2, \) the approximate confidence set can be expressed as
\[
\begin{cases}
[r_1, r_2] & \text{if } f_2 > 0, \\
(-\infty, r_2] \cup [r_1, \infty) & \text{if } f_2 \leq 0 \text{ and } D > 0, \\
(-\infty, \infty) & \text{if } f_2 \leq 0 \text{ and } D \leq 0.
\end{cases}
\tag{4.123}
\]

This confidence set can contain negative values, which must be truncated to zero since \( d_a \geq 0. \)

**Bootstrap estimates**

We can apply the parametric bootstrap method by modifying the bootstrap cycle described in Section 4.2.1.7. After computing \( (\hat{\beta}_{0b}, \hat{\beta}_{1b}) \) we need to to resample the study data parametrically as described in the algorithm 5.

Then compute the sets \( D_a = \{ d_{ab} : b = 1, \ldots, B \} \). A \((1-\alpha)100\%\) CI for \( d_a \) is \( (Q_{(\alpha/2)}(D_a), Q_{(1-\alpha/2)}(D_a)) \), where \( Q \) is as in Section 4.1.2.1.

The bootstrap estimates of the bias and standard error of \( d_a \) are
\[
\text{bias}[\hat{d}_a]^* = \overline{d}_a - \hat{d}_a,
\]
\[
\text{se}[\hat{d}_a]^* = \sqrt{1/B - 1/B \sum_{b=1}^{B} (\hat{d}_{ab} - \overline{d}_a)^2}, \tag{4.124}
\]
\[
\overline{d}_a^* = \frac{1}{B} \sum_{b=1}^{B} \hat{d}_{ab}.
\]
Algorithm 5  Bootstrap algorithm for corrected study data

for $b = 1$ to $B$(large) do
  \( \hat{x}_{ab} \leftarrow \hat{x}_a + \hat{\sigma}_a Z_{1ab}, a = 1, \ldots, A, \)
  \( \hat{\sigma}_{xab}^2 \leftarrow \hat{\sigma}_{2a}^2 \chi^2(\nu_{2a}), a = 1, \ldots, A, \)
  \( \hat{Y}_{ab} \leftarrow \hat{\beta}_0 + \hat{\beta}_1 \hat{x}_a + Z_{2ab} \sqrt{\hat{\beta}_1^2 \hat{\sigma}_{1a}^2 + \hat{\sigma}_{ea}^2}, a = 1, \ldots, A, \)
  where \((Z_{1ab}, Z_{2ab})\) are independent, each distributed as a standard normal,
  Apply (4.60) to \((\hat{x}_{ab}, \hat{Y}_{ab})\) to get \((\hat{\beta}_{0,b}, \hat{\beta}_{1,b})\)
  Apply (4.69) or (4.72) to obtain \(\hat{\sigma}_{e,b}^2\)
  If \(S_{\chi_b}^2 > \hat{\sigma}_{e,b}^2\) and \(\hat{\sigma}_{e,b}^2 \leq 0\) then apply (4.77) to \((\hat{x}_{ab}, \hat{Y}_{ab})\) to get a modified estimator \((\hat{\beta}_{0,b}, \hat{\beta}_{1,b}).\)
  If \(S_{\chi_b}^2 \leq \hat{\sigma}_{e,b}^2\) and \(\hat{\sigma}_{e,b}^2 > 0\) then apply (4.80) to \((\hat{x}_{ab}, \hat{Y}_{ab})\) to get a modified estimator \((\hat{\beta}_{0,b}, \hat{\beta}_{1,b}).\)
  If \(S_{\chi_b}^2 \leq \hat{\sigma}_{e,b}^2\) and \(\hat{\sigma}_{e,b}^2 \leq 0\) then warning: model (4.51) is not supported by the data and drop this estimate.
  generate \(W_{a,b} \sim N(\hat{\beta}_{0,b} + \hat{\beta}_{1,b} d_a, s_{Wa}^2/M_a)\)
  Apply (4.113) to obtain \(\hat{d}_{a,b} = \frac{W_{a,b} - \hat{\beta}_{0,b}}{\hat{\beta}_{1,b}}, a = 1, \ldots, A\)
end for

4.2.2.3 Inferences for \(\sigma_{V}^2\)

The distribution of \(\hat{\sigma}_{V_a}^2\) and its associated confidence interval can be approximated as

\[
\frac{(M_a - 1)s_{Wa}^2}{\hat{\beta}_1^2 \hat{\sigma}_{V_a}^2 + \hat{\sigma}_{ea}^2} \sim \chi^2(M - 1) \tag{4.125}
\]

and an approximate \(100(1 - \alpha)\%)\ CI for \(\sigma_{V_a}^2\) is

\[
\left( \frac{(M_a - 1)s_{Wa}^2}{\hat{\beta}_1^2 \chi^2(1 - \alpha/2, M_a - 1)} - \hat{\sigma}_{ea}^2, \frac{(M_a - 1)s_{Wa}^2}{\hat{\beta}_1^2 \chi^2(\alpha/2, M_a - 1)} - \hat{\sigma}_{ea}^2 \right). \tag{4.126}
\]

The approximation comes from use of the estimates \(\hat{\beta}_1, \hat{\sigma}_{ea}^2\) as if they were \(\beta_1\) and \(\sigma_{ea}^2\).

There is potential for obtaining estimates that are outside of the parameter space for \(\sigma_{V_a}^2\), such as \(0 \in (1 - \alpha)100\%)\ CI(\sigma_{V_a}^2)\) or even just negative values. In such a case the interpretation is that the true variance \(\sigma_{V_a}^2\) is zero. If \(\hat{\sigma}_{V_a}^2 = 0 \forall a\) we can conclude the study material is homogeneous.

As elsewhere a much better option which accounts for all of the uncertainty is to use the bootstrap techniques. In order to do so, we need to modify the bootstrap
cycle to include per replicate resampling of the study data \( W_{a,mb} \). We can modify the last two steps within the cycle in algorithm 5 as

**Algorithm 6** Bootstrap algorithm for corrected value and heterogeneity variance in the study data

```plaintext
for b = 1 to B (large) do
  ... (same as algorithm 5 removing the last two steps) ...
  Generate \( W_{a,mb} \sim N(\bar{W}_a, s^2_{W_a}) \equiv N(\hat{\beta}_0 + \hat{\beta}_1 \hat{d}_a, s^2_{W_a}), a = 1, \ldots, A, m = 1, \ldots, M_a, \)
  \( \bar{W}_{a,b} = \frac{1}{M_a} \sum_m W_{a,mb}, \)
  \( s^2_{W_{ab}} = \frac{1}{M_a-1} \sum_m (W_{a,mb} - \bar{W}_{a,b})^2, \)
  \( \hat{\sigma}^2_{V_{ab}} = \frac{s^2_{W_{ab}} - \hat{\beta}^2_{ib}}{\hat{\beta}^2_{ib}} \)
  Apply (4.113) to obtain \( \hat{d}_{a,b} = \frac{W_{a,b} - \hat{\beta}_{0,b}}{\hat{\beta}_{1,b}}, a = 1, \ldots, A \)
end for
```

Finally we compute the sets \( \mathcal{V}_a = \{\hat{\sigma}^2_{V_{ab}} : b = 1, \ldots, B\} \) and we proceed to obtain the percentile CI as described before. If any of the endpoints of the confidence interval is negative it must be set to zero.

### 4.2.2.4 Inferences Under Correlated Errors

Now assume that the error terms are correlated for the method while measuring the control sample and the study sample, \( \text{Corr}[\epsilon_{ak}, \delta_{ak}] = \rho_a \neq 0 \). Define the sample statistics \( \bar{Y}_a, s^2_{Y_a}, \bar{W}_a, s^2_{W_a} \) as above, and define the sample covariance as \( s_{YW_a} = \frac{1}{K_a-1} \sum_{k=1}^{K_a} (Y_{ak} W_{ak} - \bar{Y}_a \bar{W}_a)^2. \)

By using the method of moments it is straightforward to shown that

\[
E[\bar{W}_a] = \beta_0 + \beta_1 d_a, \quad E[s^2_{W_a}] = \beta_1^2 \sigma^2_{V,a} + \sigma^2_{\epsilon_a}\]

and \( E[s_{YW_a}] = \rho_a \sigma_{\epsilon a} \sigma_{\delta a} = \rho_a \sigma^2_{\epsilon a}. \)

The last equality is due to the assumption of equal variances \( \sigma^2_{\delta a} = \sigma^2_{\epsilon a} \). Hence the point estimators defined as (4.113) and (4.114) are still adequate.

However \( V[\hat{d}_a] \) changes due to a covariance term, by using the delta method this can be approximated by
\[ V[\hat{d}_a] \approx \frac{1}{\beta_1^2} \left( V[\bar{W}_a] + V[\hat{\beta}_0] - 2 \text{Cov}[\bar{W}_a, \hat{\beta}_0] + \hat{d}_a^2 \text{Cov}[\hat{\beta}_1] \right) - 2 \hat{d}_a \text{Cov}[\bar{W}_a, \hat{\beta}_1] + 2 \hat{d}_a \text{Cov}[\hat{\beta}_0, \hat{\beta}_1] \] \quad (4.127)

where the terms \( \text{Cov}[\bar{W}_a, \hat{\beta}_0] \) and \( \text{Cov}[\bar{W}_a, \hat{\beta}_1] \) depend on the specific estimators \( \hat{\beta}_0 \) and \( \hat{\beta}_1 \). If we assume the estimators \( \hat{\beta}_0 \) and \( \hat{\beta}_1 \) are as described by (4.60) then

\[
\text{Cov} \left[ \bar{W}_a, \hat{\beta}_0 \right] = \text{Cov} \left[ \bar{W}_a, \bar{Y} - \hat{\beta}_1 \bar{x} \right] = \frac{1}{A} \text{Cov}[\bar{W}_a, \bar{Y}] - \text{Cov} \left[ \bar{W}_a, \hat{\beta}_1 \bar{x} \right] = \frac{1}{A} \text{Cov}[\bar{W}_a, \bar{Y}] + E[\bar{W}_a] E \left[ \hat{\beta}_1 \bar{x} \right] - E \left[ \bar{W}_a \hat{\beta}_1 \bar{x} \right] \quad (4.128)
\]
\[ E \left[ W_a \hat{\beta}_1 \bar{x} \right] = E \left[ \bar{W}_a \frac{S_{XY}}{S_X^2 - \sigma_X^2} \bar{x} \right] \]

\[ = E \left[ \bar{W}_a \frac{S_{XY}}{S_X^2 - \sigma_X^2} \frac{1}{A - 1} \sum_{a' = 1}^{A} (\hat{x}_{a'} - \bar{x})(\bar{Y}_{a'} - \bar{Y}) \right] \]

\[ = \frac{1}{A - 1} \sum_{a' = 1}^{A} E \left[ \bar{W}_a \frac{S_{XY}}{S_X^2 - \sigma_X^2} (\hat{x}_{a'} - \bar{x}) (\bar{Y}_{a'} - \bar{Y}) \right] \]

\[ = \frac{1}{A - 1} \sum_{a' = 1}^{A} E \left[ \frac{S_{XY}}{S_X^2 - \sigma_X^2} (\hat{x}_{a'} - \bar{x}) \right] E[\bar{W}_a (\bar{Y}_{a'} - \bar{Y})] \]

\[ + \frac{1}{A - 1} \sum_{a' = 1}^{A} E \left[ \frac{S_{XY}}{S_X^2 - \sigma_X^2} (\hat{x}_{a'} - \bar{x}) \right] Cov[\bar{W}_a, \bar{Y}_{a'} - \bar{Y}] \]

\[ = E[\bar{W}_a] E \left[ \frac{S_{XY}}{S_X^2 - \sigma_X^2} \frac{1}{A - 1} \sum_{a' = 1}^{A} (\hat{x}_{a'} - \bar{x})(\bar{Y}_{a'} - \bar{Y}) \right] \]

\[ + \frac{1}{A - 1} E \left[ \frac{S_{XY}}{S_X^2 - \sigma_X^2} (\hat{x}_a - \bar{x}) \right] \left( Cov[\bar{W}_a, \bar{Y}_a] - \frac{Cov[\bar{W}_a, \bar{Y}_a]}{A} \right) \]

\[ - \frac{1}{A - 1} \sum_{a' \neq a} E \left[ \frac{S_{XY}}{S_X^2 - \sigma_X^2} (\hat{x}_{a'} - \bar{x}) \right] \frac{Cov[\bar{W}_a, \bar{Y}_a]}{A} \]

\[ = E[\bar{W}_a] E \left[ \hat{\beta}_1 \bar{x} \right] + \frac{Cov[\bar{W}_a, \bar{Y}_a]}{A - 1} E \left[ \frac{S_{XY}}{S_X^2 - \sigma_X^2} (\hat{x}_a - \bar{x}) \right] \]

\[ - \frac{Cov[\bar{W}_a, \bar{Y}_a]}{(A - 1)A} E \left[ \frac{S_{XY}}{S_X^2 - \sigma_X^2} \sum_{a'}^{A} (\hat{x}_{a'} - \bar{x}) \right]. \]

Hence

\[ Cov \left[ W_a, \hat{\beta}_0 \right] = Cov[\bar{W}_a, Y - \hat{\beta}_1 \bar{x}] \]

\[ = \frac{1}{A} Cov[\bar{W}_a, \bar{Y}_a] - \frac{Cov[\bar{W}_a, \bar{Y}_a]}{A - 1} E \left[ \frac{S_{XY}}{S_X^2 - \sigma_X^2} (\hat{x}_a - \bar{x}) \right] \]  \hspace{1cm} (4.130)

Similarly
\[ \text{Cov} \left[ \hat{W}_a, \hat{\beta}_1 \right] = \frac{\text{Cov}[\hat{W}_a, \hat{Y}_a]}{A - 1} E \left[ \frac{\left( \hat{x}_a - \bar{x} \right)}{S_X^2 - \sigma_\hat{x}^2} \right] \]  

(4.131)

The variance of the adjusted study value can be estimated as

\[ \hat{V}[\hat{d}_a] = \frac{1}{\hat{\beta}_1^2} \left( s_{Wa}^2 + \hat{\sigma}_{W0}^2 - 2\hat{\sigma}_{W\beta_0} + \hat{d}_a^2 \hat{\sigma}_{\beta_1}^2 - 2\hat{d}_a s_{Wa,\beta_1} + 2\hat{d}_a \hat{\sigma}_{\beta_0\beta_1} \right), \]  

(4.132)

where

\[
\hat{\sigma}_{W0} = \hat{\text{Cov}} \left[ \hat{W}_a, \hat{\beta}_0 \right] = \frac{1}{A} s_{WY_a} - \hat{\sigma}_{Wa,\beta_1} \bar{x},
\]

\[
\hat{\sigma}_{W1} = \hat{\text{Cov}} \left[ \hat{W}_a, \hat{\beta}_1 \right] = \frac{(\hat{x}_a - \bar{x})}{(A - 1)(S_X^2 - \sigma_\hat{x}^2)} s_{WY_a}.
\]  

(4.133)

Some additional distributional assumptions on the error terms are required in order to get confidence intervals for \( d_a \) and \( \sigma_{V_a}^2 \). Assuming normality of \( \bar{W}_a - \hat{\beta}_0 - \hat{\beta}_1 d_a \) the distribution and confidence interval of \( d_a \) can be approximated as

\[
\frac{\bar{W}_a - \hat{\beta}_0 - \hat{\beta}_1 d_a}{\sqrt{\hat{V}[\bar{W}_a - \hat{\beta}_0 - \hat{\beta}_1 d_a]}} \sim T(\nu_{da}),
\]  

(4.134)

where

\[ \nu_{da} \approx \frac{2 \left( \hat{V}[\hat{d}_a] \right)^2}{\hat{V}[\hat{d}_a]} \]  

(4.135)

The procedure to obtain \( \hat{V}[\hat{d}_a] \) is lengthy and tedious but it is based on well known theory (Anderson (1984) [2]), assuming normality of \( U_{ak}, \epsilon_{ak}, V_{ak} \) and \( \delta_{ak} \):

\[
\hat{V} \left[ \hat{V}[\hat{d}_a] \right] = \frac{2s_{Wa}^4}{K_a - 1} + \frac{2\hat{d}_a^4}{\nu_{\beta_1}} + \frac{4\hat{d}_a^2 s_{W}^2 S_Y^2}{\bar{x}_a^2 (K_a - 1)} + \frac{8\hat{d}_a^2 S_Y^2}{\bar{x}_a (K_a - 1)}
\]

\[ - \frac{8\hat{d}_a s_{W}^2 S_Y}{\bar{x}_a (K_a - 1)} - \frac{8\hat{d}_a^3 S_Y^2}{\bar{x}_a^3 (K_a - 1)}, \]  

(4.136)

A \((1 - \alpha)100\%\) CI for \( d_a \) is approximately
\[
\left( \hat{d}_a - T_{(\alpha/2,\nu_{da})} \sqrt{V[\hat{d}_a]}, \hat{d}_a + T_{(\alpha/2,\nu_{da})} \sqrt{V[\hat{d}_a]} \right).
\]  \hspace{1cm} (4.137)

**Fieller’s method**

Applying the Fieller’s method, equations similar to (4.123), replacing the conditions in (4.122) to include the correlation adjustments as

\[
P(q(d_a) \leq 0) = 1 - \alpha,
\]

\[
q(d) = f_2 d^2 - 2 f_1 d + f_0 \text{ with}
\]

\[
f_0 = (\hat{W}_a - \hat{\beta}_0)^2 - t^2_{(1-\alpha/2,\nu_{da})} (\hat{\sigma}_{W_a}^2 + \hat{\sigma}_{\beta_0}^2 - \hat{\sigma}_{W_a,\beta_0}^2),
\]

\[
f_1 = \hat{\beta}_1 (\hat{W}_a - \hat{\beta}_0) + t^2_{(1-\alpha/2,\nu_{da})} (\hat{\sigma}_{\beta_0,\beta_1} - \hat{\sigma}_{W_a,\beta_1})
\]

\[
f_2 = \hat{\beta}_2^2 - t^2_{(1-\alpha/2,\nu_{da})} \hat{\sigma}_{\beta_1}^2
\]

and using the approximate degrees of freedom under correlation in (4.135) and (4.136).

**Parametric Bootstrap method**

Applying the parametric bootstrap method, we need to extend algorithm 5 in the way the error terms are generated. This modification is described in algorithm 7.

Then compute the sets \( \mathcal{D}_a = \{ \hat{d}_{ab} : b = 1, \ldots, B \} \). A \((1 - \alpha)\)100% CI for \( d_a \) is

\[
(Q_{(\alpha/2)}(\mathcal{D}_a), Q_{(1-\alpha/2)}(\mathcal{D}_a)),
\]

\hspace{1cm} (4.139)

where \( Q \) is as in Section 4.2.1.2.

Hence (4.124) is still valid and it can be used to estimate the bias of \( \hat{d}_a \).

These confidence intervals can contain zero or even contain only negative values. In such a case the interpretation is that the true value of the compound in the study sample is below the detection limit of the analytical method, i.e., it becomes a censored data point.

As with the point estimators, the distribution and approximate confidence interval of \( \sigma_{V_a}^2 \) remain as stated by (4.125) and (4.126).
Algorithm 7 Bootstrap algorithm for corrected value and heterogeneity variance in the study data under correlated errors

for $b = 1$ to $B \text{(large)}$ do

\[ \hat{x}_{ab} \leftarrow \hat{x}_a + \hat{\sigma}_a Z_{ab}, a = 1, \ldots, A, \]

\[ \hat{\sigma}_{xab}^2 \leftarrow \hat{\sigma}_{xa}^2 \chi^2(\nu_{xa})/\nu_{xa}, a = 1, \ldots, A, \]

\[
\begin{bmatrix}
    \epsilon_{a,kb} \\
    \delta_{a,kb}
\end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \hat{\sigma}_{\epsilon a}^2 & \hat{\sigma}_{\delta a} \\ \hat{\sigma}_{\delta a} & \hat{\sigma}_{\delta a}^2 \end{bmatrix}\right),
\]

\[ \hat{Y}_{a,kb} \leftarrow \hat{\beta}_0 + \hat{\beta}_1 \hat{x}_a + \hat{\beta}_1 \hat{\sigma}_{Ua} Z_{a2,kb} + \epsilon_{a,kb}, a = 1, \ldots, A, k = 1, \ldots, K_a, \]

Apply (4.60) to $(\hat{x}_{ab}, \hat{Y}_{ab})$ to get $(\hat{\beta}_{0,b}, \hat{\beta}_{1,b})$

Apply (4.69) or (4.72) to obtain $\hat{\sigma}_{\epsilon,b}^2$

Apply (4.77) or (4.80) to obtain modified estimators $(\hat{\beta}_{0,b}, \hat{\beta}_{1,b})$

\[ W_{a,kb} \leftarrow \hat{W}_a + \hat{\beta}_1 \hat{\sigma}_{V_a} Z_{a3,kb} + \delta_{a,kb}, a = 1, \ldots, A, k = 1, \ldots, K_a, \]

\[ W_{a,b} \leftarrow \frac{1}{K_a} \sum_k W_{a,kb}, \]

\[ s_{Wab}^2 \leftarrow \frac{1}{K_a-1} \sum_k (W_{a,kb} - \hat{W}_{a,b})^2, \]

\[ \hat{\sigma}_{Vab}^2 \leftarrow \frac{s_{Wab}^2 - \hat{\sigma}_{\epsilon,b}^2}{\beta_{1,b}^2} \]

Apply (4.113) to obtain $\hat{d}_{a,b} = \frac{W_{a,b} - \hat{\beta}_{0,b}}{\beta_{1,b}}, a = 1, \ldots, A$

end for

If $\sigma_{\delta|D}^2 \neq \sigma_{\epsilon|X}^2$, as may be the case if the method’s variance depends on the concentration of the compound and the concentration of the compound in the control material and the study material are very different, then $\sigma_{\delta|D}^2$ and $\sigma_{\epsilon|X}^2$ should be modeled accordingly. In Chapter 5 we briefly review a variance model proportional to the square of the concentration.

4.2.2.5 Example

Table 4.28 contains the estimates of the adjusted study values and their variances by using the data in Table 3.1 and (4.24), (4.120) and (4.123), under the assumption of independent errors and equal within method variance for each compound across samples.

Table 4.29 contains the estimates of the adjusted study values and their variances and confidence intervals under the assumption of correlated errors and equal within method variance for each compound across samples.
In Table 4.28 it can be observed that the corrected values $\hat{d}_a$ change more or less even around the observed values $\hat{W}_a$ and their magnitudes are very similar. Recalling the results from Section 3.3.3.2 the proportional bias model produces confidence interval estimates within reasonable values in contrast to the constant bias model that can produce confidence interval estimates outside the parameter space (negative concentration levels) or in the best case a censored value although the observed value has a high likelihood of being positively present.

However the corrected variance estimate $\hat{\sigma}_{da}^2$ is larger than the observed variance $s_{Wa}^2$ in all the cases, this is mainly due to the pooled additional information with large variance. Secondarily, this can be due to an underestimation of the observed variance. About half of the corrected confidence intervals appear to be narrower than their corresponding confidence intervals based on the study data only ($\hat{W}_a \pm t_{(\nu, 1 - \alpha/2)} s_{Wa}/\sqrt{M_a}$), while the rest of the estimated CIs are wider. Under the assumption of unbiasedness the uncorrected estimates would lead to incorrect optimistic inferences of the true value. For this specific example, less than 47% of the observed confidence intervals contain the corrected estimates both under independent error terms and under correlated error terms.

When comparing the delta method estimates against the Fieller’s method and bootstrap estimates, the remarkable distinction is the symmetry of the delta method intervals against the asymmetric intervals of the bootstrap and Fieller’s intervals both probably following the skewness of the underlying distribution. As a consequence, the delta method intervals often overlap with the bootstrap and Fieller’s intervals on the shortest tail side of the re-sampled distribution. A second distinction is that in general the delta method intervals are shorter than the bootstrap and Fieller’s intervals, and is also related to the skewness of the re-sampled values.

Comparing the estimates on Tables 4.28 and 4.29, it can be observed that in general the estimated bootstrap confidence intervals under correlation of the error
terms are narrower when the correlation is positive and larger when the correlation is negative in comparison with the confidence intervals under independence of the error terms. The delta method and Fieller’s method estimates produce narrower CIs. One compound (triphenylene) shows a larger confidence interval under correlated errors while having positive correlation. For an easy comparison, the correlation coefficient is shown in Table 4.29, originally in Table 3.1.

Table 4.30 contains the estimates of the within material variance for the study data by using (4.114) and (4.126), while assuming the measurement error variance for each compound is the same across samples. The estimates of the within material variance are those on Table 3.5 scaled by the factor $1/\hat{\beta}_1^2$. The estimated probability of obtaining negative estimates is the same.
Table 4.28. Example of study values and variances corrected for linear bias using control data when errors are independent.

<table>
<thead>
<tr>
<th>Compound</th>
<th>$\bar{W}$</th>
<th>lcl$_W$</th>
<th>ucl$_W$</th>
<th>$\hat{d}$</th>
<th>$\hat{\sigma}^2_{\hat{d}DM}$</th>
<th>lcl$_{DM}$</th>
<th>ucl$_{DM}$</th>
<th>lcl$_F$</th>
<th>ucl$_F$</th>
<th>$\hat{bias}_B(\hat{d})$</th>
<th>$\hat{\sigma}_B$</th>
<th>lcl$_B$</th>
<th>ucl$_B$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2,3,5-trimethylnaphthalene</td>
<td>25.7</td>
<td>24.3</td>
<td>27.1</td>
<td>27.3</td>
<td>1.32</td>
<td>24.7</td>
<td>29.9</td>
<td>24.2</td>
<td>30.3</td>
<td>-0.0365</td>
<td>1.36</td>
<td>24.6</td>
<td>29.9</td>
</tr>
<tr>
<td>acenaphthene</td>
<td>34.7</td>
<td>33.6</td>
<td>35.8</td>
<td>36.1</td>
<td>1.25</td>
<td>33.7</td>
<td>38.6</td>
<td>33.2</td>
<td>38.9</td>
<td>-0.012</td>
<td>1.26</td>
<td>33.6</td>
<td>38.6</td>
</tr>
<tr>
<td>dibenz[a,h]anthracene</td>
<td>112</td>
<td>102</td>
<td>122</td>
<td>112</td>
<td>2.55</td>
<td>107</td>
<td>117</td>
<td>108</td>
<td>115</td>
<td>-0.0105</td>
<td>2.55</td>
<td>107</td>
<td>117</td>
</tr>
<tr>
<td>acenaphthylene</td>
<td>29.6</td>
<td>26.3</td>
<td>33</td>
<td>31.2</td>
<td>1.47</td>
<td>28.3</td>
<td>34</td>
<td>27.9</td>
<td>34.3</td>
<td>-0.0296</td>
<td>1.5</td>
<td>28.2</td>
<td>34</td>
</tr>
<tr>
<td>biphenyl</td>
<td>29.2</td>
<td>27.5</td>
<td>31</td>
<td>30.8</td>
<td>1.32</td>
<td>28.2</td>
<td>33.4</td>
<td>27.7</td>
<td>33.7</td>
<td>0.00545</td>
<td>1.35</td>
<td>28.2</td>
<td>33.3</td>
</tr>
<tr>
<td>1-methylphenanthrene</td>
<td>69.4</td>
<td>67.8</td>
<td>71</td>
<td>70.1</td>
<td>1.1</td>
<td>67.9</td>
<td>72.2</td>
<td>67.5</td>
<td>72.6</td>
<td>-0.0254</td>
<td>1.12</td>
<td>67.7</td>
<td>72.1</td>
</tr>
<tr>
<td>2,6-dimethylnaphthalene</td>
<td>35.2</td>
<td>34.1</td>
<td>36.3</td>
<td>36.6</td>
<td>1.25</td>
<td>34.2</td>
<td>39</td>
<td>33.7</td>
<td>39.4</td>
<td>0.000454</td>
<td>1.28</td>
<td>34.2</td>
<td>39.1</td>
</tr>
<tr>
<td>fluorene</td>
<td>65.5</td>
<td>63.8</td>
<td>67.2</td>
<td>66.2</td>
<td>1.12</td>
<td>64</td>
<td>68.5</td>
<td>63.6</td>
<td>68.8</td>
<td>-0.00412</td>
<td>1.17</td>
<td>63.9</td>
<td>68.5</td>
</tr>
<tr>
<td>triphenylene</td>
<td>141</td>
<td>120</td>
<td>162</td>
<td>140</td>
<td>4.92</td>
<td>130</td>
<td>150</td>
<td>136</td>
<td>145</td>
<td>0.0864</td>
<td>4.76</td>
<td>131</td>
<td>150</td>
</tr>
<tr>
<td>1-methylphenanthalene</td>
<td>49</td>
<td>46.7</td>
<td>51.2</td>
<td>50.1</td>
<td>1.24</td>
<td>47.6</td>
<td>52.5</td>
<td>47.2</td>
<td>52.9</td>
<td>-0.00791</td>
<td>1.29</td>
<td>47.4</td>
<td>52.4</td>
</tr>
<tr>
<td>anthracene</td>
<td>172</td>
<td>166</td>
<td>179</td>
<td>171</td>
<td>1.9</td>
<td>167</td>
<td>174</td>
<td>167</td>
<td>174</td>
<td>0.0759</td>
<td>1.93</td>
<td>167</td>
<td>174</td>
</tr>
<tr>
<td>benzo[j]fluoranthene</td>
<td>281</td>
<td>260</td>
<td>302</td>
<td>277</td>
<td>5.29</td>
<td>267</td>
<td>287</td>
<td>271</td>
<td>283</td>
<td>-0.00632</td>
<td>5.16</td>
<td>268</td>
<td>287</td>
</tr>
<tr>
<td>benzo[k]fluoranthene</td>
<td>406</td>
<td>384</td>
<td>428</td>
<td>399</td>
<td>6.05</td>
<td>388</td>
<td>411</td>
<td>391</td>
<td>408</td>
<td>0.307</td>
<td>5.85</td>
<td>388</td>
<td>411</td>
</tr>
<tr>
<td>2-methylnaphthalene</td>
<td>79.2</td>
<td>77.2</td>
<td>81.2</td>
<td>79.7</td>
<td>1.11</td>
<td>77.5</td>
<td>81.9</td>
<td>77.1</td>
<td>82.2</td>
<td>-0.00159</td>
<td>1.18</td>
<td>77.2</td>
<td>81.9</td>
</tr>
<tr>
<td>chrysene</td>
<td>547</td>
<td>538</td>
<td>556</td>
<td>537</td>
<td>5.47</td>
<td>527</td>
<td>548</td>
<td>526</td>
<td>549</td>
<td>0.0478</td>
<td>5.45</td>
<td>526</td>
<td>548</td>
</tr>
<tr>
<td>benzo[e]pyrene</td>
<td>590</td>
<td>560</td>
<td>620</td>
<td>579</td>
<td>8.73</td>
<td>562</td>
<td>597</td>
<td>567</td>
<td>593</td>
<td>-0.0641</td>
<td>8.37</td>
<td>563</td>
<td>595</td>
</tr>
<tr>
<td>benzo[a]anthracene</td>
<td>378</td>
<td>368</td>
<td>388</td>
<td>372</td>
<td>3.94</td>
<td>364</td>
<td>380</td>
<td>365</td>
<td>380</td>
<td>0.0838</td>
<td>3.98</td>
<td>364</td>
<td>379</td>
</tr>
<tr>
<td>benzo[ghi]perylene</td>
<td>488</td>
<td>476</td>
<td>500</td>
<td>480</td>
<td>5.22</td>
<td>469</td>
<td>490</td>
<td>470</td>
<td>490</td>
<td>0.131</td>
<td>5.12</td>
<td>470</td>
<td>489</td>
</tr>
<tr>
<td>indeno[1,2,3-cd]pyrene</td>
<td>608</td>
<td>588</td>
<td>628</td>
<td>597</td>
<td>7.25</td>
<td>583</td>
<td>611</td>
<td>584</td>
<td>611</td>
<td>0.305</td>
<td>7.4</td>
<td>582</td>
<td>611</td>
</tr>
<tr>
<td>benzo[a]pyrene</td>
<td>636</td>
<td>622</td>
<td>650</td>
<td>624</td>
<td>6.81</td>
<td>611</td>
<td>638</td>
<td>611</td>
<td>639</td>
<td>0.295</td>
<td>6.77</td>
<td>611</td>
<td>638</td>
</tr>
<tr>
<td>perylene</td>
<td>184</td>
<td>167</td>
<td>201</td>
<td>182</td>
<td>4.18</td>
<td>174</td>
<td>190</td>
<td>178</td>
<td>187</td>
<td>-0.115</td>
<td>4.13</td>
<td>174</td>
<td>190</td>
</tr>
<tr>
<td>phenanthrene</td>
<td>539</td>
<td>522</td>
<td>556</td>
<td>530</td>
<td>6.33</td>
<td>517</td>
<td>542</td>
<td>518</td>
<td>541</td>
<td>0.392</td>
<td>6.28</td>
<td>518</td>
<td>542</td>
</tr>
<tr>
<td>benzo[b]fluoranthene</td>
<td>1090</td>
<td>1040</td>
<td>1140</td>
<td>1070</td>
<td>15.6</td>
<td>1040</td>
<td>1100</td>
<td>1040</td>
<td>1090</td>
<td>0.287</td>
<td>15.6</td>
<td>1040</td>
<td>1100</td>
</tr>
<tr>
<td>pyrene</td>
<td>1130</td>
<td>1090</td>
<td>1160</td>
<td>1110</td>
<td>13.8</td>
<td>1080</td>
<td>1130</td>
<td>1080</td>
<td>1130</td>
<td>0.492</td>
<td>14.1</td>
<td>1080</td>
<td>1130</td>
</tr>
<tr>
<td>fluoranthene</td>
<td>1090</td>
<td>1070</td>
<td>1120</td>
<td>1070</td>
<td>12.3</td>
<td>1050</td>
<td>1100</td>
<td>1050</td>
<td>1100</td>
<td>0.278</td>
<td>12.2</td>
<td>1050</td>
<td>1100</td>
</tr>
<tr>
<td>naphthalene</td>
<td>121</td>
<td>115</td>
<td>128</td>
<td>121</td>
<td>1.74</td>
<td>117</td>
<td>124</td>
<td>118</td>
<td>124</td>
<td>-0.0232</td>
<td>1.83</td>
<td>117</td>
<td>125</td>
</tr>
</tbody>
</table>

For each compound, $\bar{W}$ = the observed mean study value, $(lcl_W, ucl_W)$ = a 95% confidence interval for the observed mean study value, $\hat{d}$ = the corrected mean of the study value under linear bias, $\hat{\sigma}^2_{\hat{d}DM}$ = the standard deviation of the corrected study value by the Delta Method, $(lcl_{DM}, ucl_{DM})$ = a 95% confidence interval for the corrected study values using the Delta Method and assuming large degrees of freedom, $(lcl_F, ucl_F)$ = a 95% confidence interval for the corrected study value using the Fieller’s Method and assuming large degrees of freedom. Using parametric bootstrap: $\hat{bias}_B(\hat{d})$ = the estimated bias of the corrected mean of the study value, $\hat{\sigma}_d$ = the standard deviation of the corrected study value, $(lcl_B, ucl_B)$ = the 95% confidence interval. All under the assumption of independent errors.
Table 4.29. Example of study values and variances corrected for linear bias using control data when errors are correlated.

<table>
<thead>
<tr>
<th>Compound</th>
<th>$\hat{\rho}_{WY}$</th>
<th>$\hat{\sigma}_{DM}$</th>
<th>$lcl_{DM}$</th>
<th>$ucl_{DM}$</th>
<th>$lcl_F$</th>
<th>$ucl_F$</th>
<th>$\hat{\text{bias}}_{B}(d)$</th>
<th>$\hat{\sigma}_{dB}$</th>
<th>$lcl_B$</th>
<th>$ucl_B$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2,3,5-trimethylnaphthalene</td>
<td>0.69</td>
<td>1.22</td>
<td>24.9</td>
<td>29.7</td>
<td>24.4</td>
<td>30.1</td>
<td>0.0252</td>
<td>1.31</td>
<td>24.7</td>
<td>29.8</td>
</tr>
<tr>
<td>acenaphthene</td>
<td>-0.885</td>
<td>1.19</td>
<td>33.8</td>
<td>38.4</td>
<td>33.4</td>
<td>38.8</td>
<td>0.0263</td>
<td>1.3</td>
<td>33.6</td>
<td>38.7</td>
</tr>
<tr>
<td>dibenz[a,h]anthracene</td>
<td>0.277</td>
<td>2.52</td>
<td>107</td>
<td>117</td>
<td>108</td>
<td>115</td>
<td>0.000166</td>
<td>2.66</td>
<td>107</td>
<td>118</td>
</tr>
<tr>
<td>acenaphthylene</td>
<td>-0.517</td>
<td>1.42</td>
<td>28.4</td>
<td>33.9</td>
<td>28.1</td>
<td>34.2</td>
<td>0.00234</td>
<td>1.53</td>
<td>28.1</td>
<td>34.2</td>
</tr>
<tr>
<td>biphenyl</td>
<td>-1</td>
<td>1.27</td>
<td>28.3</td>
<td>33.2</td>
<td>27.9</td>
<td>33.6</td>
<td>0.049</td>
<td>1.35</td>
<td>28.1</td>
<td>33.4</td>
</tr>
<tr>
<td>1-methylphenanthrene</td>
<td>0.916</td>
<td>1.02</td>
<td>68.1</td>
<td>72.1</td>
<td>67.7</td>
<td>72.4</td>
<td>0.0184</td>
<td>1.12</td>
<td>67.8</td>
<td>72.3</td>
</tr>
<tr>
<td>2,6-dimethylnaphthalene</td>
<td>0.525</td>
<td>1.15</td>
<td>34.3</td>
<td>38.9</td>
<td>33.9</td>
<td>39.2</td>
<td>0.0243</td>
<td>1.25</td>
<td>34.2</td>
<td>39.3</td>
</tr>
<tr>
<td>fluorene</td>
<td>-0.947</td>
<td>1.11</td>
<td>64.1</td>
<td>68.4</td>
<td>63.7</td>
<td>68.7</td>
<td>0.0247</td>
<td>1.16</td>
<td>64</td>
<td>68.6</td>
</tr>
<tr>
<td>triphenylene</td>
<td>0.489</td>
<td>4.86</td>
<td>131</td>
<td>150</td>
<td>136</td>
<td>144</td>
<td>0.162</td>
<td>4.9</td>
<td>130</td>
<td>150</td>
</tr>
<tr>
<td>1-methylphenanthalene</td>
<td>-0.426</td>
<td>1.22</td>
<td>47.7</td>
<td>52.5</td>
<td>47.3</td>
<td>52.7</td>
<td>0.0389</td>
<td>1.27</td>
<td>47.7</td>
<td>52.6</td>
</tr>
<tr>
<td>anthracene</td>
<td>1</td>
<td>1.7</td>
<td>167</td>
<td>174</td>
<td>168</td>
<td>174</td>
<td>0.111</td>
<td>2</td>
<td>167</td>
<td>175</td>
</tr>
<tr>
<td>benzo[j]fluoranthene</td>
<td>-0.0128</td>
<td>5.21</td>
<td>267</td>
<td>287</td>
<td>271</td>
<td>283</td>
<td>0.376</td>
<td>5.59</td>
<td>267</td>
<td>288</td>
</tr>
<tr>
<td>benzo[k]fluoranthene</td>
<td>-0.866</td>
<td>5.93</td>
<td>388</td>
<td>411</td>
<td>391</td>
<td>408</td>
<td>0.0732</td>
<td>6.42</td>
<td>387</td>
<td>412</td>
</tr>
<tr>
<td>2-methylnaphthalene</td>
<td>-0.567</td>
<td>1.1</td>
<td>77.5</td>
<td>81.9</td>
<td>77.2</td>
<td>82.1</td>
<td>0.0405</td>
<td>1.14</td>
<td>77.5</td>
<td>82</td>
</tr>
<tr>
<td>chrysene</td>
<td>-0.968</td>
<td>5.16</td>
<td>527</td>
<td>547</td>
<td>527</td>
<td>548</td>
<td>0.223</td>
<td>6.17</td>
<td>525</td>
<td>550</td>
</tr>
<tr>
<td>benzo[e]pyrene</td>
<td>0.921</td>
<td>8.26</td>
<td>563</td>
<td>596</td>
<td>568</td>
<td>591</td>
<td>0.375</td>
<td>9</td>
<td>562</td>
<td>598</td>
</tr>
<tr>
<td>benz[a]anthracene</td>
<td>0.961</td>
<td>3.55</td>
<td>365</td>
<td>379</td>
<td>365</td>
<td>379</td>
<td>0.174</td>
<td>4.07</td>
<td>364</td>
<td>381</td>
</tr>
<tr>
<td>benzo[ghi]perylen</td>
<td>0.454</td>
<td>4.84</td>
<td>470</td>
<td>489</td>
<td>471</td>
<td>489</td>
<td>0.13</td>
<td>5.58</td>
<td>469</td>
<td>490</td>
</tr>
<tr>
<td>indeno[1,2,3-cd]pyrene</td>
<td>-0.982</td>
<td>7.01</td>
<td>583</td>
<td>611</td>
<td>585</td>
<td>610</td>
<td>0.328</td>
<td>7.84</td>
<td>582</td>
<td>613</td>
</tr>
<tr>
<td>benzo[a]pyrene</td>
<td>0.99</td>
<td>6.21</td>
<td>612</td>
<td>637</td>
<td>612</td>
<td>637</td>
<td>0.124</td>
<td>6.89</td>
<td>612</td>
<td>638</td>
</tr>
<tr>
<td>perylene</td>
<td>0.684</td>
<td>4.04</td>
<td>174</td>
<td>190</td>
<td>178</td>
<td>186</td>
<td>0.248</td>
<td>4.22</td>
<td>174</td>
<td>191</td>
</tr>
<tr>
<td>phenanthrene</td>
<td>-0.763</td>
<td>6.17</td>
<td>517</td>
<td>542</td>
<td>519</td>
<td>541</td>
<td>-0.083</td>
<td>6.74</td>
<td>516</td>
<td>543</td>
</tr>
<tr>
<td>benzo[b]fluoranthene</td>
<td>-0.607</td>
<td>15.1</td>
<td>1040</td>
<td>1100</td>
<td>1040</td>
<td>1090</td>
<td>0.932</td>
<td>17</td>
<td>1040</td>
<td>1100</td>
</tr>
<tr>
<td>pyrene</td>
<td>-0.652</td>
<td>13.6</td>
<td>1080</td>
<td>1130</td>
<td>1080</td>
<td>1130</td>
<td>0.273</td>
<td>15.1</td>
<td>1080</td>
<td>1140</td>
</tr>
<tr>
<td>fluoranthene</td>
<td>-0.382</td>
<td>12.5</td>
<td>1050</td>
<td>1100</td>
<td>1050</td>
<td>1100</td>
<td>0.611</td>
<td>13.3</td>
<td>1050</td>
<td>1100</td>
</tr>
<tr>
<td>naphthalene</td>
<td>-0.262</td>
<td>1.66</td>
<td>118</td>
<td>124</td>
<td>118</td>
<td>124</td>
<td>0.158</td>
<td>1.81</td>
<td>117</td>
<td>124</td>
</tr>
</tbody>
</table>

For each compound, $\hat{\rho}_{WY}$ = the correlation coefficient between the control and study data, $\hat{\sigma}^2_{DM}$ = the standard deviation of the corrected study value by the Delta Method, $(lcl_{DM}, ucl_{DM})$ = a 95% confidence interval for the corrected study value using the Delta Method assuming large degrees of freedom, $(lcl_F, ucl_F)$ = a 95% confidence interval for the corrected study value using the Fieller’s Method assuming large degrees of freedom. Using parametric bootstrap: $\hat{\text{bias}}_{B}(d)$ = the estimated bias of the corrected mean of the study value, $\hat{\sigma}_{dB}$ = the standard deviation of the corrected study value, $(lcl_B, ucl_B)$ = the 95% confidence interval. All under the assumption of correlated errors.
For each compound, $\hat{\sigma}_V^2$ = the estimated within material (heterogeneity) variance, $(lcl(\sigma_V^2), ucl(\sigma_V^2))$ = a 95% confidence interval for the within material (heterogeneity) variance, $P(\hat{\sigma}_V^2 < 0)$ = the estimated probability of getting a negative estimate of the within material (heterogeneity) variance. All under the assumption of linear bias and constant variance of the error in the equation variance for each compound across samples.

<table>
<thead>
<tr>
<th>Compound</th>
<th>$\hat{\sigma}_V^2$</th>
<th>lcl($\sigma_V^2$)</th>
<th>ucl($\sigma_V^2$)</th>
<th>$Pr(\hat{\sigma}_V^2 &lt; 0)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2,3,5-trimethylnaphthalene</td>
<td>0</td>
<td>0</td>
<td>14.1</td>
<td>0.756</td>
</tr>
<tr>
<td>acenaphthene</td>
<td>0</td>
<td>0</td>
<td>3.97</td>
<td>0.957</td>
</tr>
<tr>
<td>dibenz[a,h]anthracene</td>
<td>14.9</td>
<td>3.74</td>
<td>606</td>
<td>2.22e-16</td>
</tr>
<tr>
<td>acenaphthylene</td>
<td>0</td>
<td>0</td>
<td>60.4</td>
<td>0.649</td>
</tr>
<tr>
<td>biphenyl</td>
<td>0</td>
<td>0</td>
<td>17.6</td>
<td>0.885</td>
</tr>
<tr>
<td>1-methylphenanthrene</td>
<td>0</td>
<td>0</td>
<td>15.6</td>
<td>0.404</td>
</tr>
<tr>
<td>2,6-dimethylnaphthalene</td>
<td>0</td>
<td>0</td>
<td>4.62</td>
<td>0.948</td>
</tr>
<tr>
<td>fluorene</td>
<td>0</td>
<td>0</td>
<td>6.14</td>
<td>0.963</td>
</tr>
<tr>
<td>triphenylene</td>
<td>61.8</td>
<td>11.9</td>
<td>2700</td>
<td>3.25e-05</td>
</tr>
<tr>
<td>1-methylnaphthalene</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.981</td>
</tr>
<tr>
<td>anthracene</td>
<td>0</td>
<td>0</td>
<td>219</td>
<td>0.84</td>
</tr>
<tr>
<td>benzo[j]fluoranthene</td>
<td>67.7</td>
<td>16.7</td>
<td>2760</td>
<td>3.44e-14</td>
</tr>
<tr>
<td>benzo[k]fluoranthene</td>
<td>50.1</td>
<td>0</td>
<td>2940</td>
<td>0.0498</td>
</tr>
<tr>
<td>2-methylnaphthalene</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.995</td>
</tr>
<tr>
<td>chrysene</td>
<td>0</td>
<td>0</td>
<td>356</td>
<td>0.912</td>
</tr>
<tr>
<td>benzo[e]pyrene</td>
<td>108</td>
<td>0</td>
<td>6110</td>
<td>0.0386</td>
</tr>
<tr>
<td>benz[a]anthracene</td>
<td>0</td>
<td>0</td>
<td>577</td>
<td>0.846</td>
</tr>
<tr>
<td>benzo[ghi]pyrene</td>
<td>0</td>
<td>0</td>
<td>842</td>
<td>0.439</td>
</tr>
<tr>
<td>indeno[1,2,3-cd]pyrene</td>
<td>0</td>
<td>0</td>
<td>2180</td>
<td>0.381</td>
</tr>
<tr>
<td>benzo[a]pyrene</td>
<td>6.86</td>
<td>0</td>
<td>1230</td>
<td>0.279</td>
</tr>
<tr>
<td>perylene</td>
<td>0</td>
<td>0</td>
<td>1710</td>
<td>0.588</td>
</tr>
<tr>
<td>phenanthrene</td>
<td>0</td>
<td>0</td>
<td>1610</td>
<td>0.614</td>
</tr>
<tr>
<td>benzo[b]fluoranthene</td>
<td>318</td>
<td>52.7</td>
<td>14300</td>
<td>0.000362</td>
</tr>
<tr>
<td>pyrene</td>
<td>89.4</td>
<td>0</td>
<td>7770</td>
<td>0.163</td>
</tr>
<tr>
<td>fluoranthene</td>
<td>0</td>
<td>0</td>
<td>2660</td>
<td>0.92</td>
</tr>
<tr>
<td>naphthalene</td>
<td>0</td>
<td>0</td>
<td>174</td>
<td>0.908</td>
</tr>
</tbody>
</table>

Table 4.30. Example of study within material variances for proportional bias.
CHAPTER 5
CHANGING MEASUREMENT VARIANCE

5.1 Modeling the variance as a function of concentration

So far we have assumed the variance is constant or that it changes arbitrarily with the compounds or over time. The objective of this section is to analyze the possible functional dependency of the within-laboratory variance with respect to the concentration level or time with the goal of improving the estimates of the true values for the study sample under two scenarios: (i) using the constant bias model and (ii) allowing the bias to change but without assuming anything about how the bias changes with the concentration level.

The changing variance modeling is an important problem worthy of further attention. This section presents some preliminary work of one variance model but not completely accounting for the error in the CRMs.

5.2 Variance proportional to the squared concentration

Assume that data from multiple compounds at one point in time measured by one laboratory using the same method is available. Under the assumption of normality and within-laboratory variance conditionally proportional to the squared level of the true concentration but not necessarily a constant proportion across compounds:

\[
\sigma^2_{\epsilon_a|x_a} = h(X_a, \theta) = \theta_0 a X_a^2
\]  

(5.1)

with \(\theta_0 a \geq 0\), then
\[ \sigma^2_{e_a} = E[\sigma^2_{e_a|X_a}] = E[\theta_0 a X^2_a] = \theta_0 a (x^2_a + \sigma^2_{Ua}) \] (5.2)

or

\[ \theta_0 a = \frac{\sigma^2_{e_a}}{x^2_a + \sigma^2_{Ua}} = \frac{\sigma^2_{Y_a} - \sigma^2_{Ua}}{x^2_a + \sigma^2_{Ua}}. \] (5.3)

### 5.2.1 Parameter Estimation

The parameter \( \theta_0 a \) can be estimated by substituting unbiased estimators in (5.3), given by

\[ \hat{\theta}_0 a = \frac{s^2_{Y_a} - \hat{\sigma}^2_{Ua}}{\hat{x}^2_a - \hat{\sigma}^2_{X_a} + \hat{\sigma}^2_{Ua}}. \] (5.4)

However, this estimator is biased but consistent. We can show this is so by using the independence of control data and the reference value and the Jensen’s inequality (e.g. Casella and Berger (2002) [11], Theorem 4.7.7.).

\[ E[\hat{\theta}_0 a] = E \left[ \frac{s^2_{Y_a} - \hat{\sigma}^2_{Ua}}{\hat{x}^2_a - \hat{\sigma}^2_{X_a} + \hat{\sigma}^2_{Ua}} \right] = E[s^2_{Y_a} - \hat{\sigma}^2_{Ua}] E \left[ \frac{1}{\hat{x}^2_a - \hat{\sigma}^2_{X_a} + \hat{\sigma}^2_{Ua}} \right] \geq \frac{\sigma^2_{Y_a} - \sigma^2_{Ua}}{x^2_a + \sigma^2_{Ua}} = \theta_0 a. \] (5.5)

\[ \lim_{K_a a \nu_a \nu_{Ua} \to \infty} \hat{\theta}_0 a = \lim_{K_a a \nu_a \nu_{Ua} \to \infty} \frac{s^2_{Y_a} - \hat{\sigma}^2_{Ua}}{\hat{x}^2_a - \hat{\sigma}^2_{X_a} + \hat{\sigma}^2_{Ua}} = \frac{\sigma^2_{Y_a} - \sigma^2_{Ua}}{x^2_a + \sigma^2_{Ua}} = \theta_0 a. \]

If the material is homogeneous (\( \sigma^2_{Ua} = 0 \)) and the true values are known then the estimator defined by (5.4) is unbiased for \( \theta_0 a \), since

\[ E[\hat{\theta}_0 a] = E \left[ \frac{s^2_{Y_a}}{x^2_a} \right] = \frac{E[s^2_{Y_a}]}{x^2_a} = \theta_0 a \frac{x^2_a}{x^2_a} = \theta_0 a. \] (5.6)

Also under normality

\[ (K_a - 1) \frac{\hat{\theta}_0 a}{\theta_0 a} = \frac{(K_a - 1)s^2_{Y_a}}{\theta_0 a x^2_a} \sim \chi^2(K_a - 1). \] (5.7)

If the material is heterogeneous or the true values are known with some uncertainty then the (5.7) is only an approximation.
Note that under linear bias model (5.3) becomes

$$\theta_{0a} = \frac{\sigma^2_{Ya} - \beta^2_1 \sigma^2_{Ua}}{\bar{x}^2_a + \sigma^2_{Ua}}$$

which can be estimated as

$$\hat{\theta}_{0a} = \frac{s^2_{Ya} - \hat{\beta}^2_1 \hat{\sigma}^2_{Ua}}{\bar{x}^2_a - \hat{\sigma}^2_{xa} + \hat{\sigma}^2_{Ua}}.$$  

\[5.8\]

\[5.9\]

**Pooling information across similar compounds**

If we can assume the same proportional coefficient is shared across compounds (i.e., the coefficient of variation is constant over the concentration and across compounds), that is $H_0 : \theta_{0a} = \theta_0 \forall a$, then

$$\theta_0 = \frac{1}{A} \sum_a \theta_{0a} = \frac{1}{A} \sum_a \frac{\sigma^2_{Ya} - \sigma^2_{Ua}}{\bar{x}^2_a + \sigma^2_{Ua}}$$

and this can be estimated as

$$\hat{\theta}_0 = \frac{1}{A} \sum_a \hat{\theta}_{0a} = \frac{1}{A} \sum_a \frac{s^2_{Ya} - \hat{\sigma}^2_{Ua}}{\bar{x}^2_a - \hat{\sigma}^2_{xa} + \hat{\sigma}^2_{Ua}}.$$  

\[5.10\]

\[5.11\]

This estimator is also biased but consistent.

When the material is homogeneous ($\sigma^2_{Ua} = 0 \forall a$) and the true values are constant and known ($\sigma^2_{xa} = 0 \forall a$) then the estimator (5.11) is unbiased for $\theta_0$. Also under normality

$$\frac{\nu_{\theta_0} \hat{\theta}_0}{\theta_0} \sim \chi^2(\nu_{\theta_0})$$

where $\nu_{\theta_0} = (1/A \sum_a (K_a - 1)^{-1})^{-1}$ is a Welch-Satterhwaite type degrees of freedom.

5.2.2 **Performance Evaluation**

The performance evaluation is affected by the variance model, since we can pool information among the compounds to obtain estimates.

**Point-wise evaluation:**
If the compounds in the CRM are handled and measured with different methods then it may be questionable to use a shared model both for bias modeling and variance modeling. In this case we can use the $f_{score}$ in (3.22) just adding subscripts $a$ to index by compound.

Envelope evaluation:

If multiple compounds are available from the certification it is possible to obtain the $h$ model for the within-laboratory variance component and make this model available in the certificate. This would allow pooling information across compounds and make it independent of the specific level, sampling effort, and compound. No additional measurements would be required for certification. Assume the CRM provides an estimated within-laboratory variance from the participant in the certification process and its estimated degrees of freedom $(\hat{\theta}_0c, \nu_{\hat{\theta}_0c})$, where the subscript $c$ is used to indicate that it belongs to the CRM. Then under normality and independence of the error terms

$$F = \frac{\sigma_{\epsilon a}^2}{\sigma_{\epsilon ca}^2} = \frac{\theta_0(x_a^2 + \sigma_{Ua}^2)}{\hat{\theta}_0c(x_a^2 + \sigma_{Ua}^2)} = \frac{\theta_0}{\hat{\theta}_0c},$$

(5.13)

$$\hat{\theta} \sim \frac{\theta}{\nu_{\theta}} \chi^2_{\nu_{\theta}},$$

(5.14)

and

$$\hat{\theta}_0c \sim \frac{\theta_{0c}}{\nu_{\theta_{0c}}} \chi^2_{\nu_{\theta_{0c}}}.$$

(5.15)

Note that (5.13) reduces since the measures are made on the same material at different times: one at certification time and the other at intercomparison time.

Hence using (3.22) and the same argument

$$F = \frac{\sigma_{\epsilon a}^2}{\sigma_{\epsilon ca}^2} = \frac{\hat{\theta}_0}{\hat{\theta}_0c} \sim f(\nu_{\theta_0}, \nu_{\theta_{0c}}),$$

(5.16)

And we use the same decision rule as described in (3.23) and the following discussion.
5.2.3 Exporting the Variance Model

Now assume the variance model is shared across the control sample and the study sample. This is a weaker assumption when compared to the assumption of constant variance over the same compound of the control and study samples ($\sigma^2_{\delta a} = \sigma^2_{\epsilon a}$).

5.2.3.1 Parameter Estimation

Recall from (3.102) that $\hat{d}_a = \bar{W}_a - \hat{\beta}_0$ does not depend on $\sigma^2_{\delta a}$ and $\sigma^2_{Va}$. Also note that $V[\hat{d}_a]$ in (3.105) depends on $\sigma^2_{\delta a}$ and $\sigma^2_{Va}$ only through $V[\bar{W}_a]$, then the inferences about $\hat{d}_a$ remain as described in Section 3.3.3.

*Within-material variance estimation using separate estimates*

Assuming the within-laboratory variance model is adequate and that it is shared by both the control and the study sample, but allowed to change over compounds, we have

$$\sigma^2_{\delta a} = \theta_{0a}(d_a^2 + \sigma^2_{Va}). \quad (5.17)$$

We can estimate this variance as

$$\hat{\sigma}^2_{\delta a} = \hat{\theta}_{0a}(d_a^2 - \hat{\sigma}^2_{\delta a} + \hat{\sigma}^2_{Va}) \quad (5.18)$$

or more revealingly

$$\hat{\sigma}^2_{Va} = \frac{s^2_{Wa} - \hat{\theta}_{0a}((\bar{W}_a - \bar{Y}_a + \hat{x}_a)^2 - (s^2_{Wa} + s^2_{Ya} + \hat{\sigma}^2_{x} - 2s_{Y Wa}))}{1 + \hat{\theta}_{0a}}. \quad (5.19)$$

For inferences of the variance due to heterogeneity we can update (3.34) by dropping the assumption $\sigma^2_{\delta a} = \sigma^2_{\epsilon a}$ and using the variance model (5.17) as
\[
\frac{(M_a - 1)s_{Wa}^2}{\sigma_{\delta a}^2 + \sigma_{V a}^2} = \frac{(M_a - 1)s_{Wa}^2}{\theta_0(d_a^2 + \sigma_{V a}^2) + \sigma_{V a}^2} \sim \chi_{(M_a-1)}^2.
\] (5.20)

Hence a \((1 - \alpha)100\%\) CI for \(\sigma_{V a}^2\) is

\[
\frac{1}{1 + \theta_0} \left( \frac{(M_a - 1)s_{Wa}^2}{\chi_{(1-\alpha/2;M_a-1)}} - \theta_0(d_a^2 - \hat{\sigma}_{\delta a}^2), \frac{(M_a - 1)s_{Wa}^2}{\chi_{(\alpha/2;M_a-1)}} - \theta_0(d_a^2 - \hat{\sigma}_{\delta a}^2) \right).
\] (5.21)

If the CI contains zero then it must be truncated.

Under this scenario, the information from the control material is exported via the models \(g\) and \(h\) with their respective parameters. The restriction of having similar mean values of concentration and exactly the same variance on the control and study materials as mentioned in Section 3.1.3.1 is no longer required. However, it is possible to have non-overlapping ranges of mean values and variances for the control material and the study material. In such cases there is an extrapolation risk when exporting the models and some of the estimates for the study material may not be adequate.

**Within-material variance estimation using the pooled estimate**

Similarly we can estimate the within-material variance of the study sample \(\sigma_{V a}^2\). In addition to the compound-wise individual estimates detailed in Section 3.1.3.1 now we can also consider the pooled estimate from the control sample. This requires assuming that the variance model is shared across all the compounds and (5.18), (5.19) and (5.21) need to be updated as

\[
\hat{\sigma}_{\delta a}^2 = \theta_0(d_a^2 - \hat{\delta}_{\delta a}^2 + \hat{\sigma}_{V a}^2),
\] (5.22)

\[
\hat{\sigma}_{V a}^2 = \frac{s_{Wa}^2 - \theta_0((\bar{W}_a - \bar{Y}_a + \bar{x})^2 - (s_{Wa}^2 + s_{Ya}^2 + \hat{\delta}_{\delta a}^2 - 2s_{WYa}))}{1 + \theta_0}
\] (5.23)

and a \((1 - \alpha)100\%\) CI for \(\sigma_{V a}^2\) becomes

\[
\frac{1}{1 + \theta_0} \left( \frac{(M_a - 1)s_{Wa}^2}{\chi_{(1-\alpha/2;M_a-1)}} - \theta_0(d_a^2 - \hat{\sigma}_{\delta a}^2), \frac{(M_a - 1)s_{Wa}^2}{\chi_{(\alpha/2;M_a-1)}} - \theta_0(d_a^2 - \hat{\sigma}_{\delta a}^2) \right).
\] (5.24)
We obtain a smoothness effect on the estimates by averaging out some noise from the individual estimates $\hat{\sigma}^2_{da}$. However, it is possible to obtain negative point and interval estimates, in those cases we assign a zero value. Also the pooled estimate is sensitive to outliers. The probability of the estimate taking negative values in (3.37) needs to be updated as

\begin{equation}
Pr(\hat{\sigma}^2_{Va} < 0) = Pr \left( \frac{1}{1 + \theta_0} \left( \frac{(M_a - 1)s^2_{Wa}}{\chi^2_{M_a-1}} - \hat{\theta}_0(\hat{d}^2_a - \hat{\sigma}^2_{da}) \right) < 0 \right)
= Pr \left( \frac{(M_a - 1)s^2_{Wa}}{\chi^2_{M_a-1}} - \hat{\theta}_0(\hat{d}^2_a - \hat{\sigma}^2_{da}) < 0 \right)
= 1 - Pr \left( \frac{\chi^2_{(M_a-1)}}{\hat{\theta}_0(\hat{d}^2_a - \hat{\sigma}^2_{da})} \right). \tag{5.25}
\end{equation}

**Within-material variance estimation using both samples simultaneously**

Under the hypothesis of known within-material variance of both the study and control materials all the information from both samples can be combined simultaneously to estimate the within-laboratory variance. This condition is reasonable to assume since the study materials distributed for interlaboratory studies are often intended to be essentially homogeneous (ISO GUIDE 17043 (2008) [30]) hence $\hat{\sigma}^2_{Va} = 0$.

For the constant within-laboratory variance model we can simply use all the information to obtain a pooled estimate of the within-laboratory variance, using (5.4) and (5.18) then

\begin{equation}
\hat{\theta}_{0a} = \frac{1}{2} \left( \frac{s^2_{Ya} - \hat{\sigma}^2_{Ua}}{\hat{d}^2_a - \hat{\sigma}^2_{xa}} + \frac{s^2_{Wa} - \hat{\sigma}^2_{Va}}{\hat{d}^2_a - \hat{\sigma}^2_{da} + \hat{\sigma}^2_{Va}} \right). \tag{5.26}
\end{equation}

If the estimate is negative it must be set to zero.

In addition if we can assume the model is shared across compounds then

\begin{equation}
\hat{\theta}_0 = \frac{1}{A} \sum_a \hat{\theta}_{0a}. \tag{5.27}
\end{equation}
5.3 Example

Separate estimates

Using the data in Table 3.1 and using (5.11) we obtain the separate estimates for laboratory 1, shown in Table 5.1.

<table>
<thead>
<tr>
<th>Compound</th>
<th>$\theta_{0a}$</th>
<th>$s_{Wa}^2$</th>
<th>$\hat{\sigma}_{\delta a}^2$</th>
<th>$\hat{\sigma}_{Va}^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2,3,5-trimethylnaphthalene</td>
<td>0.0017</td>
<td>0.31</td>
<td>1.39</td>
<td>0.00</td>
</tr>
<tr>
<td>acenaphthene</td>
<td>0.0033</td>
<td>0.21</td>
<td>4.49</td>
<td>0.00</td>
</tr>
<tr>
<td>dibenz[a,h]anthracene</td>
<td>0.00017</td>
<td>17.33</td>
<td>2.03</td>
<td>15.29</td>
</tr>
<tr>
<td>acenaphthylene</td>
<td>0.0015</td>
<td>1.84</td>
<td>1.01</td>
<td>0.83</td>
</tr>
<tr>
<td>1-methylphenanthrene</td>
<td>8e-05</td>
<td>0.39</td>
<td>0.40</td>
<td>0.00</td>
</tr>
<tr>
<td>biphenyl</td>
<td>0.00075</td>
<td>0.50</td>
<td>0.82</td>
<td>0.00</td>
</tr>
<tr>
<td>2,6-dimethylnaphthalene</td>
<td>0.00068</td>
<td>0.21</td>
<td>0.93</td>
<td>0.00</td>
</tr>
<tr>
<td>fluorene</td>
<td>0.0018</td>
<td>0.48</td>
<td>7.37</td>
<td>0.00</td>
</tr>
<tr>
<td>triphenylene</td>
<td>0.0006</td>
<td>72.33</td>
<td>11.42</td>
<td>60.85</td>
</tr>
<tr>
<td>1-methylnaphthalene</td>
<td>0.0026</td>
<td>0.82</td>
<td>7.88</td>
<td>0.00</td>
</tr>
<tr>
<td>anthracene</td>
<td>0.0011</td>
<td>6.33</td>
<td>32.73</td>
<td>0.00</td>
</tr>
<tr>
<td>benzo[j][fluoranthene</td>
<td>5e-05</td>
<td>72.33</td>
<td>3.81</td>
<td>68.46</td>
</tr>
<tr>
<td>benzo[k]fluoranthene</td>
<td>0.0005</td>
<td>76.00</td>
<td>82.01</td>
<td>0.00</td>
</tr>
<tr>
<td>2-methylnaphthalene</td>
<td>0.0018</td>
<td>0.65</td>
<td>7.91</td>
<td>0.00</td>
</tr>
<tr>
<td>chrysene</td>
<td>0.0019</td>
<td>14.33</td>
<td>511.02</td>
<td>0.00</td>
</tr>
<tr>
<td>benzo[ghi]perylene</td>
<td>0.00032</td>
<td>25.00</td>
<td>70.61</td>
<td>0.00</td>
</tr>
<tr>
<td>benzo[e]pyrene</td>
<td>0.00042</td>
<td>144.33</td>
<td>152.44</td>
<td>0.00</td>
</tr>
<tr>
<td>benzo[a]anthracene</td>
<td>0.00087</td>
<td>16.33</td>
<td>129.71</td>
<td>0.00</td>
</tr>
<tr>
<td>indeno[1,2,3-cd]pyrene</td>
<td>0.00057</td>
<td>63.00</td>
<td>196.04</td>
<td>0.00</td>
</tr>
<tr>
<td>benzo[a]pyrene</td>
<td>0.0002</td>
<td>32.33</td>
<td>79.07</td>
<td>0.00</td>
</tr>
<tr>
<td>perylene</td>
<td>0.00059</td>
<td>49.00</td>
<td>16.40</td>
<td>32.57</td>
</tr>
<tr>
<td>phenanthrene</td>
<td>0.00061</td>
<td>49.00</td>
<td>159.48</td>
<td>0.00</td>
</tr>
<tr>
<td>benzo[b]fluoranthene</td>
<td>0.00023</td>
<td>380.33</td>
<td>272.40</td>
<td>107.82</td>
</tr>
<tr>
<td>pyrene</td>
<td>0.00031</td>
<td>189.00</td>
<td>405.28</td>
<td>0.00</td>
</tr>
<tr>
<td>fluoranthene</td>
<td>0.0026</td>
<td>93.00</td>
<td>3231.85</td>
<td>0.00</td>
</tr>
<tr>
<td>naphthalene</td>
<td>9.1e-05</td>
<td>6.33</td>
<td>0.62</td>
<td>5.71</td>
</tr>
</tbody>
</table>

Table 5.1. Variance modeling separate estimates for laboratory 1.

For each compound, $\bar{W} = \text{mean of the study value for reference}$, $\hat{\theta}_{0a} = \text{the estimated parameter of the variance model}$, $\hat{\sigma}_{\delta a}^2 = \text{the estimated variance of the study data}$, $s_{Wa}^2 = \text{the observed variance of the study data}$, $\hat{\sigma}_{Va}^2 = \text{the estimated heterogeneity variance of the study sample}$.

Assuming $\theta_{0a} = \theta_0, \forall a$ we obtain the pooled estimate shown in Table 5.2.

Assessing the Performance
Using the envelope evaluation in (5.13), the pooled estimate \( \hat{\theta}_0 = 0.000976 \) and the pooled estimate \( \hat{\theta}_{0c} = 0.005637 \) from the CRM data we get a test statistic of \( F = 0.17315 \). The evaluation criteria described in Section 3.1.2.3 is still valid. The acceptance (satisfactory) region is \((f_{0.023;29,192}, f_{(0.977;29,192)}) = (0.5298, 1.6725)\), the rejection (unsatisfactory) region is \((f_{0.001;29,192}, f_{(0.999;29,192)})^c = (0, 0.3614] \cup [2.1857, \infty)\).

Applying the evaluation criteria as described, we must conclude the performance about the within-laboratory variance of the participant is unsatisfactory. However, under this condition we can conclude that the performance about the within variance of the participant is far superior with respect to the reference values. If the confidence interval of this ratio contains one then the observed within variance is satisfactory.

The point and interval estimates for \( d \) in Tables 3.11 and 3.12 remain the same.

Table 5.3 shows the estimates of the within-material variance of the study sample using separate variance models from the control sample.

Table 5.4 shows the estimates of the within-material variance of the study sample using the pooled estimate of the within-laboratory variance from the control sample.

Zero or negative estimates of \( \hat{\sigma}^2_{Va} \) would suggest the within-material variance of the study sample is zero (i.e. the study material is homogeneous) and we can take advantage of this situation to improve our variance model.

Table 5.5 shows the the estimates assuming the study material is homogeneous and using the control and study data simultaneously.

Comparing Tables 3.5, 5.3 and 5.4 we find that three compounds (dibenz[a,h]anthracene, triphenylene and benzo[j]fluoranthene) are consistently highlighted as suspicious for heterogeneity under the assumption of the constant bias model.
<table>
<thead>
<tr>
<th>Compound</th>
<th>$\hat{\sigma}_V^2$</th>
<th>lcl($\sigma_V^2$)</th>
<th>ucl($\sigma_V^2$)</th>
<th>Pr($\hat{\sigma}_V^2 &lt; 0$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2,3,5-trimethylnaphthalene</td>
<td>0.00</td>
<td>0.00</td>
<td>10.80</td>
<td>0.80</td>
</tr>
<tr>
<td>acenaphthene</td>
<td>0.00</td>
<td>0.00</td>
<td>3.79</td>
<td>0.95</td>
</tr>
<tr>
<td>dibenz[a,h]anthracene</td>
<td>15.30</td>
<td>2.67</td>
<td>682.00</td>
<td>0.00</td>
</tr>
<tr>
<td>acenaphthylene</td>
<td>0.83</td>
<td>0.00</td>
<td>71.70</td>
<td>0.16</td>
</tr>
<tr>
<td>1-methylphenanthrene</td>
<td>0.00</td>
<td>0.02</td>
<td>15.30</td>
<td>0.01</td>
</tr>
</tbody>
</table>
biphenyl                        | 0.00                | 0.00              | 16.10             | 0.88                     |
|2,6-dimethylnaphthalene         | 0.00                | 0.00              | 7.36              | 0.80                     |
|fluorene                        | 0.00                | 0.00              | 11.60             | 0.94                     |
triphenylene                    | 60.90               | 8.19              | 2840.00            | 0.00                     |
|1-methylnaphthalene             | 0.00                | 0.00              | 24.60             | 0.90                     |
anthracene                      | 0.00                | 0.00              | 217.00             | 0.82                     |
|benzo[j]fluoranthene            | 68.50               | 15.80             | 2850.00           | 0.00                     |
|benzo[k]fluoranthene            | 0.00                | 0.00              | 2920.00           | 0.40                     |
|2-methylnaphthalene             | 0.00                | 0.00              | 17.90             | 0.92                     |
|chrysene                        | 0.00                | 0.00              | 55.00             | 0.97                     |
|benzo[ghi]perylene              | 0.00                | 0.00              | 939.00            | 0.60                     |
|benzo[e]pyrene                  | 0.00                | 0.00              | 5610.00           | 0.21                     |
|benz[a]anthracene               | 0.00                | 0.00              | 330.00            | 0.95                     |
|indeno[1,2,3-cd]pyrene           | 0.00                | 0.00              | 2290.00           | 0.72                     |
|benzo[a]pyrene                  | 0.00                | 0.00              | 1200.00           | 0.66                     |
|perylenes                       | 32.60               | 0.00              | 1920.00           | 0.05                     |
|phenanthrene                    | 0.00                | 0.00              | 1770.00           | 0.74                     |
|benzo[b]fluoranthene            | 108.00              | 0.00              | 14700.00          | 0.25                     |
|pyrene                          | 0.00                | 0.00              | 7060.00           | 0.63                     |
|fluoranthene                    | 0.00                | 0.00              | 440.00            | 0.97                     |
naphthalene                     | 5.72                | 1.10              | 250.00             | 0.00                     |

Table 5.3. Example of point and interval estimates of study sample within-material variance using the separate variance models from the control sample.
<table>
<thead>
<tr>
<th>Compound</th>
<th>$\hat{\sigma}_V^2$</th>
<th>$lcl_{\hat{\sigma}_V^2}$</th>
<th>$ucl_{\hat{\sigma}_V^2}$</th>
<th>$Pr(\hat{\sigma}_V^2 &lt; 0)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2,3,5-trimethylnaphthalene</td>
<td>0.00</td>
<td>0.00</td>
<td>11.40</td>
<td>0.67</td>
</tr>
<tr>
<td>acenaphthene</td>
<td>0.00</td>
<td>0.00</td>
<td>6.93</td>
<td>0.85</td>
</tr>
<tr>
<td>dibenz[a,h]anthracene</td>
<td>5.82</td>
<td>0.00</td>
<td>673.00</td>
<td>0.22</td>
</tr>
<tr>
<td>acenaphthylene</td>
<td>1.19</td>
<td>0.00</td>
<td>72.00</td>
<td>0.06</td>
</tr>
<tr>
<td>1-methylphenanthrene</td>
<td>0.00</td>
<td>0.00</td>
<td>14.30</td>
<td>0.69</td>
</tr>
<tr>
<td>biphenyl</td>
<td>0.00</td>
<td>0.00</td>
<td>15.00</td>
<td>0.90</td>
</tr>
<tr>
<td>2,6-dimethylnaphthalene</td>
<td>0.00</td>
<td>0.00</td>
<td>6.96</td>
<td>0.85</td>
</tr>
<tr>
<td>fluorene</td>
<td>0.00</td>
<td>0.00</td>
<td>14.90</td>
<td>0.89</td>
</tr>
<tr>
<td>triphenylene</td>
<td>53.70</td>
<td>1.05</td>
<td>2840.00</td>
<td>0.02</td>
</tr>
<tr>
<td>1-methylnaphthalene</td>
<td>0.00</td>
<td>0.00</td>
<td>29.50</td>
<td>0.75</td>
</tr>
<tr>
<td>anthracene</td>
<td>0.00</td>
<td>0.00</td>
<td>220.00</td>
<td>0.81</td>
</tr>
<tr>
<td>benzo[j]fluoranthene</td>
<td>0.00</td>
<td>0.00</td>
<td>2780.00</td>
<td>0.38</td>
</tr>
<tr>
<td>benzo[k]fluoranthene</td>
<td>0.00</td>
<td>0.00</td>
<td>2840.00</td>
<td>0.62</td>
</tr>
<tr>
<td>2-methylnaphthalene</td>
<td>0.00</td>
<td>0.00</td>
<td>21.60</td>
<td>0.86</td>
</tr>
<tr>
<td>chrysene</td>
<td>0.00</td>
<td>0.00</td>
<td>296.00</td>
<td>0.95</td>
</tr>
<tr>
<td>benzo[ghi]perylene</td>
<td>0.00</td>
<td>0.00</td>
<td>841.00</td>
<td>0.84</td>
</tr>
<tr>
<td>benzo[e]pyrene</td>
<td>0.00</td>
<td>0.00</td>
<td>5490.00</td>
<td>0.51</td>
</tr>
<tr>
<td>benz[a]anthracene</td>
<td>0.00</td>
<td>0.00</td>
<td>291.00</td>
<td>0.95</td>
</tr>
<tr>
<td>indeno[1,2,3-cd]pyrene</td>
<td>0.00</td>
<td>0.00</td>
<td>2150.00</td>
<td>0.83</td>
</tr>
<tr>
<td>benzo[a]pyrene</td>
<td>0.00</td>
<td>0.00</td>
<td>887.00</td>
<td>0.92</td>
</tr>
<tr>
<td>perylene</td>
<td>21.70</td>
<td>0.00</td>
<td>1910.00</td>
<td>0.16</td>
</tr>
<tr>
<td>phenanthrene</td>
<td>0.00</td>
<td>0.00</td>
<td>1680.00</td>
<td>0.82</td>
</tr>
<tr>
<td>benzo[b]fluoranthene</td>
<td>0.00</td>
<td>0.00</td>
<td>13900.00</td>
<td>0.72</td>
</tr>
<tr>
<td>pyrene</td>
<td>0.00</td>
<td>0.00</td>
<td>6180.00</td>
<td>0.86</td>
</tr>
<tr>
<td>fluoranthene</td>
<td>0.00</td>
<td>0.00</td>
<td>2470.00</td>
<td>0.92</td>
</tr>
<tr>
<td>naphthalene</td>
<td>0.00</td>
<td>0.00</td>
<td>244.00</td>
<td>0.38</td>
</tr>
</tbody>
</table>

Table 5.4. Example of point and interval estimates of study sample within-material variance using the pooled estimate for the variance model from the control sample.

<table>
<thead>
<tr>
<th>$\theta_0$</th>
<th>$dof(\theta_0)$</th>
<th>$lcl(\theta_0)$</th>
<th>$ucl(\theta_0)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.000795</td>
<td>45</td>
<td>0.000502</td>
<td>0.001115</td>
</tr>
</tbody>
</table>

Table 5.5. Variance modeling assuming homogeneous materials and constant coefficient for laboratory 1, using both control and study data simultaneously.
CHAPTER 6
CONCLUSIONS AND FUTURE WORK

6.1 Conclusions

This project has made important methodological progress in the analysis of intercomparison data. We have proposed more realistic models for intercomparison data, ways to fit those models to data based on certified reference materials and then how to use that information to find consensus values for a new study material analyzed by multiple laboratories.

In Chapter 3, we consider models that allow separate biases and variances for each laboratory at each compound and each point in time (with the model of constant biases over labs or time as a special case). In this context, we obtain unbiased method of moments estimators for the bias and its variance. Based on these models we obtain unbiased estimators for the true value of the study sample corrected for bias and its variance.

The problem of obtaining a consensus value using information from multiple sources was reviewed. The one way random effects has become the de facto method for certifying reference values when the sources are assumed unbiased. However, it is hard to justify its use when there is evidence of some kind of biasedness.

In Chapter 4, we extend the work by modeling the bias as a function of the concentration. This was done primarily by exploiting data from multiple compounds. With the uncertainty in the CRMs this leads to a measurement error in linear regression problem. We use the method of moments and the bootstrap technique to obtain
biased but consistent estimators. Noting that this situation can be stated as a ratio problem we also used Fieller’s method in a novel way.

In order to analyze the performance of these estimators (method of moments, Bootstrap technique, and Fieller’s method) a simulation study was conducted. Simulation suggests that the Fieller’s method tends to perform better than the bootstrap technique. When the variance of the error in the equation is changing with compounds the Fieller’s method tends to perform better than the method of moments. However, when the variance of the error in the equation is constant across compounds, it is unclear which of these two methods performs better.

Simulation also suggests that these three methods are more sensitive to the sample size, the variance of the measurement error, and whether the variance of the error in the equation is constant or changing across the compounds rather than the number of replicates, the true intercept and slope of the bias model ($\beta_0, \beta_1$) and the heterogeneity variance. We also show that the variance of the measurement error becomes important when it is comparable to the dispersion of the predictor values.

We extended the theory for estimating the variance of the error in the equation under heterogeneity and proposed an adjusted estimator that outperforms the method of moments estimator, the estimator based on the mean squared error, and the estimator based on replicated data.

In Chapter 5, we present a preliminary analysis for one variance model. We used it to estimate the variance component associated with heterogeneity in the material of study that otherwise is ignored during this kind of analysis. More importantly, the use of a variance model allows us to weaken the assumption that a constant variance of the error in the equation is shared across the control and study samples for each compound.

For the purpose of estimating a linear bias model, we conclude that it is more convenient to allocate units of several different CRMs as controls rather than using
several allocated units of one CRM. This design would allow estimating separate bias models for each compound and would provide the required data with no additional effort.

For the purpose of obtaining a consensus value, we conclude that a biased laboratory with small variance may provide more information than an unbiased laboratory with large variance, since in the first case we can obtain estimates with reduced bias and relatively small variance.

6.2 Future Work

In the process of writing this dissertation, we have observed that there are some interesting topics that can be explored in future research.

The models that we described in this thesis have the potential of being applicable to other types of experiments and performance evaluation studies such as the proficiency testing and CRM characterization. Basically it can be applied to any experimental data that undergoes a calibration procedure and where the measurements of the different compounds are intrinsically made in batches. The strongest assumption is that the bias model is shared by those compounds measured in a batch.

A number of interesting problems arise when we consider modeling how the performance of a lab might change over time or when we try to exploit a potential trend in the true values of the study material over time.

Models where the bias and variance are functions of time only can be treated with standard linear regression theory since no measurement error is present. In this sense it is of no interest to the field of statistics but it is an important tool for intercomparison exercises.

Modeling bias as a linear function of time and concentration leads to a multiple linear regression problem with interactions and with the CRM having uncertainty this leads to the problem of measurement error in multiple linear regression. Consider,
for example, one laboratory measuring over time a set of materials of the same type and containing the same compound. Assuming the coefficients in the bias change as a linear function of time, then

\[
\beta_{0,t} = \alpha_{0,0} + \alpha_{0,1}t, \\
\beta_{1,t} = \alpha_{1,0} + \alpha_{1,1}t.
\]  

(6.1)

Then the control material model at time \( t \) becomes

\[
Y_{tk} = (\alpha_{0,0} + \alpha_{0,1}t) + (\alpha_{1,0} + \alpha_{1,1}t)(x_t + U_{tk}) + \epsilon_{tk} \\
= \alpha_{0,0} + \alpha_{0,1}t + \alpha_{1,0}x_t + \alpha_{1,1}x_tt + \eta_{tk} \\
= \beta_0 + \beta_1t + \beta_2x_t + \beta_3x_tt + \eta_{tk}
\]  

(6.2)

where \( V[\eta_{tk}] = (\alpha_{1,0} + \alpha_{1,1}t)^2\sigma_{Ut}^2 + \sigma_{\epsilon t}^2 \). This can be treated with multiple linear regression theory with measurement error in \( x \) and interaction terms.

More realistically the true bias and variance can be seen as an unobserved component model and state-space models can be used, especially a model with stochastic drift. These kinds of models have been applied extensively in areas such as ecology and economics.

A second feature that could be explored is where a dynamic model or a trend in the true study value (the \( d \)’s earlier in the thesis) over time may be present. An important area of future work would be to estimate the dynamic/trend model in this context and, secondarily to consider exploiting that model for estimation of the true value at a particular point in time.
APPENDIX A

THE WELCH-SATTERTHWAITE APPROXIMATION

Estimating the degrees of freedom of a linear combination of sample variances

Consider a general linear combination of sample variances and covariances of \( N \) normal random variables

\[
\hat{C} = \sum_{i=1}^{N} \sum_{j=1}^{N} c_{i,j} S_{i,j}
\]  \hspace{1cm} (A.1)

where \( c_{i,j} \) are fixed and known and \( S_{i,j} \) are the sample covariance terms. Also assume

\[
\hat{C} \sim \chi_{\nu}^2,
\]  \hspace{1cm} (A.2)

for some degrees of freedom \( \nu \).

We are interested on estimating \( \nu \) the approximate degrees of freedom of \( \hat{C} \) by using the Welch-Satterthwaite equation. In order to do so we need to find an estimate of \( V[\hat{C}] \). Then since \( E[S_{ij}] = \sigma_{ij} \)

\[
E[\hat{C}] = \sum_{i=1}^{N} \sum_{j=1}^{N} c_{i,j} \sigma_{i,j},
\]  \hspace{1cm} (A.3)
\[ V[\hat{C}] = V \left[ \sum_{i}^{N} \sum_{j}^{N} c_{i,j} S_{i,j} \right] \]
\[ = E \left[ \left( \sum_{i}^{N} \sum_{j}^{N} c_{i,j} S_{i,j} \right)^2 \right] - E \left[ \sum_{i}^{N} \sum_{j}^{N} c_{i,j} S_{i,j} \right]^2 \]
\[ = E \left[ \sum_{i}^{N} \sum_{j}^{N} c_{i,j}^2 S_{i,j}^2 + \sum_{i}^{N} \sum_{j}^{N} \sum_{k}^{N} \sum_{l \neq (k,l)} c_{i,j} c_{k,l} S_{i,j} S_{k,l} \right] \]
\[ - \left( \sum_{i}^{N} \sum_{j}^{N} c_{i,j} E[S_{i,j}] \right)^2 \] (A.4)
\[ = \sum_{(i,j)} c_{i,j}^2 E[S_{i,j}^2] + \sum_{(i,j) \neq (k,l)} c_{i,j} c_{k,l} E[S_{i,j} S_{k,l}] \]
\[ - \sum_{(i,j)} c_{i,j}^2 E[S_{i,j}]^2 - \sum_{(i,j) \neq (k,l)} c_{i,j} c_{k,l} E[S_{i,j}] E[S_{k,l}] \]
\[ = \sum_{(i,j)} c_{i,j}^2 V[S_{i,j}] + \sum_{(i,j) \neq (k,l)} c_{i,j} c_{k,l} Cov[S_{i,j}, S_{k,l}], \]

and

\[ \nu \approx \frac{2 \left( E[\hat{C}] \right)^2}{V[\hat{C}]} \] (A.5)

Under independence of the random variables (A.4) reduces to

\[ V[\hat{C}] = \sum_{(i)} c_{i,i}^2 V[S_{i,i}]. \] (A.6)

In particular for \( Z = aX + bY \) and \( C = V[Z], \)

\[ \hat{C} = a^2 S_X^2 + b^2 S_Y^2 + 2abS_{X,Y}, \]
\[ E[\hat{C}] = a^2 V[X] + b^2 V[Y] + 2ab Cov[X, Y] = C, \] (A.7)
\[ V[\hat{C}] = a^4 V[S_X^2] + b^4 V[S_Y^2] + 4a^2 b^2 V[S_{X,Y}] + 2a^2 b^2 Cov[S_X^2, S_Y^2] \]
\[ + 4a^3 b Cov[S_X^2, S_{X,Y}] + 4ab^3 Cov[S_Y^2, S_{X,Y}]. \]
The expressions involved in (A.7) can be rewritten in terms of the moments, for a general distribution with sample variances and covariances built out of \( n \) replicates. They become

\[
V[S_X^2] = \frac{1}{n} \mu_{40} - \frac{(n-3)}{n(n-1)} \mu_{20}^2, \\
V[S_Y^2] = \frac{1}{n} \mu_{04} - \frac{(n-3)}{n(n-1)} \mu_{02}^2, \\
V[S_{X,Y}] = \frac{1}{n} \mu_{22} - \frac{(n-2)}{n(n-1)} \mu_{11}^2 + \frac{1}{n(n-1)} \mu_{20}\mu_{02}, \\
Cov[S_X^2, S_Y^2] = \frac{1}{n} (\mu_{22} - \mu_{20}\mu_{02}) - \frac{2}{n(n-1)} \mu_{11}^2, \\
Cov[S_X^2, S_{X,Y}] = \frac{1}{n} \mu_{31} - \frac{n-3}{n(n-1)} \mu_{20}\mu_{11}, \\
Cov[S_Y^2, S_{X,Y}] = \frac{1}{n} \mu_{13} - \frac{n-3}{n(n-1)} \mu_{02}\mu_{11}.
\]

Under normality these are reduced to

\[
V[S_X^2] = \frac{2}{n-1} \mu_{20}^2, \\
V[S_Y^2] = \frac{2}{n-1} \mu_{02}^2, \\
V[S_{X,Y}] = \frac{1}{n-1} (\mu_{20}\mu_{02} + \mu_{11}^2), \\
Cov[S_X^2, S_Y^2] = \frac{2}{n-1} \mu_{11}^2, \\
Cov[S_X^2, S_{X,Y}] = \frac{2}{n-1} \mu_{20}\mu_{11}, \\
Cov[S_Y^2, S_{X,Y}] = \frac{2}{n-1} \mu_{02}\mu_{11}.
\]

Furthermore
\[ V[\hat{C}] = \frac{2}{n-1} \left( a^4 \mu_{20}^2 + b^4 \mu_{02}^2 + 2a^2b^2 \mu_{20}\mu_{02} + 4a^2b^2 \mu_{11}^2 \right. \\
+ 4ab(a^2 \mu_{20} + b^2 \mu_{02}) \mu_{11} \right) \\
= \frac{2}{n-1} \left( a^2 \mu_{20} + b^2 \mu_{02} + 2ab \mu_{11} \right)^2 \\
= \frac{2}{n-1} \left( E[\hat{C}] \right)^2. \] (A.10)

Hence \( \nu \approx n - 1 \) and \( \hat{C} = a^2 S_X^2 + b^2 S_Y^2 + 2ab S_{X,Y} \) has approximately \( n - 1 \) degrees of freedom under balancedness, independently of the correlation between \( X, Y \).
APPENDIX B
LEVENE TEST UNDER HETEROGENEITY

Given $Y_{ak} = \mu_a + \epsilon_{ak}$, where $\mu_a$ are neither known nor assumed equal and $\epsilon_{ak}$ are independent and similarly distributed with zero mean and possibly unequal variances $V[\epsilon_{ak}] = \sigma_{\epsilon a}^2$.

We are interested in testing $H_0$: all the $\sigma_{\epsilon a}^2$ are equal. Levene’s Test (Brown and Forsythe (1974) [5]) proceeds as follows. Define $Z_{ak} = |Y_{ak} - \bar{Y}_a|$ then define the statistic

$$W_0 = \frac{\sum_a K_a (\bar{Z}_a. - \bar{Z}_.)^2 / (A - 1)}{\sum_a \sum_k (Z_{ak} - Z_a.)^2 / \sum_a (K_a - 1)}$$

where $\bar{Z}_a. = \sum_k Z_{ak} / K_a$ and $\bar{Z}_. = \sum_a \sum_k Z_{ak} / \sum_a K_a$. Then

$$W_0 \sim F(A - 1, \sum_a (K_a - 1))$$

and we can reject the null hypothesis if $W_0 > F(1 - \alpha; A - 1, \sum_a (K_a - 1))$ at the $\alpha$ level.

Consider now the case when we cannot observe $Y_{ak}$ directly, instead we observe $Y'_{ak} = Y_{ak} + U_{ak}$ where $U_{ak}$ are independent and similarly distributed with zero mean and unequal variances $\sigma_{Ua}^2$ and we have estimates of them, $\hat{\sigma}_{Ua}^2$. We are still interested on testing the null hypothesis $H_0: \sigma_{\epsilon a}^2 = \sigma_{\epsilon a}^2$.

Levene (1960) [35] stated that (B.2) holds for any function of $Z_{ak}$ monotonically increasing on $(0, \infty)$. In particular it holds for $Z_{ak} = (Y'_{ak} - \bar{Y}_a')^2 - \hat{\sigma}_{Ua}^2$ then $E[Z_{ak}'] \approx \sigma_{\epsilon a}^2$ and we can apply the Levene’s Test.
APPENDIX C
THE VARIANCE AND COVARIANCE OF THE SAMPLE
VARIANCE AND SAMPLE COVARIANCE OF
INDEPENDENT RANDOM VARIABLES

Theorem 1 Let $X_i$ be $n$ independent random variables with $E[X_i] = \mu_i$, $V[X_i] = \mu_{2,i}$, $E[(X_i - \mu_i)^3] = \mu_{3,i}$ and $E[(X_i - \mu_i)^4] = \mu_{4,i}$, the second, third and fourth finite central moments, respectively. Also define $\Sigma = \text{diag}(\mu_{2,i})$, $M_3 = \text{diag}(\mu_{3,i})$ and $M_4 = \text{diag}(\mu_{4,i})$. If $A$ is any $n \times n$ symmetric matrix and $a$ is the column vector of the diagonal elements of $A$, then

$$V[X'AX] = a'(M_4 - 3\Sigma^2)a + 2\text{tr}((A\Sigma)^2) + 4\mu'AX\mu + 4\mu'AM_3a.$$ \hfill (C.1)

Note that this quantity exists and is well defined for any $n$ independent random variables with only the assumption of having finite first four central moments; sharing the same distribution is not a requirement, neither is having constant variance (homo-skedasticity), skewness or kurtosis.

Proof

We proceed as in Theorem 1.6 from Seber&Lee (2003) [48]. By definition

$$V[X'AX] = E[(X'AX)^2] - E[X'AX]^2.$$ \hfill (C.2)

Setting $Y = X - \mu$, we have $E[Y] = 0$, $V[Y] = \Sigma$, $\text{diag}(E[Y^3]) = M_3$, $\text{diag}(E[Y^4]) = M_4,$
\[ E[X'A X] = E[(\mu + Y)'A(\mu + Y)] = \mu' A \mu + 2 \mu' A E[Y] + E[Y'A Y] \]
\[ = \mu' A \mu + tr(A \Sigma). \]  
(C.3)

Also
\[ X'A X = (X - \mu)' A (X - \mu) + 2 \mu' A (X - \mu) - \mu' A \mu \]
\[ = (X - \mu)' A (X - \mu) + 2 \mu' A (X - \mu) + \mu' A \mu \]  
(C.4)

so that squaring gives
\[ (X'A X)^2 = ((X - \mu)' A (X - \mu))^2 + 4(\mu' A (X - \mu))^2 + (\mu' A \mu)^2 \]
\[ + 2 \mu' A \mu (X - \mu)' A (X - \mu) + 4 \mu' A \mu' A (X - \mu) \]
\[ + 4 \mu' A (X - \mu) (X - \mu)' A (X - \mu). \]  
(C.5)

Taking the expected value of
\[ E[(X'A X)^2] = E[(Y'A Y)^2] + 4 E[(\mu' A Y)^2] + (\mu' A \mu)^2 + 2 \mu' A \mu E[Y'A Y] \]
\[ + 4 E[\mu' A YY' A Y]. \]  
(C.6)

As a first step in evaluating the expression above we note that
\[ (Y'A Y)^2 = \sum \sum \sum \sum a_{ij} a_{kl} Y_i Y_j Y_k Y_l. \]  
(C.7)

Since the \( Y_i \) are mutually independent with different second, third and fourth moments about the origin, we have
\[ E[Y_i Y_j Y_k Y_l] = \begin{cases} 
\mu_{4,i}, & i = j = k = l, \\
\mu_{2,i} \mu_{2,k}, & i = j \neq k = l, \\
\mu_{2,i} \mu_{2,j}, & i = k \neq j = l \text{ or } i = l \neq j = k, \\
0, & \text{otherwise}. 
\end{cases} \]  
(C.8)
Hence

\[ E[(Y'AY)^2] = \sum_{i} \sum_{j} \sum_{k} \sum_{l} a_{ij}a_{kl}E[Y_iY_jY_kY_l] \]

\[ = \sum_{i} a_{ii}a_{ii}\mu_4,i + \sum_{j \neq k} a_{ij}a_{ij}E[Y_i^2Y_j^2] + \sum_{i} a_{ij}E[Y_i^2Y_j^2] \]

\[ + \sum_{i \neq k} a_{ij}a_{ji}E[Y_i^2Y_j^2] \]

\[ = \sum_{i} a_{ii}^2\mu_4,i + \sum_{i} \left( \sum_{k \neq i} a_{ii}a_{kk}\mu_2,i\mu_2,k + 2 \sum_{j \neq i} a_{ij}^2\mu_2,i\mu_2,j \right) \]

\[ = \sum_{i} a_{ii}^2\mu_4,i + \sum_{i} \left( a_{ii}\mu_2,i \sum_{j \neq i} a_{jj}\mu_2,j + 2\mu_2,i \sum_{j \neq i} a_{ij}^2\mu_2,j \right) \]

\[ = \sum_{i} a_{ii}^2\mu_4,i + \sum_{i} \left( a_{ii}\mu_2,i \sum_{j \neq i} a_{jj}\mu_2,j + 2\mu_2,i \sum_{j \neq i} a_{ij}^2\mu_2,j \right) \]

\[ + 3 \sum_{i} a_{ii}^2\mu_2,i - 3 \sum_{i} a_{ii}^2\mu_2,i \]

\[ = \sum_{i} a_{ii}^2(\mu_4,i - 3\mu_2,i) + \sum_{i} \left( a_{ii}\mu_2,i \sum_{j \neq i} a_{jj}\mu_2,j + 2\mu_2,i \sum_{j \neq i} a_{ij}^2\mu_2,j \right) \]

\[ = a'(M_4 - 3\Sigma^2)a + tr(\Sigma) + 2 \sum_{i} \sum_{j} a_{ij}^2\mu_2,i\mu_2,j \]

where we have used the symmetry property of \( A \).

Also

\[ (\mu'AY)^2 = (b'Y)^2 = \sum_{i} \sum_{j} b_ib_jY_iY_j, \]

\[ \mu'AYY'AY = \sum_{i} \sum_{j} \sum_{k} b_iY_iY_jY_k, \]

so that

\[ E[(\mu'AY)^2] = \sum_{i} \sum_{j} b_ib_jE[Y_iY_j] = \sum_{i} b_i^2\mu_2,i = b'\Sigma b = \mu'\Sigma A\mu, \]

\[ E[\mu'AYY'AY] = \sum_{i} \sum_{j} \sum_{k} b_iY_iY_jY_k = \sum_{i} b_i\mu_3,i = \mu'M_3a, \]

\( (C.10) \)
and

\[
tr((A\Sigma)^2) = tr\left( \begin{pmatrix} a_{11} \ldots a_{1n} \\ \vdots \ldots \vdots \\ a_{n1} \ldots a_{nn} \end{pmatrix} \begin{pmatrix} \mu_{2,1} & 0 \\ \vdots & \ddots & \vdots \\ 0 & \mu_{2,n} \end{pmatrix} \right)^2
= tr\left( \begin{pmatrix} a_{11}\mu_{2,1} \ldots a_{1n}\mu_{2,n} \\ \vdots \ldots \vdots \\ a_{n1}\mu_{2,1} \ldots a_{nn}\mu_{2,n} \end{pmatrix} \begin{pmatrix} a_{11}\mu_{2,1} \ldots a_{1n}\mu_{2,n} \\ \vdots \ldots \vdots \\ a_{n1}\mu_{2,1} \ldots a_{nn}\mu_{2,n} \end{pmatrix} \right)
\]

(C.12)

Finally, collecting all the terms leads to

\[
V[X'AX] = E[(X'AX)^2] - E[X'AX]^2
= E[(Y'AY)^2] + 4E[(\mu'A\mu)^2] + (\mu'A\mu)^2 + 2\mu'A\mu E[Y'AY]
+ 4E[\mu'A Y'Y'A Y] - (\mu'A\mu + tr(A\Sigma))^2
= a'(M_4 - 3\Sigma^2)a + tr(A\Sigma)^2 + 2\sum_i\sum_j a^2_{ij}\mu_{2,i}\mu_{2,j}
\]

(C.13)
Theorem 2 Let \( X_i, Y_i \) be \( n \) independent random variables with expected values \( E[X] = \mu_X = (\mu_{X_1}, \ldots, \mu_{X_n}), \ E[Y] = \mu_Y, \) and \( V[X] = \Sigma_{XX} = \text{diag}(\mu_{X_2}, \ldots, \mu_{X_n}), \) and \( V[Y] = \Sigma_{YY} = \text{diag}(\mu_{Y_2}, \ldots, \mu_{Y_n}) \), the second finite central moments. If \( A \) is any \( n \times n \) symmetric matrix, then

\[
V[X'A Y] = \mu_X' \Sigma_{YY} A \mu_X + \mu_Y' \Sigma_{XX} A \mu_Y + \text{tr}(A \Sigma_{XX} A \Sigma_{YY}). \tag{C.14}
\]

Proof

We proceed as in Theorem 1.6 from Seber&Lee (2003) [48].

\[
E[X' A Y] = \text{tr}(E[X' A Y]) = E[\text{tr}(X' A Y)] = E[\text{tr}(A Y X')] = \text{tr}(E[AY X'])
\]

\[
= \text{tr}(A E[YX']) = \text{tr}(A(E[X] + \mu_X \mu_X'))
\]

\[
= \text{tr}(A \Sigma_{XX}) + \mu_X' A \mu_Y
\]

\[
= \mu_X' A \mu_Y. \tag{C.15}
\]

Setting \( Z_1 = (X - \mu_X), \ Z_2 = (Y - \mu_Y) \) leads us to

\[
X' A Y = (\mu_X + Z_1)'(\mu_Y + Z_2)
\]

\[
= \mu_X' A \mu_Y + \mu_X' A Z_2 + \mu_Y' A Z_1 + Z_1' A Z_2. \tag{C.16}
\]

Then

\[
V[X' A Y] = V[\mu_X' A \mu_Y + \mu_X' A Z_2 + \mu_Y' A Z_1 + Z_1' A Z_2] \]

\[
= V[\mu_X' A Z_2] + V[\mu_Y' A Z_1] + V[Z_1' A Z_2]
\]

\[
= V[\mu_X' A Z_2] + 2 \text{Cov}[\mu_X' A Z_2, \mu_Y' A Z_1] + 2 \text{Cov}[\mu_X' A Z_2, Z_1' A Z_2] + 2 \text{Cov}[\mu_Y' A Z_1, Z_1' A Z_2]. \tag{C.17}
\]
As a first step in evaluating the expression above we note that for any \( n \times n \) matrix \( \mathbf{M} \) using the law of total covariance

\[
\text{Cov}[Z_1, Z_1' \mathbf{M} Z_2] = E[\text{Cov}[Z_1, Z_1' \mathbf{M} Z_2 | Z_1]] + \text{Cov}[E[Z_1 | Z_1], E[Z_1' \mathbf{M} Z_2 | Z_1]]
\]

\[
= E[Z_1 \text{Cov}[I, Z_2' Z_1] \mathbf{M} Z_1] + \text{Cov}[Z_1, Z_1' \mathbf{M} E[Z_2 | Z_1]]
\]

\[
= \text{Cov}[Z_1, Z_1'] \mathbf{M} E[Z_2] = 0
\]

(C.18)

where we use the independence of \( Z_1, Z_2 \).

If \( \mathbf{M} \) is symmetric also and using the law of total variance

\[
\text{V}[Z_1' \mathbf{M} Z_2] = \text{V}[E[Z_1' \mathbf{M} Z_2 | Z_1]] + \text{E}[\text{V}[Z_1' \mathbf{M} Z_2 | Z_1]]
\]

\[
= \text{V}[Z_1' \mathbf{M} E[Z_2 | Z_1]] + \text{E}[Z_1' \mathbf{M} \text{V}[Z_2 | Z_1] \mathbf{M} Z_1]
\]

\[
= \text{V}[Z_1' \mathbf{M} E[Z_2]] + \text{E}[Z_1' \mathbf{M} \text{V}[Z_2 | Z_1] \mathbf{M} Z_1]
\]

\[
= \text{V}[Z_1' \mathbf{M} \text{V}[Z_2] \mathbf{M} Z_1] = \text{tr}(\mathbf{M} \text{V}[Z_2] \mathbf{M} \text{V}[Z_1]).
\]

(C.19)

Hence

\[
\text{V}[\mathbf{X}' \mathbf{A} \mathbf{Y}] = \text{V}[\mu_X' \mathbf{A} Z_2] + \text{V}[\mu_Y' \mathbf{A} Z_1] + \text{V}[Z_1' \mathbf{A} Z_2]
\]

\[
= \mu_X' \mathbf{A} \Sigma_{YY} \mu_X + \mu_Y' \mathbf{A} \Sigma_{XX} \mu_Y + \text{tr}(\mathbf{A} \Sigma_{YY} \mathbf{A} \Sigma_{XX}).
\]

(C.20)

\[\square\]
Theorem 3 Let $X_i, Y_i$ be $n$ independent random variables with $E[X_i] = \mu_{X_i}$, and $\Sigma_{XX}, M_{X3}$ and $M_{X4}$ defined as in Theorem 1. Similarly define $\mu_{Y_i}, \Sigma_{YY}, M_{Y3}, M_{Y4}$ using the respective first moment and the second, third and fourth central moments of $Y$. If $A, B$ are any $n \times n$ symmetric matrices and $b$ is the column vector of the diagonal elements of $B$, then

$$\text{Cov}[X'AY, Y'BY] = 2\mu_X' A \Sigma_{YY} B \mu_Y + \mu_X' A M_{Y3} b.$$  \hfill (C.21)

Proof

Using the law of total covariance leads us to

$$\text{Cov}[X'AY, Y'BY] = E[\text{Cov}[X'AY, Y'BY|X]]$$

$$+ \text{Cov}[E[X'AY|X], E[Y'BY|X]]$$

$$= E[X'ACov[Y, Y'BY|X]]$$

$$+ \text{Cov}[X'AE[Y|X], E[Y'BY|X]]$$

$$= E[X'ACov[Y, Y'BY] + \text{Cov}[X'AY, 1] E[Y'BY]$$

$$= E[X'ACov[Y, Y'BY].$$

(C.22)

Setting $Z = (Y - \mu_Y)$, we have $E[Z] = 0$ and using the identity found in (C.11) leads us to

$$\text{Cov}[Y, Y'BY] = \text{Cov}[\mu_Y + Z, (\mu_Y + Z)'B(\mu_Y + Z)]$$

$$= \text{Cov}[Z, \mu_Y' B \mu_Y + 2\mu_Y' BZ + Z'BZ]$$

$$= \text{Cov}[Z, 2\mu_Y' BZ + Z'BZ]$$

$$= \text{Cov}[Z, 2\mu_Y' BZ] + \text{Cov}[Z, Z'BZ]$$

$$= 2\text{Cov}[Z, Z] B \mu_Y + \text{Cov}[Z, Z'BZ]$$

$$= 2\Sigma_{YY} B \mu_Y + M_{Y3} b.$$  \hfill (C.23)
Hence

\[
Cov[X'AY, Y'BY] = 2\mu_X' A \Sigma_{YY} B \mu_Y + \mu_X' A M \Sigma_{YY} \mu_Y + \mu_X' A M \mu_Y.
\]

(C.24)


