

# Fiducial approach for assessing agreement between two instruments

C M Wang<sup>1</sup> and Hari K Iyer<sup>1,2</sup>

<sup>1</sup> Statistical Engineering Division, National Institute of Standards and Technology, Boulder, CO 80305, USA

<sup>2</sup> Department of Statistics, Colorado State University, Fort Collins, CO 80523, USA

E-mail: [jwang@boulder.nist.gov](mailto:jwang@boulder.nist.gov)

Received 11 March 2008

Published 9 July 2008

Online at [stacks.iop.org/Met/45/415](http://stacks.iop.org/Met/45/415)

## Abstract

This paper presents an approach for making inferences about the intercept and slope of a linear regression model when both variables are subject to measurement errors. The approach is based on the principle of fiducial inference. A procedure is presented for computing uncertainty regions for the intercept and slope that can be used to assess agreement between two instruments. Computer codes for performing these calculations, written using open-source software, are listed.

(Some figures in this article are in colour only in the electronic version)

## 1. Introduction

In metrology it is often required to compare two different methods and/or instruments for measuring some quantity of interest. One type of comparison occurs when one method is a *reference* method and the other method is *calibrated* against this reference method. However, in many instances we simply have two different methods or instruments to measure the same quantity and we are interested in comparing the two methods. The topic of *assessing agreement* between two methods has been studied by many authors. For example, Bland and Altman [1–5] discussed the analysis of agreement between methods of clinical measurement. Lin *et al* [6] compared several measures of agreement, including mean squared deviation, total deviation index, concordance correlation coefficient and coverage probability, and derived the accuracy and precision components, which can be used to identify the sources of disagreement, of each measure. Choudhary [7] developed a tolerance interval approach for assessing agreement between two methods. In the metrological literature, Astrua *et al* [8] discussed and compared several techniques for assessing the degree of conformity between methods. They also pointed out the merits and drawbacks of each technique. The approaches included in their study were graphical procedures, procedures based on Pearson and concordance correlation coefficients, and regression procedures. In this paper we propose a *fiducial* approach for assessing the agreement between two methods.

The paper is organized as follows. In section 2 we review a simple *errors-in-variables* model that can be used to describe measurement models for the two methods being compared. Section 3 provides a fiducial solution for method comparison. The approach is based on generalizations of fiducial inference proposed by Fisher [9] and *structural inference* introduced by Fraser [10]. Section 4 uses an example to illustrate the fiducial solution to the agreement problem. Finally, we conclude with some summary remarks in section 5.

## 2. Models and assumptions

The comparison is carried out by measuring the same collection of artefacts by use of each of the methods. Let  $(X_i, Y_i)$ ,  $i = 1, \dots, n$ , be pairs of random variables representing the measurements of the  $n$  artefacts using two methods. The lower-case versions,  $(x_i, y_i)$ , represent the corresponding realized values. Since both  $X_i$  and  $Y_i$  are affected by measurement errors, we write

$$X_i = \theta_i + \epsilon_i, \quad Y_i = \mu_i + \delta_i, \quad i = 1, \dots, n, \quad (1)$$

where  $\theta_i$  and  $\mu_i$  are the expected values of  $X_i$  and  $Y_i$ , respectively, and  $\epsilon_i$  and  $\delta_i$  are independent errors. It is assumed that  $\epsilon_i \sim N(0, \sigma_x^2)$ ,  $\delta_i \sim N(0, \sigma_y^2)$ , and at least two  $\theta$ s are distinct. The total number of artefacts available is  $n$ , and each artefact is measured by both methods under study.

Many definitions of *agreement* are possible. Ideally, perfect agreement would mean that both methods yield the

same result as long as the same artefact is being measured. This is not a suitable concept of agreement for practical applications. A more reasonable definition of agreement is to require that if the same artefact is measured by each method a large number of times (mathematically speaking, an infinite number of times), then the average values of the two methods are equal. Accordingly, we say that two measurement methods agree with one another, on average, if the expected values  $E(X)$  and  $E(Y)$  are equal for every artefact that is measured by the two methods. In particular, agreement between two methods will imply that  $\mu_i = \theta_i$  for  $i = 1, \dots, n$ . Conversely, the conditions  $\mu_i = \theta_i, i = 1, \dots, n$ , would support a claim of agreement, on average, between the two methods. If, however, the relationship between the two methods is linear, that is,

$$\mu_i = \beta_0 + \beta_1 \theta_i, \quad i = 1, \dots, n, \quad (2)$$

then the question of whether the two methods agree reduces to assessing whether  $(\beta_0, \beta_1) = (0, 1)$ . Experience suggests that this linearity assumption is often a reasonable approximation of reality. Thus the aim of this paper is to develop a procedure, using fiducial inference methods, for testing whether or not  $(\beta_0, \beta_1) = (0, 1)$ , under the following errors-in-variables regression model:

$$\begin{aligned} Y_i &= \beta_0 + \beta_1 \theta_i + \delta_i \\ X_i &= \theta_i + \epsilon_i, \end{aligned} \quad (3)$$

and based on the  $n$  pairs of measurements  $(x_i, y_i)$ . Our approach will be to develop a method for constructing confidence regions for  $(\beta_0, \beta_1)$  and to use these confidence regions to assess the claim of agreement between the two methods.

The parameters of the model in (3) are  $\beta_0, \beta_1, \sigma_x, \sigma_y$  and  $\theta_i, i = 1, \dots, n$ . One additional assumption we need to make is that an estimate of  $\sigma_x^2$ , denoted by  $S_x^2$ , and the degrees of freedom  $\nu_x$  associated with  $S_x^2$ , reflecting the amount of information used to arrive at this estimate of  $\sigma_x^2$ , are available. If  $\nu_x$  is infinity,  $\sigma_x$  is then assumed known exactly. Also, we assume that

$$W_x = \frac{\nu_x S_x^2}{\sigma_x^2} \sim \chi^2(\nu_x), \quad (4)$$

where  $\chi^2(\nu)$  stands for a  $\chi^2$  distribution with  $\nu$  degrees of freedom.

In practice, the established method, such as a standard or a reference method, is often used as variable  $X$ . In such a situation, it is reasonable to assume the existence of an estimate of  $\sigma_x^2$  and its associated degrees of freedom.

### 3. Generalized fiducial inference

Statistical procedures based on fiducial inference have been developed for various applications in metrology, for example, see [11–14]. In this approach a probability distribution, called the *fiducial distribution*, for the parameters of interest conditional on the data is obtained based on the *structural equation* that relates the measurements with model parameters and error processes whose distributions are fully known. Once

obtained, the fiducial distribution can often be used to make inferences about the parameters of interest.

We use a simple example to illustrate the basic ideas of fiducial inference. Suppose  $Y \sim N(\mu, 1)$ . We write  $Y = \mu + Z$ , where  $\mu$  is the parameter of interest and  $Z \sim N(0, 1)$ . If  $y$  is a realized value of  $Y$  corresponding to the realized value  $z$  of  $Z$ , then we have  $\mu = y - z$ . Of course  $Z$  is not observable. Nevertheless, the fact that we know the distribution from which  $z$  was generated helps us determine a set of values of  $\mu$  that we consider plausible. We can use the probabilities associated with  $Z$  to infer the probabilities for  $\mu$ . The process of transferring the relationship  $\mu = y - z$  to the relationship  $\mu = y - Z$  is what constitutes the fiducial argument. The fiducial distribution of  $\mu$  is the distribution of the random variable  $\tilde{\mu} = y - Z$  with  $y$  fixed. That is,  $\tilde{\mu} \sim N(y, 1)$ . The random variable  $\tilde{\mu}$  is called the *fiducial quantity* (FQ) for  $\mu$ . The fiducial distribution is similar to the *posterior* distribution in the Bayesian framework. A central value of this fiducial distribution may be used as the estimated value of  $\mu$  and the spread of the fiducial distribution may be used to compute an uncertainty interval for  $\mu$ .

To obtain a joint fiducial distribution for  $\beta_0$  and  $\beta_1$ , we rely on a two-stage procedure for constructing the FQ [15]. The procedure is outlined below.

Let  $\xi = (\xi_1, \xi_2)$  and  $\xi_1$  be the parameter of interest. If

1.  $\tilde{\xi}_1(\xi_2)$  is a FQ for  $\xi_1$  given that  $\xi_2$  is known and
2.  $\tilde{\xi}_2$  is a FQ for  $\xi_2$ ,

then  $\tilde{\xi}_1(\tilde{\xi}_2)$  is a FQ for  $\xi_1$ .

The first step of the above procedure is to obtain a FQ for the parameter of interest  $\xi_1$  assuming that the nuisance parameter  $\xi_2$  is known. The resulting FQ for  $\xi_1$  would depend on  $\xi_2$ . The next step is to obtain a FQ for  $\xi_2$  and use this FQ to replace  $\xi_2$  in the FQ for  $\xi_1$  obtained in the first step.

To apply this procedure in our errors-in-variables regression problem, we first obtain a FQ for  $\beta = (\beta_0, \beta_1)^t$  assuming  $\theta_i, i = 1, \dots, n$ , are known. Since this FQ depends on  $\theta = (\theta_1, \dots, \theta_n)$ , we denote it by  $\tilde{\beta}_\theta$ . Next we obtain a FQ for  $\theta$  and denote it by  $\tilde{\theta}$ . The desired FQ for  $\beta$  is then given by

$$\tilde{\beta} = \tilde{\beta}_{\tilde{\theta}}.$$

Under the assumption of known  $\theta_i$ , the errors-in-variables regression model in (3) reduces to the ordinary linear regression model, and the following results are well known from the standard regression analysis. Let  $\mathbf{B}_\theta = (\mathbf{B}_0, \mathbf{B}_1)^t$  denote the least-squares estimator of  $\beta$  in the model

$$Y_i = \beta_0 + \beta_1 \theta_i + \delta_i.$$

Then

$$\mathbf{B}_\theta \sim N(\beta, \sigma_y^2 \mathbf{V}_\theta), \quad (5)$$

where

$$\begin{aligned} \mathbf{V}_\theta &= \begin{pmatrix} n & \sum_{i=1}^n \theta_i \\ \sum_{i=1}^n \theta_i & \sum_{i=1}^n \theta_i^2 \end{pmatrix}^{-1} \\ &= \frac{1}{D_\theta} \begin{pmatrix} \sum_{i=1}^n \theta_i^2 & -\sum_{i=1}^n \theta_i \\ -\sum_{i=1}^n \theta_i & n \end{pmatrix} \end{aligned}$$

and  $D_\theta = n \sum_{i=1}^n \theta_i^2 - (\sum_{i=1}^n \theta_i)^2$ . From (5) we can write

$$\mathbf{B}_\theta = \beta + \sigma_y \mathbf{T}_\theta \Phi, \tag{6}$$

where  $\Phi$  is a bivariate standard normal random variable and  $\mathbf{T}_\theta$  is the Cholesky factor of  $\mathbf{V}_\theta$ ; that is,

$$\mathbf{T}_\theta \mathbf{T}_\theta^t = \mathbf{V}_\theta.$$

In this example

$$\mathbf{T}_\theta = \frac{1}{\sqrt{D_\theta}} \begin{pmatrix} \sqrt{\sum_{i=1}^n \theta_i^2} & 0 \\ -\frac{\sum_{i=1}^n \theta_i}{\sqrt{\sum_{i=1}^n \theta_i^2}} & \sqrt{n - \frac{(\sum_{i=1}^n \theta_i)^2}{\sum_{i=1}^n \theta_i^2}} \end{pmatrix}. \tag{7}$$

Equation (6) is a structural equation, which relates the observable random variable  $\mathbf{B}_\theta$  to the model parameters  $\beta$ ,  $\sigma_y$  and the error process  $\mathbf{T}_\theta \Phi$ , whose distribution is fully known (a bivariate normal with mean  $\mathbf{0}$  and a known covariance matrix  $\mathbf{V}_\theta$ ). Consequently, if  $\sigma_y$  is known, a FQ for  $\beta$  can be obtained by solving (6) for  $\beta$  and replacing  $\mathbf{B}_\theta$  with its corresponding realized value  $\mathbf{b}_\theta$ . If  $\sigma_y$  is not known, we need to replace it by a FQ. The FQ for  $\sigma_y$  can be obtained in two different ways. If  $S_y^2$ , obtained from replicate measurements or other means, is available and considered to be an estimator of  $\sigma_y^2$  such that

$$W_y = \frac{\nu_y S_y^2}{\sigma_y^2} \sim \chi^2(\nu_y), \tag{8}$$

then (8) is a structural equation and can be used to obtain a FQ for  $\sigma_y$  as

$$\tilde{\sigma}_y = \frac{s_y}{\sqrt{W_y/\nu_y}}, \tag{9}$$

where  $s_y$  is the realized value of  $S_y$ . If such an estimator is not available, we can use the residuals from the least-squares fit to estimate  $\sigma_y$ . Specifically, let

$$S_\theta^2 = \frac{\sum_{i=1}^n (Y_i - B_0 - B_1 \theta_i)^2}{n - 2};$$

then  $S_\theta^2$  is independent of  $\mathbf{B}_\theta$  and

$$W_r = \frac{(n - 2) S_\theta^2}{\sigma_y^2} \sim \chi^2(n - 2).$$

A FQ for  $\sigma_y$  can be obtained as

$$\tilde{\sigma}_y = \frac{s_\theta}{\sqrt{W_r/(n - 2)}}. \tag{10}$$

Together with (6), we obtain a FQ for  $\beta$  with  $\theta$  fixed as

$$\tilde{\beta}_\theta = \mathbf{b}_\theta - \tilde{\sigma}_y \mathbf{T}_\theta \Phi, \tag{11}$$

where  $\tilde{\sigma}_y$  can be either (9) or (10).

The above development accomplishes the first stage of the construction. For the second stage we need to obtain a FQ for  $\theta$ . From (4) a FQ for  $\sigma_x$  is obtained as

$$\tilde{\sigma}_x = \frac{s_x}{\sqrt{W_x/\nu_x}}. \tag{12}$$

From (3) we can write

$$X_i = \theta_i + \sigma_x Z_i,$$

where  $Z_i$  are independent standard normal random variables. Together with (12), this leads to a FQ for  $\theta_i$  given by

$$\tilde{\theta}_i = x_i - \frac{s_x}{\sqrt{W_x/\nu_x}} Z_i. \tag{13}$$

Finally, let

$$\tilde{\theta} = (\tilde{\theta}_1, \dots, \tilde{\theta}_n);$$

a FQ for  $\beta$  is given by

$$\tilde{\beta} = \mathbf{b}_{\tilde{\theta}} - \tilde{\sigma}_y \mathbf{T}_{\tilde{\theta}} \Phi. \tag{14}$$

The FQ in (14) allows us to generate realizations from the proposed fiducial distribution of  $\beta$ . A single realization of  $\tilde{\beta}$  may be generated as follows.

1. Generate  $Z_i \sim N(0, 1)$ ,  $i = 1, \dots, n$  and  $W_x \sim \chi^2(\nu_x)$ .
2. Compute  $\tilde{\theta}_i$ ,  $i = 1, \dots, n$ , as in (13).
3. Obtain  $\mathbf{b}_{\tilde{\theta}} = (b_0 \ b_1)^t$ , the least-squares solution of  $y_i = b_0 + b_1 \tilde{\theta}_i$  and the residual mean squared error  $s_{\tilde{\theta}}^2$ .
4. Compute  $\tilde{\sigma}_y$  as in (9) or in (10) (need to generate  $W_y \sim \chi^2(\nu_y)$  or  $W_r \sim \chi^2(n - 2)$ ).
5. Compute  $\mathbf{T}_{\tilde{\theta}}$  as in (7).
6. Generate  $\Phi \sim N(\mathbf{0}, \mathbf{I}_2)$ .
7. Compute  $\tilde{\beta}$  as in (14).

A program for generating realizations of  $\tilde{\beta}$ , based on R software [16], is listed in the appendix.

Once the realizations of  $\tilde{\beta}$  are obtained, we can use them to construct a  $1 - \gamma$  fiducial region for  $\beta$ . Let  $\tilde{\beta}_1, \dots, \tilde{\beta}_M$  be the  $M$  independent realizations of  $\tilde{\beta}$ . Let  $\hat{\beta}$  and  $\mathbf{S}$  denote, respectively, the mean vector and covariance matrix computed from  $\tilde{\beta}_i$ ,  $i = 1, \dots, M$ . Let  $d_i$  be the Mahalanobis distance [17, p 595], of  $\tilde{\beta}_i$  from  $\hat{\beta}$ , which is given by

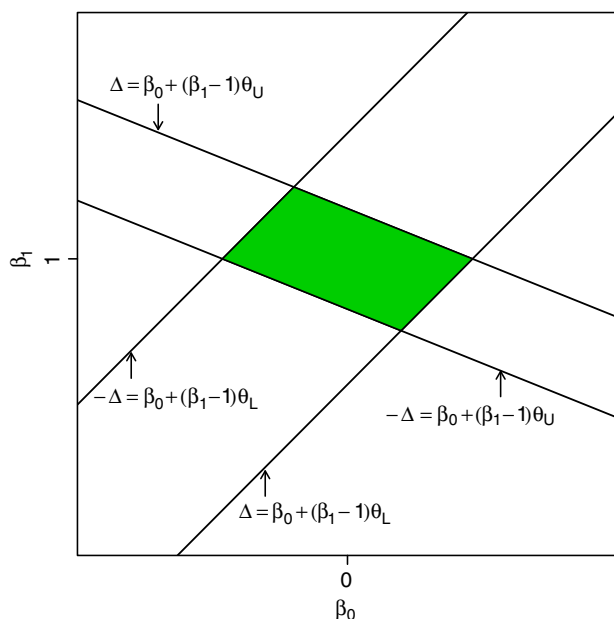
$$d_i = (\tilde{\beta}_i - \hat{\beta})^t \mathbf{S}^{-1} (\tilde{\beta}_i - \hat{\beta}) \tag{15}$$

and let  $d_{1-\gamma}$  denote the  $1 - \gamma$  quantile computed from the collection of  $d_i$ ,  $i = 1, \dots, M$ . A  $1 - \gamma$  fiducial region for  $\beta$  may then be taken to be the region defined by the equation

$$(\beta - \hat{\beta})^t \mathbf{S}^{-1} (\beta - \hat{\beta}) \leq d_{1-\gamma}. \tag{16}$$

A program for constructing fiducial regions for  $\beta$  is provided in the appendix.

One possible approach for assessing agreement between the two methods is based on whether or not the point  $(\beta_0, \beta_1) = (0, 1)$  is contained in the fiducial region. However, this is equivalent to testing the null hypothesis  $(\beta_0, \beta_1) = (0, 1)$ . That is, we assume that the two methods are in agreement to begin with. In this situation, we are determining whether there is evidence to conclude disagreement. This may not be desirable in practice. An alternative approach for assessing agreement can be based on the following idea: if  $(\beta_0, \beta_1)$  is not ‘too different’ from  $(0, 1)$ , then, for practical purposes, the two methods are equivalent. To make this idea precise, a region with  $(0, 1)$  as the centre is specified. If  $(\beta_0, \beta_1)$  is within this



**Figure 1.** The equivalence region for  $(\beta_0, \beta_1)$  determined by the maximum allowable difference  $\Delta$ .

region then the methods are regarded as in agreement. We call this region the *equivalence* region. The equivalence region is specified by the user. It can be an ellipse, parallelogram, rectangle, or a region of some other appropriate shape. The way we use the fiducial region in this method is as follows. If the  $1 - \gamma$  fiducial region that we construct is *completely inside* the equivalence region then we have established agreement. Note that the equivalence region needs to be specified at the time of problem formulation and especially before the fiducial region is obtained.

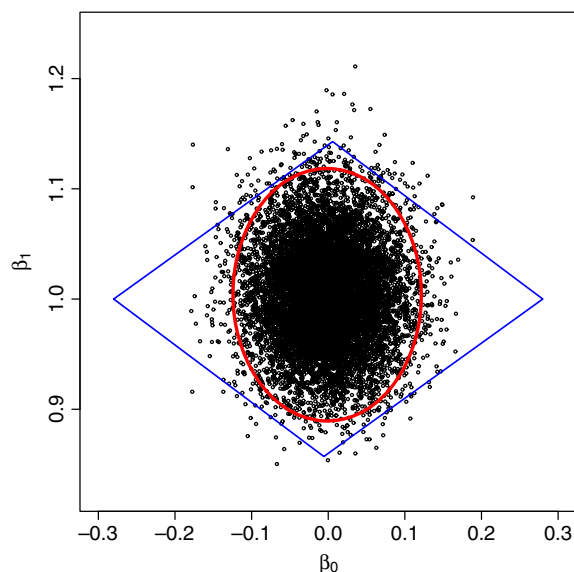
An equivalence region can be constructed based on the *maximum allowable difference* of the two equivalent methods. If the maximum allowable difference is  $\Delta$  then  $|\mu_i - \theta_i|$  must be less than or equal to  $\Delta$  for all  $i = 1, \dots, n$  in order for the two methods to be regarded as in agreement. To develop an equivalence region for a given value of  $\Delta$ , we determine a region such that if  $(\beta_0, \beta_1)$  is in this region, then

$$|\mu_i - \theta_i| = |\beta_0 + (\beta_1 - 1)\theta_i| \leq \Delta, \quad i = 1, \dots, n. \quad (17)$$

Suppose that the values  $\theta_i, i = 1, \dots, n$  are expected to lie in a known interval  $[\theta_L, \theta_U]$ . The region determined by the inequalities in (17) is the same as that determined by the following two inequalities:

$$\begin{aligned} -\Delta &\leq \beta_0 + (\beta_1 - 1)\theta_L \leq \Delta \\ -\Delta &\leq \beta_0 + (\beta_1 - 1)\theta_U \leq \Delta. \end{aligned}$$

The points in the  $(\beta_0, \beta_1)$  plane that satisfy these two inequalities belong to a parallelogram having the vertices  $(-\Delta, 1), (\Delta(\theta_L + \theta_U)/(\theta_U - \theta_L), 1 - 2\Delta/(\theta_U - \theta_L)), (\Delta, 1)$  and  $(-\Delta(\theta_L + \theta_U)/(\theta_U - \theta_L), 1 + 2\Delta/(\theta_U - \theta_L))$ . A parallelogram with  $\theta_L < 0$  and  $\theta_U > 0$  is shown as the shaded region in figure 1.



**Figure 2.** Realizations from the fiducial distribution of  $(\beta_0, \beta_1)$ , a 95% fiducial region for  $(\beta_0, \beta_1)$  and an equivalence region.

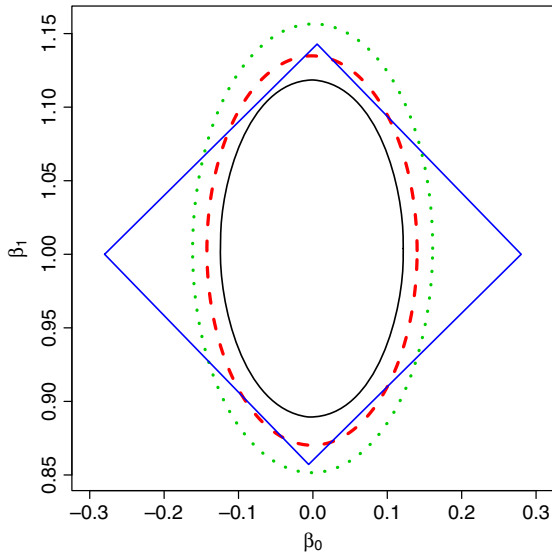
### 4. Examples

The examples are taken from [8]. Comparison studies were carried out for two angle-measuring instruments. The standard uncertainties for both instruments have been determined to be  $0.1''$ . The comparisons were based on the angle measurements of 8-sided and 12-sided polygons. Since the analysis is identical for both polygons, we present only the results corresponding to the 8-sided polygon. Deviations (in seconds) from the nominal value of the angles of the 8-sided polygon, measured by the two instruments, are given below:

$$\begin{aligned} x: & -0.11 \ 0.25 \ -2.00 \ -0.76 \ 1.92 \ 0.88 \ -0.18 \ 0.01 \\ y: & -0.22 \ 0.31 \ -2.00 \ -0.81 \ 1.92 \ 0.90 \ -0.20 \ 0.10 \end{aligned}$$

We use the program listed in the appendix to generate 100 000 realizations from the proposed fiducial distribution of  $(\beta_0, \beta_1)$ . We use  $\hat{\sigma}_x = \hat{\sigma}_y = 0.1''$ . Since no information was given in [8] concerning the degrees of freedom associated with  $\hat{\sigma}_x$  and  $\hat{\sigma}_y$ , we assume, for illustration only, infinite degrees of freedom for them. Figure 2 displays these realizations. Also, a 95% fiducial region for  $(\beta_0, \beta_1)$  and an equivalence region are shown. The equivalence region is obtained from  $\theta_L = -2'', \theta_U = 1.92''$  and  $\Delta = 0.28''$ . The maximum allowable difference is chosen to be the expanded uncertainty of the difference, i.e.  $2\sqrt{0.1^2 + 0.1^2}$ . This value of  $\Delta$  is chosen just for illustration. In practice, the user will have to take the specific application into consideration before deciding how to pick the maximum allowable difference. Since the 95% fiducial region is completely inside the equivalence region, we conclude that the two instruments are in agreement.

A larger fiducial region will result if finite degrees of freedom associated with  $\hat{\sigma}_x$  and/or  $\hat{\sigma}_y$  are used. Figure 3 plots three fiducial regions corresponding to three different sets of values on  $\hat{\sigma}_x, \hat{\sigma}_y, \nu_x$  and  $\nu_y$  used in the analysis. The inner ellipse corresponds to  $\hat{\sigma}_x = \hat{\sigma}_y = 0.1''$  and  $\nu_x = \nu_y = \infty$ .



**Figure 3.** 95% fiducial regions for  $(\beta_0, \beta_1)$  based on different uncertainty estimates used and an equivalence region.

The middle ellipse corresponds to  $\hat{\sigma}_x = \hat{\sigma}_y = 0.1''$  and  $\nu_x = \nu_y = 10$ . The outer ellipse corresponds to  $\hat{\sigma}_x = 0.1''$ ,  $\nu_x = \infty$  and  $\hat{\sigma}_y$  computed from the residual mean squares error with  $8 - 2 = 6$  degrees of freedom. All three fiducial regions contain  $(0, 1)$ . This reveals that the approach, based on whether  $(0, 1)$  is within the fiducial region, is not a desirable way for assessing agreement. A large, and hence not informative, fiducial region that includes the point  $(0, 1)$  does not provide evidence of agreement. Nor does it provide evidence of disagreement. On the other hand, the approach based on comparing the fiducial region with an equivalence region provides a useful procedure for assessing agreement; it requires that agreement be demonstrated. Figure 3 shows that only the inner fiducial region is completely inside the equivalence region, indicating that the quality of the uncertainty estimates does affect the conclusion of the analysis in this example.

Given  $\Delta$ , the vertices of the equivalence region constructed using the maximum allowable difference depend only on the range  $\theta_U - \theta_L$  and the mid-range  $(\theta_L + \theta_U)/2$  of  $\theta_i$ ,  $i = 1, \dots, n$ . If such information is available either from the previous studies or from other requirements and considerations, they may be used to derive the equivalence region.

The smallest rectangle that contains the fiducial region may also be obtained to aid in assessing agreement. This rectangle is formed by the two vertical and two horizontal tangent lines of the elliptical fiducial region. The two vertical tangent lines result in a fiducial interval for  $\beta_0$  and the two horizontal tangent lines result in a fiducial interval for  $\beta_1$ . These intervals provide a convenient information on how  $(\beta_0, \beta_1)$  differs from  $(0, 1)$ . For this example, the projected fiducial intervals for  $\beta_0$  and  $\beta_1$  are  $(-0.1246, 0.1219)$  and  $(0.8897, 1.1189)$ , respectively.

## 5. Conclusion

In this paper we have provided an approach for making inference on the intercept  $\beta_0$  and slope  $\beta_1$  of a linear regression model with both  $X$  and  $Y$  subject to measurement errors. Specifically, we have provided procedures for constructing uncertainty regions for  $(\beta_0, \beta_1)$  that can be used to assess agreement between two methods. The approach is based on fiducial inference. Recent research results [15, 18] show that, under certain regularity conditions, fiducial inference is a valid statistical method with good operating characteristics.

We described a fiducial recipe based on a two-stage construction of a fiducial quantity. This recipe is particularly useful in the general calibration problems. The statistical issues related to the fiducial inference of general calibration problems will be discussed in forthcoming papers.

## Acknowledgments

The authors are grateful to the referees for their valuable suggestions. This work is a contribution of the National Institute of Standards and Technology and is not subject to copyright in the United States.

## Appendix

We list two R functions that were used in this paper. The first one is an R function `fab` that generates realizations from the fiducial distribution of  $(\beta_0, \beta_1)$ .

```
fab <- function(X, Y, ux, dfx, uy=NA,
               dfy=NA, nrun=10000) {
  # this program generates nrun realizations
  # from the joint fiducial distribution of
  # (beta0, beta1) of the following model
  # Y_i = beta0 + beta1*theta_i + epsilon_i
  # X_i = theta_i + delta_i
  #   epsilon_i ~ N(0, sigma_y^2)
  #   delta_i ~ N(0, sigma_x^2)
  #   dfx * ux^2/sigma_x^2 ~ chi^2(dfx)
  #   dfy * uy^2/sigma_y^2 ~ chi^2(dfy)
  #
  # input
  # X, Y - measurements
  # ux - estimate of sigma_x
  # dfx - degrees of freedom for ux
  #       use negative number for infinity
  # uy - estimate of sigma_y (optional)
  # dfy - degrees of freedom for uy
  # nrun - number of desired realizations
  #
  nobs = length(X)
  if (dfx < 0) Wx = ux
  else Wx = ux/sqrt(rchisq(nrun, dfx)/dfx)
  Ei = matrix(rnorm(nobs*nrun), ncol=nobs)
  Ei = t(Wx * Ei)
  thetai = X - Ei
  out = apply(thetai, 2,
```

```

function(x, y){lsfit(x, y)}, y=Y)
out = unlist(out)
names(out) = NULL
Acoef = out[seq(from=1, by=9+4*nobs,
length=nrun)]
Bcoef = out[seq(from=2, by=9+4*nobs,
length=nrun)]
# if uy is not provided, use residual
# mean squares
if (is.na(uy) | is.na(dfy)) {
  Ti = (1:(nobs*nrun))
  for (i in (1:nrun))
    Ti[(1+(i-1)*nobs):(i*nobs)] =
      (3+(i-1)*(9+4*nobs)):
      (3+(i-1)*(9+4*nobs)+nobs-1)
  Resid = matrix(out[Ti], ncol=nrun)
  sse = apply(Resid, 2,
function(x){sqrt(sum(x^2))})
  sigmaY = sse/sqrt(rchisq(nrun, nobs-2))
}
# if uy and dfy are provided
else {
  if (dfy < 0) sigmaY = uy
  else sigmaY = uy/sqrt(rchisq(nrun,
dfy)/dfy)
}
sT = apply(thetai, 2, sum)
sT2 = apply(thetai, 2,
function(x){sum(x^2)})
tempT = sqrt(nobs*sT2 - sT^2)
Z1 = rnorm(nrun)
Z2 = rnorm(nrun)
beta0 = Acoef - sigmaY*sqrt(sT2)*Z1/tempT
beta1 = Bcoef - sigmaY*(-sT*Z1/sqrt(sT2) +
sqrt(nobs - sT^2/sT2)*Z2)/tempT
list(beta0=beta0, beta1=beta1)
}

```

The function has seven mandatory and optional arguments:

1. Measurements for  $x$ .
2. Measurements for  $y$ .
3. Value of  $S_x$ , the estimate of  $\sigma_x$ .
4. Degrees of freedom associated with  $S_x^2$ . Use negative number for infinity.
5. Value of  $S_y$ , the estimate of  $\sigma_y$ . This is optional. If it is not provided, the residual mean squares from the least-squares fit will be used as the estimate of  $\sigma_y^2$ .
6. Degrees of freedom associated with  $S_y^2$ . This is optional.
7. Number of fiducial samples desired. The default is 10 000.

The output contains the desired number of samples from the fiducial distribution of  $(\beta_0, \beta_1)$ . With this function, the following commands may be used to generate 100 000 realizations for the example in section 4:

```

> X = c(-0.11, 0.25, -2.00, -0.76,
1.92, 0.88, -0.18, 0.01)
> Y = c(-0.22, 0.31, -2.00, -0.81,
1.92, 0.90, -0.20, 0.10)
> fs = fab(X, Y, 0.1, -1, 0.1, -1, 100000)

```

The output  $fs$  consists of two components:  $\beta_0$  and  $\beta_1$ . Once the fiducial samples are obtained, we may use the following R function `fregion` to compute a 95% fiducial region for  $(\beta_0, \beta_1)$ .

```

fregion <- function(fs, gamma=.95, n=200) {
# generate n points on a 100*gamma%
# fiducial region based on fiducial
# samples in fs
#
Z = cbind(fs[[1]], fs[[2]])
mean.vector = apply(Z, 2, mean)
cov.mat = cov(Z)
dval = quantile(mahalanobis(Z,
mean.vector, cov.mat), gamma)
a = mean.vector[1]
b = mean.vector[2]
ct = cov.mat
angle = atan(-2*ct[1,2], (ct[2,2] -
ct[1,1]))/2
temp = 2*dval*(ct[1,1]*ct[2,2] -
ct[1,2]^2)
am = sqrt(temp/(ct[2,2] + ct[1,1] -
2*ct[1,2]/sin(2*angle)))
an = sqrt(temp/(ct[2,2] + ct[1,1] +
2*ct[1,2]/sin(2*angle)))
major = max(am, an)
minor = min(am, an)
if(am < an) angle = angle - pi/2
phi = seq(from=0, to=2*pi, length=n)
rr = sqrt(major^2/(1 + ((major/minor)^2
- 1) * (sin(phi - angle))^2))
xx = rr * cos(phi) + a
yy = rr * sin(phi) + b
XY = list(x=xx, y=yy)
XY
}

```

Figure 2 in section 4 was produced using the following commands:

```

> plot(fs$beta0, fs$beta1,
xlab=expression(beta[0]),
ylab=expression(beta[1]))
> fellipse = fregion(fs)
> lines(fellipse, col=2)
> polygon(c(-0.28, -0.006, 0.28, 0.006),
c(1, 0.857, 1, 1.143),
density=0, col=4)

```

The first command plots the realizations. The second command calls `fregion` to compute 200 points on a 95% fiducial region based on the fiducial samples  $fs$ . The third command plots the fiducial region in red colour. The fourth command adds the equivalence region with vertices  $(-0.28, 1)$ ,  $(-0.006, 0.857)$ ,  $(0.28, 1)$  and  $(0.006, 1.143)$ , in blue colour, to the plot.

The vertical and horizontal tangent lines of the fiducial region are obtained from the calculated values in function `fregion`. Specifically, the  $\beta_0$ -coordinates of the two vertical tangent lines are

$$a - \sqrt{\text{minor}^2 * \sin(\text{angle})^2 + \text{major}^2 * \cos(\text{angle})^2}$$

and

$$a + \sqrt{\text{minor}^2 * \sin(\text{angle})^2 + \text{major}^2 * \cos(\text{angle})^2}$$

while the  $\beta_1$ -coordinates of the two horizontal tangent lines are

$$b - \sqrt{\text{major}^2 * \sin(\text{angle})^2 + \text{minor}^2 * \cos(\text{angle})^2}$$

and

$$b + \sqrt{\text{major}^2 * \sin(\text{angle})^2 + \text{minor}^2 * \cos(\text{angle})^2}$$

## References

- [1] Altman D G and Bland J M 1983 *Statistician* **32** 307–17
- [2] Bland J M and Altman D G 1986 *Lancet* **i** 307–10
- [3] Altman D G and Bland J M 1987 *Appl. Statistics* **36** 224–5
- [4] Bland J M and Altman D G 1995 *Lancet* **346** 1085–7
- [5] Bland J M and Altman D G 1999 *Stat. Methods Med. Res.* **8** 135–60
- [6] Lin L, Hedayat A S, Sinha B and Yang M 2002 *J. Am. Stat. Assoc.* **97** 257–70
- [7] Choudhary P K 2008 *J. Stat. Planning Inference* **138** 1102–15
- [8] Astrua M, Ichim D, Penneccchi F and Pisani M 2007 *Metrologia* **44** 385–92
- [9] Fisher R A 1930 *Proc. Camb. Phil. Soc.* **26** 528–35
- [10] Fraser D A S 1968 *The Structure of Inference* (New York: Wiley)
- [11] Wang C M and Iyer H K 2005 *Metrologia* **42** 145–53
- [12] Wang C M and Iyer H K 2006 *Measurement* **39** 856–63
- [13] Wang C M and Iyer H K 2006 *Metrologia* **43** 486–94
- [14] Hannig J, Iyer H K and Wang C M 2007 *Metrologia* **44** 476–83
- [15] Hannig J, Iyer H K and Patterson P D 2006 *J. Am. Stat. Assoc.* **101** 254–69
- [16] R Development Core Team 2003 *R: A Language and Environment for Statistical Computing* (Vienna, Austria: R Foundation for Statistical Computing) ISBN 3-900051-00-3, <http://www.R-project.org>
- [17] Rao C R 1973 *Linear Statistical Inference and its Applications* 2nd edn (New York: Wiley)
- [18] Hannig J 2006 On fiducial inference *Technical Report* [http://www.stat.colostate.edu/research/2006\\_3.pdf](http://www.stat.colostate.edu/research/2006_3.pdf)