

Research Article

An Improved Measurement Error Model for Analyzing Unreplicated Method Comparison Data under Asymmetric Heavy-Tailed Distributions

Jeevana Duwarahan ^{1,2} and Lakshika S. Nawarathna ³

¹Postgraduate Institute of Science, University of Peradeniya, Peradeniya, Sri Lanka

²Department of Mathematics and Statistics, University of Jaffna, Jaffna, Sri Lanka

³Department of Statistics and Computer Science, Faculty of Science, University of Peradeniya, Peradeniya, Sri Lanka

Correspondence should be addressed to Lakshika S. Nawarathna; lakshikas@sci.pdn.ac.lk

Received 22 August 2022; Revised 28 October 2022; Accepted 16 November 2022; Published 15 December 2022

Academic Editor: Muhammad Ahsan

Copyright © 2022 Jeevana Duwarahan and Lakshika S. Nawarathna. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Method comparison studies mainly focus on determining if the two methods of measuring a continuous variable are agreeable enough to be used interchangeably. Typically, a standard mixed-effects model uses to model the method comparison data that assume normality for both random effects and errors. However, these assumptions are frequently violated in practice due to the skewness and heavy tails. In particular, the biases of the methods may vary with the extent of measurement. Thus, we propose a methodology for method comparison data to deal with these issues in the context of the measurement error model (MEM) that assumes a skew- t (ST) distribution for the true covariates and centered Student's t (cT) distribution for the errors with known error variances, named STcT-MEM. An expectation conditional maximization (ECM) algorithm is used to compute the maximum likelihood (ML) estimates. The simulation study is performed to validate the proposed methodology. This methodology is illustrated by analyzing gold particle data and then compared with the standard measurement error model (SMEM). The likelihood ratio (LR) test is used to identify the most appropriate model among the above models. In addition, the total deviation index (TDI) and concordance correlation coefficient (CCC) were used to check the agreement between the methods. The findings suggest that our proposed framework for analyzing unreplicated method comparison data with asymmetry and heavy tails works effectively for modest and large samples.

1. Introduction

Evaluation of the two methods for measuring a continuous response variable attracts greater attention in health science such as biomedical engineering, clinical research, and medical imaging. The methods may include an assay, medical device, clinical observer, and measurement technique, and the variables of interest, e.g., blood pressure, heart rate, level of cholesterol, and the concentration of the chemical. Generally, the new methods are compared with already established methods to identify sufficient agreement between them. With so many advancements in the field of medical sciences, new measurement methods and techniques are available that

may be cheaper, faster, easier to use, and less invasive. Before using these new measurement methods, the accuracy and precision must be confirmed. Therefore, detailed research in this sector will enable health professionals to choose the most appropriate and effective treatment method. Suppose the study reveals a satisfactory agreement between the methods, it could either be used interchangeably or the most appropriate method is selected. These studies are widespread in the research of health sciences. The Web of Knowledge citation database now has over 50,000 citations for Bland and Altman [1]. They proposed the limits of agreement methodology for evaluating the agreement between two methods, which is a testimony to the above.

In method comparison studies, each method generally takes measurements on each subject. At times the measurements may be replicated. The data from the same subject are considered dependent, but data from different subjects are considered independent. A two-step technique may be used to analyze these data. Modeling the method comparison data is the first step. For this purpose, the mixed-effects model [2] is commonly used and assumes an independent normal distribution for both random effects and errors when the variability of the measurement remains constant over the entire measurement range [3–9]. The second step is the agreement evaluation between the methods. The agreement evaluation is performed on one or more measures of agreement that indicate how much of these methods agree with themselves. Slight differences in measurements refer to a good agreement between the two methods. In the literature, there are several agreement measures available for evaluating the agreement between two methods, including limits of agreement [1, 3, 4, 10], CCC [11–14], TDI [9], coverage probability (CP) or tolerance interval [12–16], mean squared deviation (MSD) [17, 18], and coefficient of individual agreement [11, 17].

A linear mixed-effects model is typically employed when comparing a novel approach to an existing reference or standard. However, it is important to note that this method cannot be used when the biases of the methods vary with the magnitude of measurement [3, 19–23]. In the case of the above-nature data, a MEM [24] should be used instead of a mixed-effects model. In the literature, the majority of the studies mentioned above are under the normal distribution. However, method comparison data frequently exhibit skewness and heavy tails in practice, meaning tails that are longer than those of a normal distribution, illustrated by analyzing a real dataset based on a method comparison study by Tomaya and de Castro [25] and is discussed later in this article. For this scenario, data transformation is feasible to ensure that the normality assumption is met. Nonetheless, it may make it difficult to interpret the differences in measurements between the two methods. It is a common issue in method comparison data analysis. To overcome this problem, some alternative approaches have been considered [26–34].

Recently, Choudhary et al. [35] developed a general skew- t (GST) mixed model that assumes an ST distribution for the random effects and an independent multivariate t distribution for the errors. Later, Sengupta et al. [36] developed this methodology to analyze the method comparison data with skewness and heavy tails with unknown error variances. Here, they have developed a methodology to assess how well the methods agree when measuring in the same nominal unit. This means that the true (error-free) values of the method may differ only by a constant. The above models cannot be used when the methods have different measurement scales/methods. When collecting data, various measuring scales might lead to measurement errors in covariates and response variables. There would be some change in the statistical inferences if these errors were not taken into consideration. As a result, compared to a mixed-effects model, the MEM provides a more adaptable framework for modeling method comparison data.

The study of method comparison under heavy-tailed distributions has not received much attention in the literature because of the complexity of the likelihood function. Recently, Cao et al. [37] proposed MEM for replicated data under asymmetric and heavy-tailed distributions with the same degrees of freedom for true covariate and error terms. Further, this model is unable to account for different degrees of heaviness in the tails of true covariates and error distributions. Further, Tomaya and de Castro [25] developed STcT-MEM that assumes an ST distribution for the true covariate and a cT distribution [38] for the error terms with known error variances and considered the different degrees of freedom for true covariate and errors. In this paper, our main goal is to adapt the above model to unreplicated method comparison data with different levels of heaviness in the tails of true covariates and errors, especially if we know the error variances. This approach will enable us to model the method comparison data with better flexibility and higher accuracy, accommodating skewness and heavy tails.

The rest of the paper is set out as follows. In Section 2, we present the STcT-MEM for method comparison data. Section 3 deals with the proposed methodology for the evaluation of the agreement under STcT-MEM. Section 4 explores the performance of the proposed model using simulation studies. Section 5 provides an application utilizing data on gold particles to illustrate our methodology, and the final section discusses the findings and conclusions. All calculations presented in this paper were carried out using the R programming language [39].

2. Modeling of Method Comparison Data

This section outlines an approach for analyzing studies that compare two methods that use single measurements on each subject, implying that the measurements are not replicated. The measurement of the j^{th} method on the i^{th} subject is denoted by Y_{ij} , $j = 1, 2$; $i = 1, \dots, n$. Here, n is the number of subjects in the study. The standard method is assumed to be Method 1, and the test method is supposed to be Method 2.

2.1. An Overview of STcT-MEM. This section briefly describes the STcT-MEM in general terms before being presented for method comparison data. The details can be found in Tomaya and de Castro [25]. In this article, we use boldface letters to refer to vectors and matrices. Let $\mathcal{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $S\mathcal{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\lambda}), t_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, v)$, and $ST_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\lambda}, v)$, respectively, denote p -dimensional normal, skew-normal, t , and ST distributions. Here, $\boldsymbol{\mu} \in \mathbb{R}^p$ is a location vector, $\boldsymbol{\Sigma}$ is a $p \times p$ positive-definite scale matrix, $\boldsymbol{\lambda} \in \mathbb{R}^p$ is a vector of skewness parameters, and $v (> 0)$ is the degrees of freedom. Let $\mathcal{G}(\alpha, \beta)$ denote the gamma distribution with parameters $\alpha, \beta > 0$. Furthermore, we use $\boldsymbol{\Sigma}^{1/2}$ to denote the square root of a symmetric, positive-definite matrix $\boldsymbol{\Sigma}$ so that $\boldsymbol{\Sigma}^{1/2} \boldsymbol{\Sigma}^{1/2T} = \boldsymbol{\Sigma}$, where $\boldsymbol{\Sigma}^{1/2T}$ is the transpose of $\boldsymbol{\Sigma}^{1/2}$, $\boldsymbol{\Sigma}^{-(1/2)}$ is the inverse of $\boldsymbol{\Sigma}^{(1/2)}$, and $\det(\boldsymbol{\Sigma})$ is the determinant of $\boldsymbol{\Sigma}$.

Let $ct_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, v)$ denote a p -variate cT distribution if its probability density function (pdf) is given by

$$f_p(\mathbf{z}; \boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu) = K(\nu, p) \left[\det(\boldsymbol{\Sigma})^{-1/2} \right] \left[(\nu - 2) + (\mathbf{z} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{z} - \boldsymbol{\mu}) \right]^{-((\nu+p)/2)}; \mathbf{z} \in \mathbb{R}^p, \quad (1)$$

where $\boldsymbol{\mu} \in \mathbb{R}^p$ is a mean vector, $\boldsymbol{\Sigma}$ is a $p \times p$ covariance matrix, $K(\nu, p)$ is the normalizing constant given by $((\nu - 2)^{(v/2)} \text{gam}((\nu + p)/2) / (\pi^{(p/2)} \text{gam}(\nu/2)))$; $\nu > 2$, and $\text{gam}(\cdot)$ denotes the gamma function. It is a centered parametric version of the Student's t distribution, where the parameters are the mean vector and covariance matrix, whereas, in the usual parameterization, they are the location vector and the scale matrix. If the model assumes that the variances of the errors are known, the distribution that would best support this variances assumption is the centered version. A brief introduction about SN , t , ST can be seen in Appendix A, and detailed information about these distributions can be found in Azzalini and Capitanio [40] and Azzalini and Capitanio [33].

The SMEM can be written as

$$\begin{aligned} X_i &= b_i + u_i, \\ Y_i &= y_i + \varepsilon_i, \\ y_i &= \beta_0 + \beta_1 b_i, \quad i = 1, \dots, n, \end{aligned} \quad (2)$$

where β_0 and β_1 are the intercept and slope parameters, respectively, b_i and y_i are unobserved true covariate and unobserved true response variables, respectively, X_i and Y_i are the observed variables, u_i and ε_i are the error terms, and n is the sample size. Model (2) can be written as

$$\mathbf{Z}_i = \mathbf{A} + \mathbf{B}b_i + \mathbf{e}_i, \quad i = 1, \dots, n, \quad (3)$$

where $\mathbf{Z}_i = (X_i, Y_i)^T$, $\mathbf{A} = (0, \beta_0)^T$, $\mathbf{B} = (1, \beta_1)^T$, and $\mathbf{e}_i = (u_i, \varepsilon_i)^T$. It is standard to assume that b_i and \mathbf{e}_i are independent and

$$\begin{aligned} b_i &\sim N(\xi, \psi), \\ \mathbf{e}_i &\sim N_2(\mathbf{0}, \boldsymbol{\Sigma}_i). \end{aligned} \quad (4)$$

Normality assumption is sometimes unfeasible due to the skewness, heavy-tailed ness, and outliers. Thus, Tomaya and de Castro [25] developed the STcT-MEM with more general distributions as follows:

$$\begin{aligned} b_i &\sim ST(\xi, \psi, \lambda, \eta_b), \\ \mathbf{e}_i &\sim ct_2(\mathbf{0}, \boldsymbol{\Sigma}_i, \eta_e), \end{aligned} \quad (5)$$

where b_i and \mathbf{e}_i are mutually independent. Inverse transformations have been considered for the degrees of freedom to enhance the inference process. The hierarchical representations of b_i and \mathbf{e}_i are defined in Appendix B. Next, (ψ, λ) is reparameterized as (Γ, γ) , where

$$\begin{aligned} \delta &= \frac{\lambda}{(1 + \lambda^T \lambda)^{(1/2)}}, \\ \gamma &= \psi^{(1/2)} \delta, \\ \Gamma &= \psi - \gamma \gamma^T. \end{aligned} \quad (6)$$

Then, the mean vector and covariance matrix of \mathbf{Z}_i are

$$\begin{aligned} E(\mathbf{Z}_i) &= \mathbf{A} + \mathbf{B}\mu_b, \quad 0 < \eta_b, \eta_e < 1, \\ \text{Var}(\mathbf{Z}_i) &= \left(\frac{\Gamma + \gamma^2}{1 - 2\eta_b} - \zeta^2 \gamma^2 \right) \mathbf{B}\mathbf{B}^T + \boldsymbol{\Sigma}_i, \quad 0 < \eta_b, \eta_e < \frac{1}{2}, \end{aligned} \quad (7)$$

where $\mu_b = \xi + \zeta\gamma$, with $\zeta = [\text{gam}((1 - \eta_b)/2\eta_b) / \text{gam}(1/2\eta_b)] / \sqrt{\eta_b\pi}$.

Since the pdf of \mathbf{Z}_i is not in a closed form, the one-dimensional numerical integration is used to solve this issue, which is explained in Tomaya and de Castro [25] and Choudhary et al. [35]. This can be carried out by using the *numDeriv* package [41] in R. Furthermore, due to the complexity of the log-likelihood function, Tomaya and de Castro [25] used the ECM algorithm [42] to estimate the parameters. It is a variant of the expectation-maximization (EM) algorithm.

2.2. STcT-MEM for Method Comparison Data. It follows from (3) that the model for the paired measurements (Y_{i1}, Y_{i2}) can be written as

$$\begin{aligned} Y_{i1} &= b_i + e_{i1}, \\ Y_{i2} &= \beta_0 + \beta_1 b_i + e_{i2}, \\ i &= 1, \dots, n, \end{aligned} \quad (8)$$

where β_0 and β_1 are fixed regression coefficients known as fixed bias and proportional bias of method 2, respectively, b_i denote the true unobservable measurement for the i^{th} subject, and e_{ij} is the random error of the j^{th} method ($j = 1, 2$) on the i^{th} subject. The methods are scaled differently in this case, and the methods have the same scale if $\beta_1 = 1$. In this case, the model reduces to a mixed-effects model discussed by Sengupta et al. [36].

Further, b_i and e_{ij} are mutually independent, and we assume that

$$b_i \sim ST(\xi, \psi, \lambda, \eta_b), \quad (9)$$

$$e_{i1}, e_{i2} \sim ct_2(\mathbf{0}, \boldsymbol{\Sigma}, \eta_e),$$

where $\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix}$ and assumed as known.

The SMEM becomes a special case of the STcT-MEM (5) when the skewness parameter $\lambda = 0$, and the degree of freedom parameters $\eta_b, \eta_e \rightarrow 0$.

3. Evaluation of Agreement

The evaluation of agreement in a method comparison study examines the joint distribution of the method, and the evaluation of similarity is a comparison of the marginal characteristics of the measurement methods, such as their biases and precisions. Let $(\widetilde{Y}_1, \widetilde{Y}_2)$ indicate a pair of observations taken using two methods on a subject selected at random from a target population. The closeness of the two

methods' measurements is referred to as agreement. When the methods have equal means, variances, and correlation one, they have a perfect agreement. Here, the bivariate distributions of \widetilde{Y}_1 and \widetilde{Y}_2 are concentrated on the 45° line.

To quantify the extent of agreement, we first determine how far apart the paired measures are from the line of equality. This is performed through measures of agreement. Several agreement measures can be found in the literature, such as limits of agreement, CCC, TDI, and MSD. Here, we consider only CCC and TDI to evaluate the agreement between the methods. The CCC was proposed by Lin [12], and it is defined as

$$\text{CCC} = \frac{2\text{cov}(\widetilde{Y}_1, \widetilde{Y}_2)}{\{E(\widetilde{Y}_1) - E(\widetilde{Y}_2)\}^2 + \text{Var}(\widetilde{Y}_1) + \text{Var}(\widetilde{Y}_2)}. \quad (10)$$

The CCC ranges from -1 to $+1$. A large positive value of CCC indicates good agreement. The value of 1 implies perfect positive agreement, and the value of -1 represents perfect negative agreement. Detailed information on this measure can be found in Barnhart et al. [11] and Carrasco and Jover [5]. The TDI was proposed by Lin [18], and it is defined as

$$\text{TDI}(p) = 100p^{\text{th}} \text{percentile} |\widetilde{D} = \widetilde{Y}_1 - \widetilde{Y}_2| \text{ for a specified } p. \quad (11)$$

Generally, p varies from 0.80 to 0.95 . It is nonnegative, and a small value indicates high agreement between the methods, and it is perfect when $\text{TDI} = 0$. It has been used by Lin [18] and Choudhary [9, 15]. The TDI can be calculated by solving the following equation:

$$P\{|\widetilde{D}| \leq \text{TDI}(p)\} = p. \quad (12)$$

We often employ one-sided confidence intervals for agreement measurements to evaluate the agreement. It is possible to choose a lower or upper confidence limit. We can compute the $100(1 - \alpha)\%$ upper confidence limit, where a small value (nearing zero) for an agreement measure that indicates good agreement, for example, TDI. Similarly, we may compute a $100(1 - \alpha)\%$ lower confidence limit, such as CCC, where a large positive value (nearing one) for an agreement measure indicates good agreement [43]. Let $\mathbf{I} = -\partial^2 \log\{L(\boldsymbol{\theta})\}/\partial\boldsymbol{\theta}^2$ denote the observed information matrix [14, 35, 43] of $\boldsymbol{\theta}$ (model parameter vector) evaluated at ML estimates. Note that $L(\boldsymbol{\theta})$ is the likelihood function. It can be computed using numerical differentiation techniques. When n is large, $\boldsymbol{\theta}$ can be approximated by a normal distribution with mean $\boldsymbol{\theta}$ and variance \mathbf{I}^{-1} , according to the large sample theory. Next, let φ be a scalar measure of agreement between the two methods. Its ML estimator $\widehat{\varphi}$ is obtained by substituting $\boldsymbol{\theta}$ with $\widehat{\boldsymbol{\theta}}$. From the large sample theory, the sampling distribution of $\widehat{\varphi}$ can be approximated as

$$\widehat{\varphi} \sim N(\varphi, \mathbf{D}^T \mathbf{I}^{-1} \mathbf{D}), \quad (13)$$

where $\mathbf{D} = \partial\varphi/\partial\boldsymbol{\theta}$ is the Jacobian matrix evaluated at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$.

Then,

$$\begin{aligned} \widehat{\varphi} + c_{1-\alpha} SE(\widehat{\varphi}) &\text{ as the } 100(1 - \alpha)\% \text{ upper confidence limit for } \varphi, \\ \widehat{\varphi} - c_{1-\alpha} SE(\widehat{\varphi}) &\text{ as } 100(1 - \alpha)\% \text{ lower confidence limit for } \varphi, \end{aligned} \quad (14)$$

where $c_{1-\alpha}$ is the critical point and $SE(\widehat{\varphi}) = (\mathbf{D}^T \mathbf{I}^{-1} \mathbf{D})^{(1/2)}$. These confidence limits are generated by applying Fisher's z -transformation to CCC and the log transformation to TDI, then inverting the findings back to the original scale for greater accuracy in the estimate.

3.1. CCC and TDI under STcT-MEM. To define CCC and TDI, first, note that the hierarchical representation of $\widetilde{\mathbf{Y}} = (\widetilde{Y}_1, \widetilde{Y}_2)$ from Appendix B is

$$\begin{aligned} \widetilde{\mathbf{Y}}|U, V &\sim SN_2(\mathbf{A} + \mathbf{B}\xi, \widetilde{\boldsymbol{\Pi}}_V/U, \widetilde{\boldsymbol{\lambda}}_V), \\ U &\sim \mathcal{G}\left(\frac{1}{2\eta_b}, \frac{1}{2\eta_b}\right), \\ \frac{U}{V} &\sim \mathcal{G}\left(\frac{1}{2\eta_e}, \frac{1}{2c(\eta_e)}\right), \end{aligned} \quad (15)$$

where $\mathbf{A} = (0, \beta_0)^T$, $\mathbf{B} = (1, \beta_1)^T$, $\widetilde{\boldsymbol{\Pi}}_V$ and $\widetilde{\boldsymbol{\lambda}}_V$ are counterparts of $\boldsymbol{\Pi}_V$ and $\boldsymbol{\lambda}_V$ from Appendix B, and $\mathcal{G}(\alpha, \beta)$ denotes gamma distribution with parameters $\alpha, \beta > 0$.

The mean and variance of $\widetilde{\mathbf{Y}}$ can be represented as $E(\widetilde{\mathbf{Y}}) = \mathbf{A} + \mathbf{B}\mu_b$, where $\mu_b = \xi + \zeta\gamma$.

$$\text{Var}(\widetilde{\mathbf{Y}}) = \psi \mathbf{B} \mathbf{B}^T \left[\frac{1}{1 - 2\eta_b} - (\zeta\delta)^2 \right] + \boldsymbol{\Sigma}; \quad 0 < \eta_b, \eta_e < \frac{1}{2}, \quad (16)$$

where

$$\zeta = \frac{[\text{gam}((1 - \eta_b)/2\eta_b)/\text{gam}(1/2\eta_b)]}{\sqrt{\eta_b\pi}}. \quad (17)$$

Let $\widetilde{\mathbf{D}} = \widetilde{Y}_1 - \widetilde{Y}_2$ and $\mathbf{m} = (1, -1)^T$. As a result of Appendix C,

$$\widetilde{D}|U, V \sim SN_1\left(\mathbf{m}^T(\mathbf{A} + \mathbf{B}\xi), \frac{\mathbf{m}^T \widetilde{\boldsymbol{\Pi}}_V \mathbf{m}}{U}, \frac{\mathbf{m}^T \widetilde{\boldsymbol{\Pi}}_V^{(1/2)} \widetilde{\boldsymbol{\delta}}_V}{(\mathbf{m}^T \widetilde{\boldsymbol{\Gamma}}_V \mathbf{m})^{(1/2)}}\right), \quad (18)$$

where $\widetilde{\boldsymbol{\delta}}_V = \widetilde{\boldsymbol{\lambda}}_V / (\mathbf{I} + \widetilde{\boldsymbol{\lambda}}_V^T \widetilde{\boldsymbol{\lambda}}_V)^{(1/2)}$ and $\widetilde{\boldsymbol{\Gamma}}_V = \widetilde{\boldsymbol{\Pi}}_V - \widetilde{\boldsymbol{\Pi}}_V^{(1/2)} \widetilde{\boldsymbol{\delta}}_V \widetilde{\boldsymbol{\delta}}_V^T \widetilde{\boldsymbol{\Pi}}_V^{(1/2)}$.

Now, the STcT version of CCC can be defined from (10) as

$$\text{CCC} = \frac{2\beta_1\psi[(1/1 - 2\eta_b) - (\zeta\delta)^2]}{\{\mathbf{m}^T(\mathbf{A} + \mathbf{B}\xi)\}^2 + \{\psi[(1/1 - 2\eta_b) - (\zeta\delta)^2] + \sigma_1^2\} + \{\beta_1^2\psi[(1/1 - 2\eta_b) - (\zeta\delta)^2] + \sigma_2^2\}}, \quad (19)$$

where $0 < \eta_b, \eta_e < (1/2)$.

Next, using (11), the TDI for STcT-MEM can be defined as

$$P(|\bar{D}| \leq t) = \int_0^\infty \int_0^\infty \{F(t) - F(-t)\} f(u, v | (\eta_b, \eta_e)) du dv, \quad t > 0, \quad (20)$$

where F is the distribution function of the $\bar{D}|U, V$ and $f(u, v | (\eta_b, \eta_e))$ is the joint density of (U, V) appearing in (15).

3.2. CCC and TDI under SMEM. Under the SMEM (4),

$$\bar{\mathbf{Y}} = \begin{pmatrix} \bar{Y}_1 \\ \bar{Y}_2 \end{pmatrix} \sim N_2(\mathbf{A} + \mathbf{B}\xi, \bar{\mathbf{\Pi}}); \text{ where } \bar{\mathbf{\Pi}} = \psi\mathbf{B}\mathbf{B}^T + \mathbf{\Sigma}. \quad (21)$$

Next, the difference $\bar{D} = \bar{Y}_1 - \bar{Y}_2$ can be represented as

$$\bar{D} \sim N_1(\mathbf{m}^T(\mathbf{A} + \mathbf{B}\xi), \mathbf{m}^T\bar{\mathbf{\Pi}}\mathbf{m}) \text{ where } \mathbf{m} = (1, -1)^T. \quad (22)$$

Now, the SMEM version of CCC can be defined from (10) as

$$\text{CCC} = \frac{2\beta_1\psi}{[\xi - (\beta_0 + \beta_1\xi)]^2 + [\psi + \sigma_1^2] + [\beta_1^2\psi + \sigma_2^2]}. \quad (23)$$

Moreover, TDI under SMEM defined by (11) can be determined as

$$P(|\bar{D}| \leq t) = \Phi\left\{\frac{t - E(\bar{D})}{s d(\bar{D})}\right\} - \Phi\left\{\frac{-t - E(\bar{D})}{s d(\bar{D})}\right\}, \quad (24)$$

where $\Phi(\cdot)$ denotes the cumulative distribution function (CDF) of a standard normal distribution.

4. Simulation Study

In this section, the Monte-Carlo simulation study is conducted to examine the behaviour of the ML estimators using the ECM algorithm based on STcT-MEM and SMEM. We generate 500 datasets from the STcT-MEM with sample sizes $n = 25, 50, \text{ and } 100$. The skewness parameter in STcT-MEM is set as follows: $\lambda = 2.5, 5 \text{ and } 10$. Other parameters in STcT-MEM are set as follows: $\beta = (0.05, 0.9), \mu_b = -0.01, \log(\psi) = -3, \eta_b = 0.3, \eta_e = 0.1$, derived from the gold particles data set. Here, we considered the inverse of the degrees of freedom to enhance the inference process, and we kept them constant throughout the simulation to save computing time.

For each sample size, the variances of the measurement errors are picked from uniform distributions on $(0.004 \text{ and } 0.008)$ and $(0.001 \text{ and } 0.003)$, respectively, and then assumed as known values. Based on 500 random samples, we compute the ML estimators and their CPs with the nominal 95% confidence intervals through the ECM algorithm under

STcT-MEM. Then, we calculate the sample bias (BIAS), the standard deviation (SD), and the root mean squared error (RMSE) as assessments for the estimates based on STcT-MEM and SMEM under the simulated dataset generated by STcT-MEM. The results are summarized in Tables 1 and 2, respectively. The R programming language was used for all calculations [39].

Table 1 shows the ML estimates, asymptotic standard errors (SEs), and CPs of 95% confidence intervals. Concerning β_0 and β_1 , the CPs are fairly close to 99%, even for small, moderate, and large sample sizes. In the case of other parameters, most entries are close to 95%, and some even fall below 90% for moderate and large samples, and the values for the small samples are not accurate. However, the CPs increase when the sample size increases for all cases, and it can also be seen that the CPs have good performance when the skewness is moderate or heavy ($\lambda = 5$ or 10).

Table 2 shows that when the sample size and skewness (λ) increase, the values of BIAS, SD, and RMSE decrease, as expected. The BIAS, SD, and RMSE values under STcT-MEM are small for all cases, revealing the efficiency and accuracy of the ML estimates, and for all settings, the values of SD and RMSE are nearly equal in STcT-MEM. In SMEM, the biases of the estimates of μ_b and $\log(\psi)$ are not negligible. Moreover, for all cases, the bias, SD, and RMSE of the ML estimates based on STcT-MEM are smaller than those of the SMEM estimates. Thus, the performance of the ST distribution is better than that of the normal distribution, which may be due to their heavy-tailed characteristics.

For the model comparison, we compute their relative efficiency by dividing the MSE of SMEM by the MSE of STcT-MEM, and if it is greater than one, it means STcT-MEM is better. These values are displayed in Table 3, and it can be observed that the relative efficiencies increase with sample size in all situations. Moreover, the relative efficiencies improve as λ increases. Furthermore, all entries are greater than 1, which indicates that STcT-MEM is better than SMEM for skew and heavy-tailed data. Additionally, the results of the Akaike information criterion (AIC) and Bayesian information criterion (BIC) values based on the STcT-MEM and SMEM when the data generating model is STcT-MEM are displayed in Table 3. We find from Table 3 that AIC and BIC values under STcT-MEM are smaller than the standard model (SMEM). It means STcT-MEM performs better than SMEM for skewed data.

5. Data Analysis

This part considers gold particle data [25] as a numerical example. This data set investigates the concentration of the gold particles (in gt^{-1}) using Classical and Screen Fire Assay (SFA) methods. The measurements during this study are not replicated. There are $501 \times 2 = 1002$ observations, which are

TABLE 1: ML estimates, asymptotic standard errors (SEs), and coverage probabilities (CPs) of 95% confidence interval of STcT-MEM.

λ	n	Parameter	STcT-MEM		
			Estimates	SE	CP
2.5	25	β_0	0.051	0.022	0.998
		β_1	0.898	0.076	0.999
		μ_b	-0.007	0.080	0.840
		Log (ψ)	-3.024	0.696	0.866
		λ	2.491	2.452	0.840
	50	β_0	0.050	0.015	0.996
		β_1	0.900	0.048	0.994
		μ_b	-0.013	0.055	0.896
		Log (ψ)	-3.012	0.496	0.896
		λ	2.499	1.809	0.892
	100	β_0	0.050	0.010	0.990
		β_1	0.900	0.032	0.988
		μ_b	-0.012	0.038	0.905
		Log (ψ)	-3.004	0.342	0.900
		λ	2.492	1.266	0.918
5	25	β_0	0.050	0.023	0.996
		β_1	0.898	0.081	0.997
		μ_b	-0.008	0.053	0.808
		Log (ψ)	-3.013	0.551	0.924
		λ	4.955	5.922	0.842
	50	β_0	0.050	0.015	0.993
		β_1	0.900	0.050	0.991
		μ_b	-0.012	0.034	0.904
		Log (ψ)	-3.008	0.381	0.928
		λ	4.993	4.063	0.908
	100	β_0	0.050	0.011	0.987
		β_1	0.900	0.034	0.985
		μ_b	-0.011	0.025	0.958
		Log (ψ)	-3.004	0.268	0.944
		λ	4.996	2.952	0.936
10	25	β_0	0.049	0.024	0.996
		β_1	0.898	0.082	0.996
		μ_b	-0.009	0.037	0.964
		Log (ψ)	-3.008	0.494	0.980
		λ	9.965	10.723	0.872
	50	β_0	0.049	0.016	0.992
		β_1	0.899	0.050	0.989
		μ_b	-0.011	0.024	0.954
		Log (ψ)	-3.006	0.334	0.972
		λ	9.988	7.420	0.910
	100	β_0	0.049	0.011	0.986
		β_1	0.899	0.034	0.984
		μ_b	-0.010	0.017	0.948
		Log (ψ)	-3.005	0.236	0.970
		λ	9.985	5.788	0.944

paired and starting from 0.038 to 4.523 gt^{-1} . Since these measurements were made in a chemical lab, where a variety of variables, including the operator and the subject's location, may affect the outcomes, they are prone to mistakes, and these errors are inevitable. As a result, the proposed model is applicable. Furthermore, as required by our methodology, these measurement errors have known error variances. It was calculated using the formulas $\sigma_1^2 = 0.1568Y_{i1}^{0.5031}$ and $\sigma_2^2 = 0.0744Y_{i2}^{0.5047}$ [44].

Figure 1 shows the histograms and normal Q-Q plots of gold particle data. It is observed that the data are asymmetric and heavy-tailed. Figure 2 illustrates their trellis plot. It shows that the measures of both methods do not overlap, and the measures of the SFA often have the greatest ones. Some subjects have disproportionately big differences, implying a skewed distribution of disparities. Furthermore, it shows that the within-subject variations of both methods tend to increase with the concentration level. This means that the data are heteroscedastic.

Moreover, Figure 3 shows the scatter plot and Bland-Altman plot of these data. The scatter plot shows a modest correlation between the methods. The Bland-Altman plot shows that the vertical scatter appears to rise with average, which indicates heteroscedasticity. All the above plots show two extreme outliers. In this case, we performed the analysis by replacing the outlier with the mean value.

At the outset, we fit the mixed-effects model to the data. Figure 4 depicts the normal Q-Q plot of standardized residuals and random effects. The box plot and histogram of standardized residuals are also presented. These graphs appear skewness and heavy-tailed ness, suggesting that the assumption of normality is inadequate for error terms and random effects. Thus, we use the proposed STcT-MEM model to fit the data, where b_i follows the ST distribution and e_i follows a cT distribution. SMEM is also taken into account simultaneously for comparative purposes.

Firstly, we fit the STcT-MEM (9) by ML using the ECM algorithm, where Y_{i1} and Y_{i2} are the gold particles measurements taken from the Classical and SFA, respectively, on the i^{th} subject, $i = 1, 2, \dots, 501$. Here, the degrees of freedom (η_b, η_e) are assumed to be known parameters determined by the Schwarz information criteria [45], and the error variances ($\Sigma = \text{diag}(\sigma_1^2, \sigma_2^2)$) are also assumed as known. This model has five parameters. Secondly, we fit the SMEM model (4) when $\lambda = 0$ and $\eta_b, \eta_e \rightarrow 0$ by ML using the ECM algorithm, and it has four parameters.

Table 4 shows the parameter estimates, SEs, and 95% confidence intervals for these parameters. The 95% confidence intervals for intercept β_0 and slope β_1 are (0.051 and 0.062) and (0.835 and 0.905), respectively. In the SMEM, they are (0.057 and 0.067) and (0.816 and 0.874), respectively. Neither of the intercept intervals covers zero in both models, indicating considerable fixed biases. Likewise, neither of the slope intervals covers one, despite one being near the right border in both intervals. It provides evidence of moderately proportional biases.

The next step is to evaluate the agreement between the methods. Table 5 shows CCC and TDI (0.90) estimates and 95% one-sided confidence limits for both models discussed in section 3. The lower bound applies to CCC, and the upper bound applies to TDI. These measures are first computed using Fisher's z transformation of CCC and log transformation of TDI, and then the results are inverted back to the original scale. In STcT-MEM, the estimate of 0.940 and the lower bound of 0.932 for CCC imply a higher agreement between the methods. Further, the estimate and upper bound for TDI (0.90) are 0.123 and 0.135, respectively. The TDI bound reveals that 90% of the discrepancy between

TABLE 2: Simulated bias, sample standard deviation (SD), and root mean square error (RMSE) of ML for the STcT-MEM and SMEM.

λ	n	Parameter	STcT-MEM			SMEM		
			BIAS	SD	RMSE	BIAS	SD	RMSE
2.5	25	β_0	0.001	0.009	0.009	0.001	0.010	0.010
		β_1	-0.002	0.014	0.014	-0.005	0.030	0.031
		μ_b	0.003	0.007	0.008	-0.015	0.050	0.052
		Log (ψ)	-0.024	0.180	0.182	-0.123	0.622	0.634
		λ	-0.009	0.366	0.366	—	—	—
		β_0	0.000	0.004	0.004	-0.001	0.005	0.005
	50	β_1	0.000	0.007	0.007	0.000	0.014	0.014
		μ_b	-0.003	0.005	0.006	-0.015	0.035	0.038
		Log (ψ)	-0.012	0.083	0.084	-0.035	0.473	0.474
		λ	-0.001	0.197	0.197	—	—	—
		β_0	0.000	0.002	0.002	-0.001	0.003	0.003
		β_1	0.000	0.004	0.004	0.000	0.009	0.009
100	μ_b	-0.002	0.003	0.004	-0.014	0.024	0.028	
	Log (ψ)	-0.004	0.043	0.043	0.024	0.333	0.334	
	λ	-0.008	0.102	0.102	—	—	—	
	β_0	0.000	0.008	0.008	0.001	0.009	0.009	
	β_1	-0.002	0.012	0.013	-0.006	0.029	0.029	
	μ_b	0.002	0.003	0.004	-0.013	0.049	0.051	
5	25	Log (ψ)	-0.013	0.075	0.076	-0.129	0.653	0.666
		λ	-0.045	0.393	0.395	—	—	—
		β_0	0.000	0.004	0.004	0.000	0.005	0.005
		β_1	0.000	0.006	0.006	-0.001	0.013	0.013
		μ_b	-0.002	0.002	0.003	-0.013	0.033	0.035
		Log (ψ)	-0.008	0.020	0.021	-0.022	0.500	0.500
	50	λ	-0.007	0.167	0.168	—	—	—
		β_0	0.000	0.002	0.002	-0.001	0.003	0.003
		β_1	0.000	0.004	0.004	-0.001	0.008	0.008
		μ_b	-0.001	0.001	0.001	-0.012	0.023	0.026
		Log (ψ)	-0.004	0.009	0.010	0.018	0.353	0.353
		λ	-0.004	0.083	0.084	—	—	—
10	25	β_0	-0.001	0.008	0.008	0.001	0.009	0.009
		β_1	-0.002	0.012	0.013	-0.006	0.029	0.029
		μ_b	0.001	0.001	0.001	-0.010	0.049	0.051
		Log (ψ)	-0.008	0.014	0.017	-0.159	0.664	0.683
		λ	-0.035	0.264	0.266	—	—	—
		β_0	-0.001	0.004	0.004	0.000	0.005	0.005
	50	β_1	-0.001	0.006	0.006	-0.001	0.013	0.013
		μ_b	-0.001	0.001	0.001	-0.010	0.033	0.034
		Log (ψ)	-0.006	0.006	0.008	-0.019	0.510	0.510
		λ	-0.012	0.115	0.115	-	-	-
		β_0	-0.001	0.002	0.002	-0.001	0.003	0.003
		β_1	-0.001	0.004	0.004	-0.001	0.008	0.008
100	μ_b	-0.000	0.001	0.001	-0.010	0.023	0.025	
	Log (ψ)	-0.005	0.003	0.005	0.016	0.362	0.362	
	λ	-0.015	0.075	0.076	—	—	—	

Classical and SFA measurements is within ± 0.135 , with 95% confidence. Since the readings range between 0.03 and 3, the difference of 0.135 cannot be acceptable when the real value is 0.03, but it may be acceptable when the real value is 3. Thus, we can conclude that the tests exhibit good agreement for large values but not for small values. Focusing on the SMEM, the CCC estimate and lower bound are 0.938 and 0.930, respectively. Moreover, the estimate of TDI (0.90) is 0.174, and its 95% upper confidence bound is 0.180. Compared to the STcT counterpart, the CCC bound has not changed substantially, but the TDI bound has changed to 0.180. This value suggests that 90% of the differences in

measurements from the methods fall within ± 0.180 . Such differences are quite large compared to the STcT-MEM values. From this, we can conclude that STcT-MEM shows a satisfactory agreement between methods for large values than SMEM.

Additionally, we perform the LR test where the null hypothesis H_0 : SMEM model is preferable against the alternative hypothesis and H_1 : STcT-MEM model is preferable since both models are nested. The test statistic of LR is $2[l(\hat{\theta}) - l(\hat{\theta}_0)]$ under the null hypothesis that follows $\chi^2_{1,0.05}$ distribution, where $l(\hat{\theta})$ and $l(\hat{\theta}_0)$ are the log-likelihood functions assessed at ML estimates using the ECM algorithm

TABLE 3: AIC and BIC values while the STcT-MEM is the data generating model and relative efficiencies of STcT-MEM based estimators relative to the SMEM.

λ	n	STcT-MEM		SMEM		MSE _{SMEM} /MSE _{STcT-MEM}			Log (ψ)
		AIC	BIC	AIC	BIC	β_0	β_1	μ_b	
2.5	25	-54.674	-48.580	-50.669	-45.793	1.265	4.812	42.25	12.135
	50	-116.862	-107.302	-101.293	-93.645	1.624	4.029	40.111	31.842
	100	-240.606	-227.580	-204.249	-193.828	2.636	4.473	49.000	60.334
5	25	-59.361	-53.267	-53.186	-48.310	1.407	5.571	>100	76.793
	50	-127.384	-117.824	-105.669	-98.021	1.908	4.441	>100	>100
	100	-263.333	-250.307	-213.607	-203.186	2.848	4.703	>100	>100
10	25	-61.472	-55.378	-53.884	-49.008	1.536	5.276	>100	>100
	50	-132.193	-122.633	-106.934	-99.285	1.853	4.249	>100	>100
	100	-273.921	-260.895	-216.579	-206.158	2.201	3.949	>100	>100

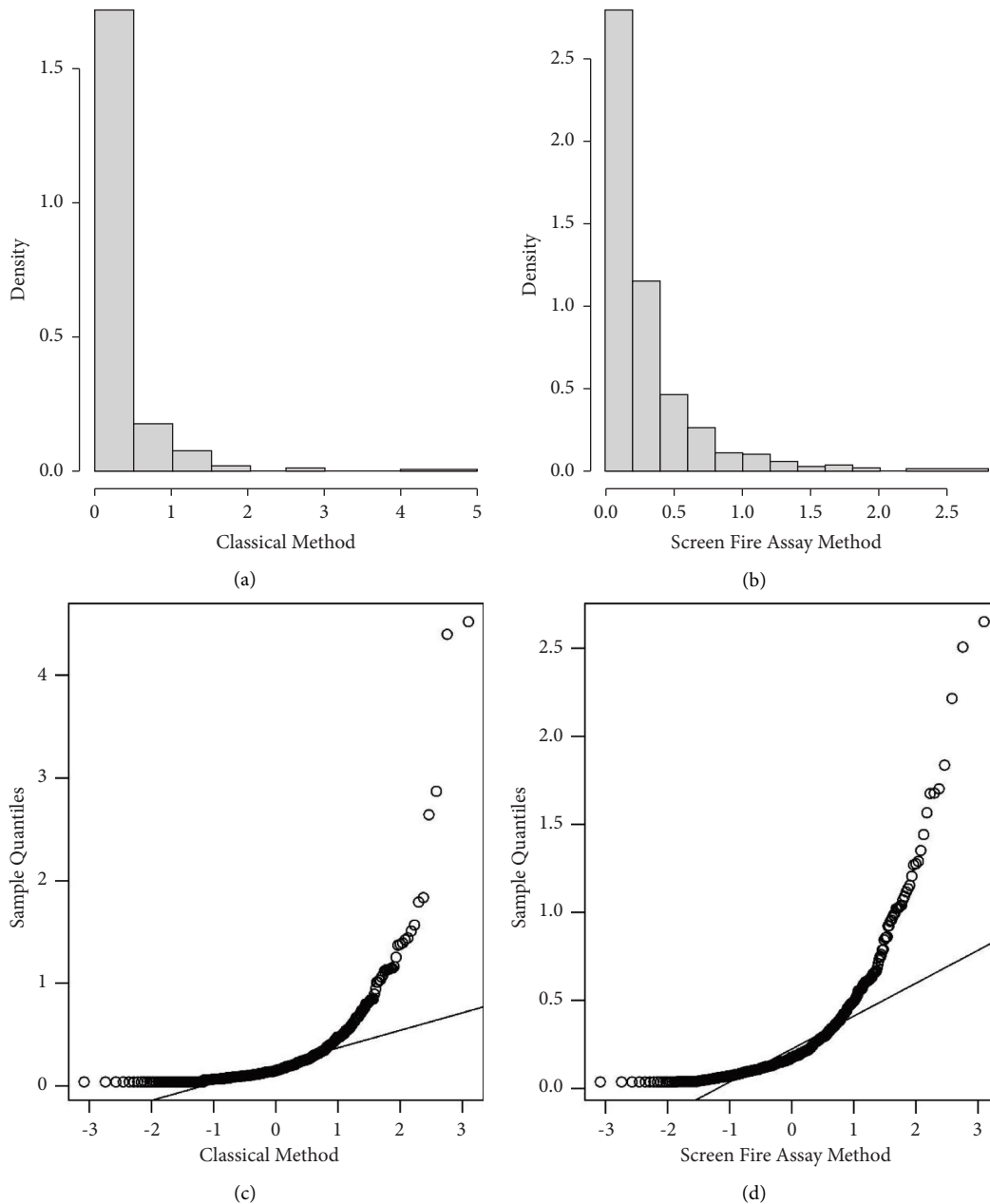


FIGURE 1: Histogram (a, b) and normal Q-Q plot (c, d) of the gold particle data taken by two measurement methods using the Classical method and the Screen Fire Assay method.

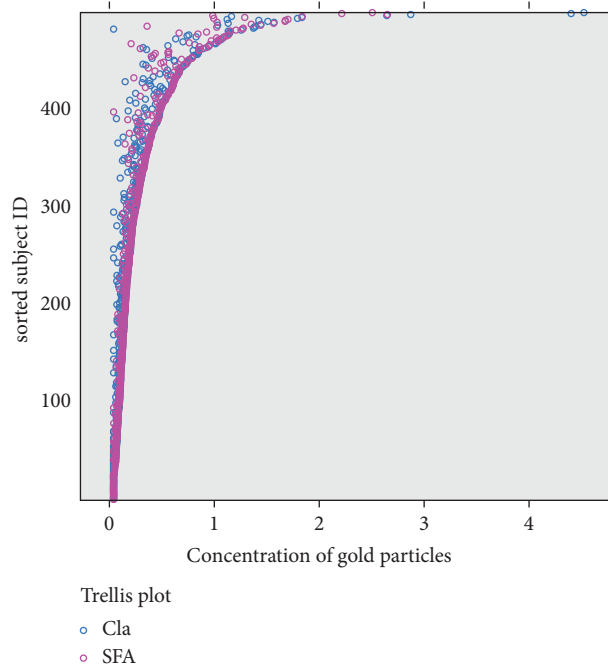


FIGURE 2: Trellis plot for gold particle data.

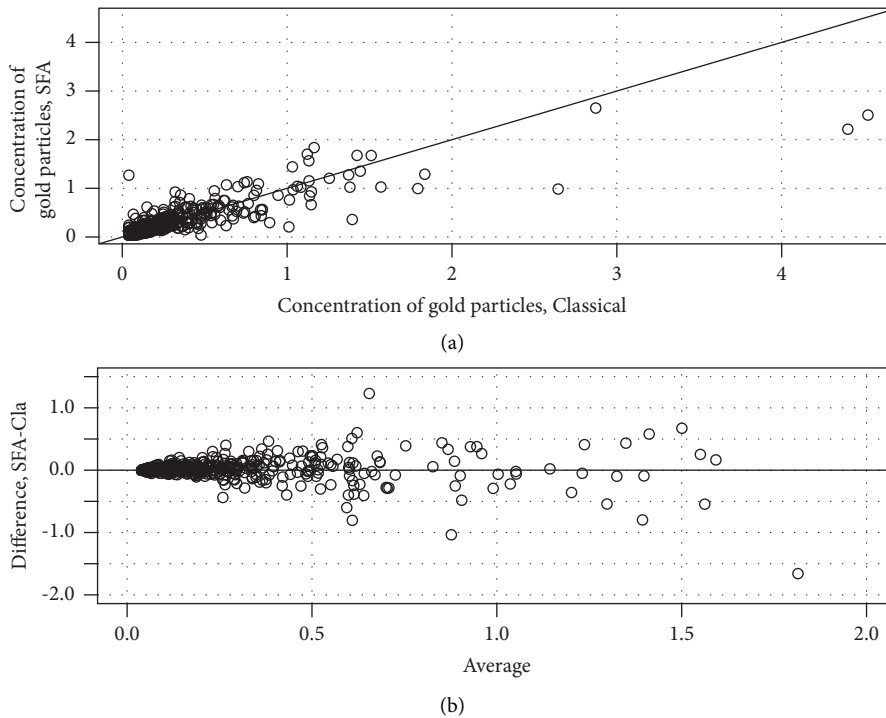


FIGURE 3: (a) A scatter plot with a line of equality and (b) Bland-Altman plot with zero line for gold particle data.

based on STcT-MEM and SMEM, respectively. The p value = 0 was obtained by applying the LR test, which is less than 0.05. Thus, STcT-MEM is better than SMEM for the

gold particle data. Furthermore, the AIC and BIC values are also included in Table 5, demonstrating that STcT-MEM performs well over the SMEM.

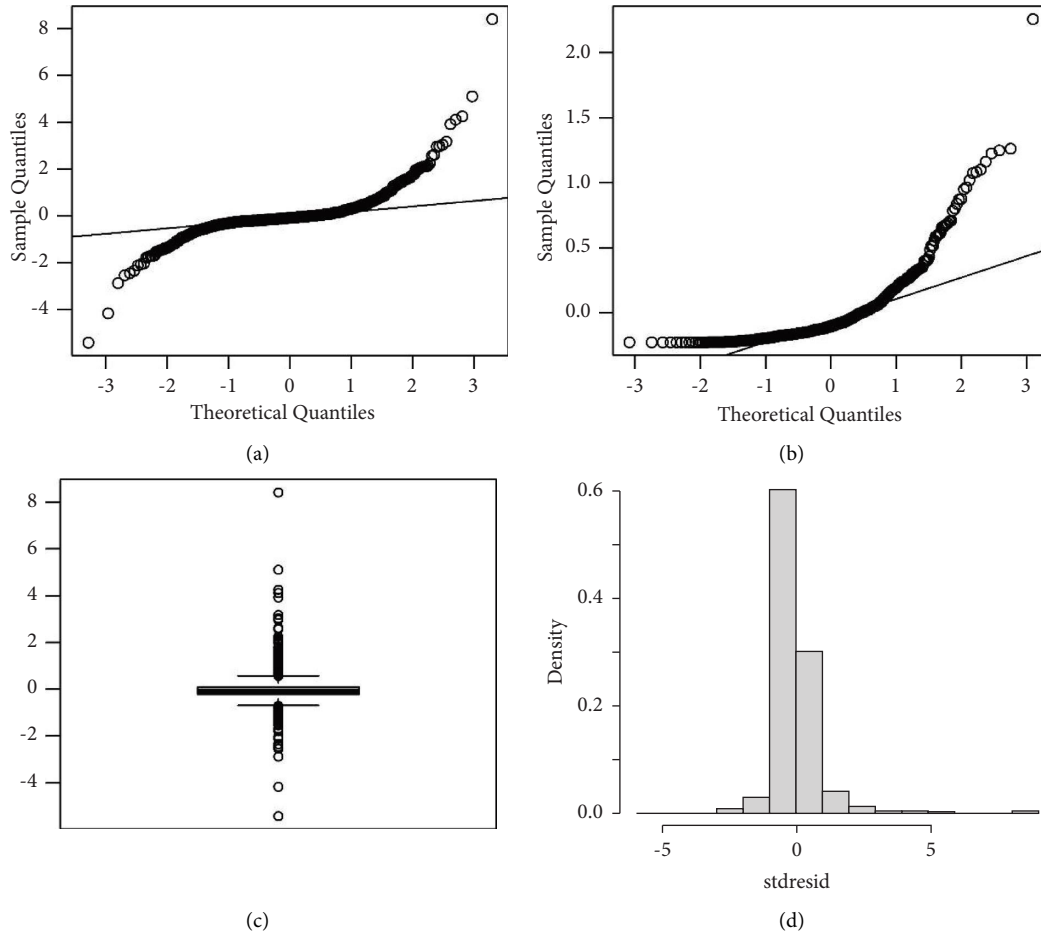


FIGURE 4: (a) Normal Q-Q plot of standardized residuals, (b) normal Q-Q plot of random effects, and (c, d) plots for box plot and histogram of standardized residuals.

TABLE 4: ML estimates, their standard errors (SEs), and 95% confidence intervals for parameters of STcT-MEM and SMEM for gold particle data.

Parameter	STcT-MEM				SMEM			
	Estimate	SE	Confidence interval		Estimate	SE	Confidence interval	
			LCL	UCL			LCL	UCL
β_0	0.057	0.003	0.051	0.062	0.062	0.003	0.057	0.067
β_1	0.870	0.018	0.835	0.905	0.845	0.015	0.816	0.874
μ_b	-0.010	0.006	-0.021	0.003	0.247	0.015	0.218	0.276
Log (ψ)	-2.725	0.107	-2.935	-2.515	-2.281	0.076	-2.430	-2.131
λ	5.538	0.692	4.182	6.893	—	—	—	—

TABLE 5: Measures of agreement and model selection criteria for gold particle data.

Models	CCC		TDI		AIC	BIC
	Estimate	95% lower bound	Estimate	95% upper bound		
STcT-MEM	0.940	0.932	0.123	0.135	-997.312	-972.763
SMEM	0.938	0.930	0.174	0.180	-275.061	-255.422

6. Conclusion

This article presents a methodology for method comparison data based on the distributions of ST and cT, called STcT-MEM, which provides excellent flexibility in considering asymmetry and heavy tails in the data. This model can also be used for normally distributed data. The ECM algorithm is performed to obtain the ML estimates of parameters. This also helped to adapt the SMEM mentioned in this article with some tweaks. The simulation results show that STcT-MEM-based ML estimates performed well in moderate and large sample sizes. We also demonstrated our approach using real data set and showed that the STcT-MEM model performed better than the SMEM. The improved model is expected to give satisfactory results for analyzing method comparison data for moderate and large samples in the presence of measurement errors, skewness, and heavy tails,

commonly found in many areas, especially in health-related fields. Our proposed methodology can be used only for unreplicated data. However, our methodology can be expanded to account for replicated measurements and multiple methods of measurement.

Appendix

A. Definition

A random vector \mathbf{Z} is said to follow the SN distribution, that is, $\mathbf{Z} \sim S\mathcal{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\lambda})$, if its density function is $f(\mathbf{z}; \boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\lambda}) = 2\phi_p(\mathbf{z}; \boldsymbol{\mu}, \boldsymbol{\Sigma})\Phi(\boldsymbol{\lambda}^T \mathbf{z}^*)$, $\mathbf{z} \in \mathbb{R}^p$, where $\mathbf{z}^* = \boldsymbol{\Sigma}^{-(1/2)}(\mathbf{z} - \boldsymbol{\mu})$, $\phi_p(\cdot; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ be the density function of $\mathcal{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ distribution, and $\Phi(\cdot)$ denote the CDF of a standard normal distribution.

A random vector \mathbf{Z} is said to follow the t distribution, that is, $\mathbf{Z} \sim t_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu)$, if its pdf is

$$f_t(\mathbf{z}; \boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu) = \frac{1}{|\boldsymbol{\Sigma}|^{(1/2)}} \frac{\text{gam}((\nu + p)/2)}{(\pi\nu)^{(p/2)} \text{gam}(\nu/2)} \left(1 + \frac{\mathbf{z}^{*T} \mathbf{z}^*}{\nu}\right)^{-((\nu+p)/2)}, \quad \mathbf{z} \in \mathbb{R}^p. \quad (\text{A.1})$$

A random vector \mathbf{Z} is said to follow the ST distribution, that is, $\mathbf{Z} \sim ST_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\lambda}, \nu)$, if its pdf is

$$f(\mathbf{z}; \boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\lambda}, \nu) = 2f_t(\mathbf{z}; \boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu)F_t\left(\boldsymbol{\lambda}^T \mathbf{z}^* \left\{ \left(\nu + p/\nu + \mathbf{z}^{*T} \mathbf{z}^* \right) \right\}^{(1/2)}; \nu + p\right), \quad \mathbf{z} \in \mathbb{R}^p, \quad (\text{A.2})$$

where $F_t(\cdot; \nu)$ denotes the CDF of the univariate Student's t distribution with ν degrees of freedom.

B. Hierarchical Representation for STcT-MEM

Consider \mathbf{Z}_i as defined in (3) where $b_i \sim ST(\xi, \psi, \lambda, \eta_b)$, $\mathbf{e}_i \sim ct_2(\mathbf{0}, \boldsymbol{\Sigma}_i, \eta_e)$.

A hierarchical representation for \mathbf{Z}_i is defined as follows:

$$\begin{aligned} \mathbf{Z}_i | U_i, V_i &\sim SN_2\left(\mathbf{A} + \mathbf{B}\xi, \frac{\boldsymbol{\Pi}_{V_i}}{U_i}, \boldsymbol{\lambda}_{V_i}\right), U_i \sim \mathcal{G}\left(\frac{1}{2\eta_b}, \frac{1}{2\eta_b}\right), \\ \frac{U_i}{V_i} &\sim \mathcal{G}\left(\frac{1}{2\eta_e}, \frac{1}{2c(\eta_e)}\right), \end{aligned} \quad (\text{B.1})$$

where $c(\eta_e) = \eta_e / (1 - 2\eta_e)$ and defined for $\nu_i > 0$.

$$\boldsymbol{\Pi}_{V_i} = \boldsymbol{\psi} \mathbf{B} \mathbf{B}^T + \nu_i \boldsymbol{\Sigma}_i,$$

$$\boldsymbol{\lambda}_{V_i} = \frac{\boldsymbol{\Pi}_{V_i}^{-(1/2)} \mathbf{B} \boldsymbol{\psi}^{(1/2)} \boldsymbol{\lambda}}{\left[1 + \boldsymbol{\lambda}^2 \boldsymbol{\psi}^{-1} (\boldsymbol{\psi}^{-1} + (\mathbf{B}^T \boldsymbol{\Sigma}_i^{-1} \mathbf{B} / \nu_i))^{-1}\right]^{(1/2)}}, \quad (\text{B.2})$$

and $\mathcal{G}(\alpha, \beta)$ is a gamma distribution with parameters $\alpha, \beta > 0$.

C. Linear Combination of Skew-Normals

Let $\mathbf{Y} \sim SN_q(\boldsymbol{\beta}, \boldsymbol{\psi}, \boldsymbol{\lambda})$, $\boldsymbol{\delta} = (\boldsymbol{\lambda} / (1 + \boldsymbol{\lambda}^T \boldsymbol{\lambda}))^{(1/2)}$, $\boldsymbol{\gamma} = \boldsymbol{\psi}^{(1/2)} \boldsymbol{\delta}$, and $\boldsymbol{\Gamma} = \boldsymbol{\psi} - \boldsymbol{\gamma} \boldsymbol{\gamma}^T$.

If $\mathbf{a} \in \mathbb{R}^q$ with at least one nonzero element, then

$$\mathbf{a}^T \mathbf{Y} \sim SN_1\left(\mathbf{a}^T \boldsymbol{\beta}, \mathbf{a}^T \boldsymbol{\psi} \mathbf{a}, \frac{\mathbf{a}^T \boldsymbol{\psi}^{(1/2)} \boldsymbol{\delta}}{(\mathbf{a}^T \boldsymbol{\Gamma} \mathbf{a})^{(1/2)}}\right). \quad (\text{C.1})$$

Data Availability

The dataset [25] used in the analysis is available from the author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this article.

Acknowledgments

The authors would also like to thank Professor L. C. Tomaya and Professor M. de Castro for providing the gold particle measurements dataset. This work was supported by the University of Peradeniya, Sri Lanka, under grant number URG/2022/59/S.

References

- [1] J. M. Bland and D. Altman, "Statistical methods for assessing agreement between two methods of clinical measurement," *The Lancet*, vol. 327, no. 8476, pp. 307–310, 1986.
- [2] J. Pinheiro and D. Bates, *Mixed-Effects Models in S and S-PLUS*, Springer, Berlin, Germany, 2000.
- [3] J. M. Bland and D. Altman, "Measuring agreement in method comparison studies," *Statistical Methods in Medical Research*, vol. 8, no. 2, pp. 135–160, 1999.
- [4] J. M. Bland and D. G. Altman, "Agreement between methods of measurement with multiple observations per individual," *Journal of Biopharmaceutical Statistics*, vol. 17, no. 4, pp. 571–582, 2007.
- [5] J. L. Carrasco and L. Jover, "Estimating the generalized concordance correlation coefficient through variance components," *Biometrics*, vol. 59, no. 4, pp. 849–858, 2003.
- [6] B. Carstensen, J. Simpson, and L. C. Gurrin, "Statistical models for assessing agreement in method comparison studies with replicate measurements," *International Journal of Biostatistics*, vol. 4, no. 1, p. 16, 2008.
- [7] J. L. Carrasco, T. S. King, and V. M. Chinchilli, "The concordance correlation coefficient for repeated measures estimated by variance components," *Journal of Biopharmaceutical Statistics*, vol. 19, no. 1, pp. 90–105, 2009.
- [8] A. S. Hedayat, C. Lou, and B. K. Sinha, "A Statistical approach to assessment of agreement involving multiple raters," *Communications in Statistics - Theory and Methods*, vol. 38, no. 16-17, pp. 2899–2922, 2009.
- [9] P. K. Choudhary, "A tolerance interval approach for assessment of agreement in method comparison studies with repeated measurements," *Journal of Statistical Planning and Inference*, vol. 138, no. 4, pp. 1102–1115, 2008.
- [10] D. M. Hawkins, "Diagnostics for conformity of paired quantitative measurements," *Statistics in Medicine*, vol. 21, no. 13, pp. 1913–1935, 2002.
- [11] H. X. Barnhart, Y. Lokhnygina, A. S. Kosinski, and M. Haber, "Comparison of concordance correlation coefficient and coefficient of individual agreement in assessing agreement," *Journal of Biopharmaceutical Statistics*, vol. 17, no. 4, pp. 721–738, 2007.
- [12] L. I. K. Lin, "A concordance correlation coefficient to evaluate reproducibility," *Biometrics*, vol. 45, no. 1, pp. 255–268, 1989.
- [13] L. S. Nawarathna and P. K. Choudhary, "Measuring agreement in method comparison studies with heteroscedastic measurements," *Statistics in Medicine*, vol. 32, no. 29, pp. 5156–5171, 2013.
- [14] L. S. Nawarathna and P. K. Choudhary, "A heteroscedastic measurement error model for method comparison data with replicate measurements," *Statistics in Medicine*, vol. 34, no. 7, pp. 1242–1258, 2015.
- [15] P. K. Choudhary, *Interrater agreement Methods and Applications of Statistics in the Life and Health Sciences* John Wiley, Hoboken, NJ, USA, 2009.
- [16] P. K. Choudhary and K. Yin, "Bayesian and frequentist methodologies for analyzing method comparison studies with multiple methods," *Statistics in Biopharmaceutical Research*, vol. 2, no. 1, pp. 122–132, 2010.
- [17] M. Haber and H. X. Barnhart, "A general approach to evaluating agreement between two observers or methods of measurement from quantitative data with replicated measurements," *Statistical Methods in Medical Research*, vol. 17, no. 2, pp. 151–169, 2007.
- [18] L. I. K. Lin, "Total deviation index for measuring individual agreement with applications in laboratory performance and bioequivalence," *Statistics in Medicine*, vol. 19, no. 2, pp. 255–270, 2000.
- [19] E. Alanen, "Everything all right in method comparison studies?" *Statistical Methods in Medical Research*, vol. 21, no. 4, pp. 297–309, 2010.
- [20] J. Ludbrook, "Confidence in Altman-Bland plots: a critical review of the method of differences," *Clinical and Experimental Pharmacology and Physiology*, vol. 37, no. 2, pp. 143–149, 2010.
- [21] G. Dunn, *Statistical Evaluation of Measurement Errors*, John Wiley & Sons, Hoboken, NJ, USA, 2009.
- [22] G. Dunn, "Regression models for method comparison data," *Journal of Biopharmaceutical Statistics*, vol. 17, no. 4, pp. 739–756, 2007.
- [23] G. Dunn and C. Roberts, "Modelling method comparison data," *Statistical Methods in Medical Research*, vol. 8, no. 2, pp. 161–179, 1999.
- [24] C.-L. Cheng and J. W. Van Ness, *Statistical Regression with Measurement Error*, John Wiley, Hoboken, NJ, USA, 1999.
- [25] L. C. Tomaya and M. de Castro, "A heteroscedastic measurement error model based on skew and heavy-tailed distributions with known error variances," *Journal of Statistical Computation and Simulation*, vol. 88, no. 11, pp. 2185–2200, 2018.
- [26] T. S. King and V. M. Chinchilli, "Robust estimators of the concordance correlation coefficient," *Journal of Biopharmaceutical Statistics*, vol. 11, no. 3, pp. 83–105, 2001.
- [27] T. S. King, V. M. Chinchilli, and J. L. Carrasco, "A repeated measures concordance correlation coefficient," *Statistics in Medicine*, vol. 26, no. 16, pp. 3095–3113, 2007.
- [28] P. K. Choudhary, "A Unified approach for nonparametric evaluation of agreement in method comparison studies," *International Journal of Biostatistics*, vol. 6, no. 1, p. 19, 2010.
- [29] H. X. Barnhart, M. Haber, and J. Song, "Overall concordance correlation coefficient for evaluating agreement among multiple observers," *Biometrics*, vol. 58, no. 4, pp. 1020–1027, 2002.
- [30] H. X. Barnhart, J. Song, and M. J. Haber, "Assessing intra, inter and total agreement with replicated readings," *Statistics in Medicine*, vol. 24, no. 9, pp. 1371–1384, 2005.
- [31] L. Lin, A. S. Hedayat, and W. Wu, "A unified approach for assessing agreement for continuous and categorical data," *Journal of Biopharmaceutical Statistics*, vol. 17, no. 4, pp. 629–652, 2007.

- [32] A. Azzalini, "A class of distributions which includes the normal ones," *Scandinavian Journal of Statistics*, vol. 12, no. 2, pp. 171–178, 1985.
- [33] A. Azzalini and A. Capitanio, "Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution," *Journal of the Royal Statistical Society: Series B*, vol. 65, no. 2, pp. 367–389, 2003.
- [34] K. R. Aravind and L. S. Nawarathna, "A statistical method for assessing agreement between two methods of heteroscedastic clinical measurements using the copula method," *Journal of Medical Statistics and Informatics*, vol. 5, no. 1, p. 3, 2017.
- [35] P. K. Choudhary, D. Sengupta, and P. Cassey, "A general skew-t mixed model that allows different degrees of freedom for random effects and error distributions," *Journal of Statistical Planning and Inference*, vol. 147, pp. 235–247, 2014.
- [36] D. Sengupta, P. K. Choudhary, and P. Cassey, *Modeling and analysis of method comparison data with skewness and heavy tails Ordered Data Analysis, Modeling and Health Research Methods*, pp. 169–187, Springer, Berlin, Germany, 2015.
- [37] C.-Z. Cao, Y. Wang, J. Q. Shi, and J. Lin, "Measurement error models for replicated data under asymmetric heavy-tailed distributions," *Computational Economics*, vol. 52, no. 2, pp. 531–553, 2017.
- [38] B. Sutradhar, "Score test for the covariance matrix of the elliptic t-distribution," *Journal of Multivariate Analysis*, vol. 46, no. 1, pp. 1–12, 1993.
- [39] R. C. Team, *Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2018.
- [40] A. Azzalini and A. Capitanio, "Statistical applications of the multivariate skew normal distribution," *Journal of the Royal Statistical Society: Series B*, vol. 61, no. 3, pp. 579–602, 1999.
- [41] P. Gilbert and R. Varadhan, "numDeriv: accurate numerical derivatives," *R package version*, vol. 1, pp. 1–8, 2019.
- [42] X. L. Meng and D. B. Rubin, "Maximum likelihood estimation via the ECM algorithm: a general framework," *Biometrika*, vol. 80, no. 2, pp. 267–278, 1993.
- [43] P. K. Choudhary and H. N. Nagaraja, *Measuring Agreement: Models, Methods, and Applications*, John Wiley & Sons, Hoboken, NJ, USA, 2017.
- [44] M. Galea-Rojas, M. V. de Castilho, H. Bolfarine, and M. de Castro, "Detection of analytical bias," *Analyst*, vol. 128, no. 8, pp. 1073–1081, 2003.
- [45] G. Schwarz, "Estimating the dimension of a model," *Annals of Statistics*, vol. 6, no. 2, pp. 461–464, 1978.